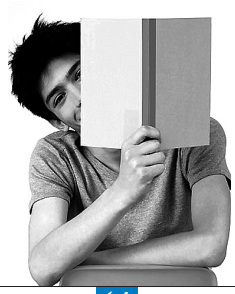


16

Scaling Procedures and Construct Validation of Context Questionnaire Data

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

OVERVIEW

The PISA 2006 context questionnaires included numerous items on student characteristics, student family background, student perceptions, school characteristics and perceptions of school principals. In 16 countries (optional) parent questionnaires were administered to the parents of the tested students.

Some of the items were designed to be used in analyses as single items (for example, gender). However, most questionnaire items were designed to be combined in some way so as to measure latent constructs that cannot be observed directly. For these items, transformations or scaling procedures are needed to construct meaningful indices.

This chapter describes how student, school and parent questionnaire indices were constructed and validated. As in previous PISA surveys, two different kinds of indices can be distinguished:

- Simple indices: These indices were constructed through the arithmetical transformation or recoding of one or more items;
- Scale indices: These indices were constructed through the scaling of items. Typically, scale scores for these indices are estimates of latent traits derived through IRT scaling of dichotomous or Likert-type items.

This chapter (i) outlines how simple indices were constructed, (ii) describes the methodology used for construct validation and scaling, (iii) details the construction and validation of scaled indices and (iv) illustrates the computation of the index on economic, social and cultural status (ESCS), including a discussion of some modifications from the PISA 2003 ESCS index. Some indices had already been used in previous PISA surveys and are constructed based on a similar scaling methodology (see Schulz, 2002; and OECD 2005). Most indices, however, were based on the elaboration of a questionnaire framework and are related to science as the major domain of the third PISA survey (see Chapter 3).

SIMPLE QUESTIONNAIRE INDICES

Student questionnaire 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 and the student tracking forms. If the month of testing was not know for a particular student, the median month of testing for that country was used in the calculation. The formula for computing *AGE* was

16.1

$$AGE = (100 + T_y - S_y) + \frac{(T_m - S_m)}{12}$$

where T_y and S_y are the year of the test and the year of the students' birth of the tested student, respectively in two-digit format (for example "06" or "92"), and T_m and S_m are the month of the test and month of the students' birth respectively. The result is rounded to two decimal places.

Study programme indices

PISA 2006 collected data on study programmes available to 15-year-old students in each country. This information was obtained through the student tracking form and the student questionnaire. In the final database, all national programmes will be included in a separate variable (*PROGN*) where the first three digits are the ISO code for a country, the next two digits are the sub-national category, and the last two digits

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

are the nationally specific programme code. All study programmes were classified using the international standard classification of education (ISCED) (OECD, 1999). The following indices are derived from the data on study programmes: programme level (*ISCDL*) indicating whether students are on the lower or upper secondary level (ISCDE 2 or ISCED 3); programme designation (*ISCEDD*) indicating 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*) indicating whether the programme's curricular content is general, pre-vocational or vocational.

Highest occupational status of parents

Occupational data for both the student's father and student's mother were obtained by asking open-ended questions. The response were coded to four-digit ISCO codes (ILO,1990) and then mapped to the international socio-economic index of occupational status (*ISEI*) (Ganzeboom *et al.*, 1992). Three indices were obtained from these scores: father's occupational status (*BFMJ*); mother's occupational status (*BMMJ*); 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.

Educational level of parents

Parental education is a second family background variable that is often used in the analysis of educational outcomes. Theoretically, it has been argued that parental education is a more relevant influence on a student's outcomes than is parental occupation. Like occupation, the collection of internationally comparable data on parental education poses significant challenges, and less work has been done on internationally comparable measures of educational outcomes than has been done on occupational status. The core difficulties with parental education relate to international comparability (education systems differ widely between countries and within countries over time), response validity (students are often unable to accurately report their parents' level of education) and, especially with increasing immigration, difficulties in the national mapping of parental qualifications gained abroad.

Parental education is classified using ISCED (OECD,1999). Indices on parental education are constructed by recoding educational qualifications into the following categories: (0) None; (1) ISCED 1 (primary education); (2) ISCED 2 (lower secondary); (3) ISCED Level 3B or 3C (vocational/pre-vocational upper secondary); (4) ISCED 3A (upper secondary) and/or ISCED 4 (non-tertiary post-secondary); (5) ISCED 5B (vocational tertiary); and (6) ISCED 5A, 6 (theoretically oriented tertiary and post-graduate). Indices with these categories were provided for the students' mother (*MISCED*) and the students' father (*FISCED*). In addition, the index on the highest educational level of parents (*HISCED*) corresponds to the higher ISCED level of either parent.

The index scores for highest educational level of parents were also recoded into estimated years of schooling (*PARED*). A mapping of ISCED levels of years of schooling is provided in Appendix 5.

Immigration background

As in PISA 2000 and PISA 2003, information on the country of birth of the students and their parents was collected. Included in the database are three country-specific variables relating to the country of birth of the student, mother, and father (*CTNUMS, CTNUMM,* and *CTNUMF*). Also, the items ST11Q01, ST11Q02 and ST11Q03 have been recoded for the database into the following categories: (1) country of birth is same as country of assessment, and (2) otherwise.

The index on immigrant background (*IMMIG*) is calculated from these variables, and has the following categories: (1) native students (those students who had at least one parent born in the country), (2) first-generation students (those students born outside the country of assessment and whose parents were also born in another country). and (3) second generation' students (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 have been given missing values for this variable.

Language spoken at home

Similar to PISA 2003, students also indicated what language they usually spoke at home, and the database includes a variable (*LANGN*) containing country-specific codes for each language. In addition, the item ST12Q01 has be recoded for the international database into the following categories: (1) language at home is same as the language of assessment for that student, (2) language at home is a national language of the country but the student was assessed in a different language, and (3) language at home is another (foreign) language.

Expected occupational status

As in PISA 2000 and 2003, students were asked to report their expected occupation at age 30 and a description of this job. The responses were coded to four-digit ISCO codes (ILO, 1990) and then mapped to the *ISEI* index (Ganzeboom *et al.*, 1992). Recoding of ISCO codes into *ISEI* index results in scores for the students' expected occupational status (*BSMJ*), where higher scores of *ISEI* indicate higher levels of expected occupational status.

Blue-collar/white-collar parental occupation

As in 2003, the ISCO codes of parents were recoded into 4 categories: (1) white collar high skilled, (2) white collar low skilled, (3) blue collar high skilled, and (4) blue collar low skilled. Three variables are included, one indicating the mother's employment category (*MSECATEG*), another indicating father's employment category (*FSECATEG*), and another indicating the highest employment category of either parent (*HSECATEG*).

Table 16.1
ISCO major group white-collar/blue-collar classification

ISCO Major Group	White-collar/blue-collar classification
1	White-collar high-skilled
2	White-collar high-skilled
3	White-collar high-skilled
4	White-collar low-skilled
5	White-collar low-skilled
6	Blue-collar high-skilled
7	Blue-collar high-skilled
8	Blue-collar low-skilled
9	Blue-collar low-skilled

Science-related occupations for parents and students

The ISCO data were used to compute four variables indicating whether or not the student expects to have a science-related career at age 30 (*SCIS5*), whether their mother (*SCIM1*) or father (*SCIM2*) are in a science career, or whether either or both parents are in a science related career (*SCIH12*). Values of 1 on these indicate "yes", while values of 0 indicate "no or undetermined".

To reduce the amount of missing data for parents' career status, parents with the following responses for occupations were recoded to "no/undetermined': home makers, social beneficiaries and students.

Furthermore, to reduce the amount of missing data on students' expected career status at age 30, students indicating "don't know" were recoded from missing to "no/undetermined". Also, students who responded to the items immediately subsequent to this question, but who did not respond to expected job at 30 were recoded to "no/undetermined'.

Since the ISCO coding scheme is rather broad for this purpose (*e.g.* some teaching professionals may be in a science-related career, but the scheme does not distinguish between teachers in different subject areas and disciplines), these science-related career variables should be interpreted as broad indicators rather than precise classifications. The ISCO occupation categories that were classified as science-related occupations are shown in Table 16.2.

Table 16.2
ISCO occupation categories classified as science-related occupations

ISCO Group Number	Occupation Category
1236	Computing services department managers
1237	Research and development department managers
211	Physicists, chemists and related professionals
2122	Statisticians
213	Computing professionals
214	Architects, engineers etc, professionals
221	Life science professionals
222	Health professionals except nursing
223	Nursing and midwifery professionals
2442	Sociologists, anthropologists etc, professionals
2445	Psychologists
2446	Social work professionals
311	Physical and engineering science associate professionals
313	Optical and electronic equipment operators
3143	Aircraft pilots etc, associate professionals
3144	Air traffic controllers
3145	Air traffic safety technicians
315	Safety and quality inspectors
321	Life science etc, associate professionals
322	Modern health professionals except nursing
323	Nursing and midwifery associate professionals

School questionnaire indices

School size

As in previous surveys, the PISA 2006 index of school size (*SCHLSIZE*) contains the total enrolment at school based on the enrolment data provided by the school principal, summing the number of girls and boys at a school.

Class size

The average class size (*CLSIZE*) is derived from one of nine possible categories, ranging from "15 students or fewer" to "More than 50 students'. *CLSIZE* takes the midpoint of each response category, a value of 13 for the lowest category, and a value of 53 for the highest.

Proportion of girls enrolled at school

As in previous surveys, the PISA 2006 index on the proportion of girls at school (*PCGIRLS*) is based on the enrolment data provided by the school principal, dividing the number of girls by the total of girls and boys at a school.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

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. As in previous PISA surveys, the index on school type (*SCHLTYPE*) has three categories: (1) public schools controlled and managed by a public education authority or agency, (2) government-dependent private schools controlled by a non-government organisation or with a governing board not selected by a government agency which receive more than 50% of their core funding from government agencies, (3) government-independent private schools controlled by a non-government organisation or with a governing board not selected by a government agency which receive less than 50% of their core funding from government agencies.[1]

Availability of computers

As in PISA 2000 and PISA 2003, school principals were asked to report the number of computers available at school. However, the question wording was modified for 2006 where principles were asked to report on the total number of computers, the number of computers available for instruction and the number of computers connected to the internet. The index of availability of computers (*RATCOMP*) is obtained by dividing the number of computers at school by the number of students at school. The overall ratio of computers to school size (*IRATCOMP*) was obtained by dividing the number of computers available for instruction at school by the number of students at school. The proportion of computers connected to the Internet (*COMPWEB*) was obtained by dividing the total number of computers connected to the Web by the total number of computers.

Quantity of teaching staff at school

As in previous PISA surveys, principles were asked to report the number of full-time and part-time teachers at school. However, the number of items was reduced in 2006 to capture only teachers in total, certified teachers, and teachers with an ISCED 5A qualification.

The student-teacher ratio (*STRATIO*) was obtained by dividing the school size by the total number of teachers. The number of part-time teachers is weighted by 0.5 and the number of full-time teachers is weighted by 1.0. The proportion of fully certified teachers (*PROPCERT*) was computed by dividing the number of fully certified teachers by the total number of teachers. The proportion of teachers who have an ISCED 5A qualification (*PROP5A*) was calculated by dividing the number of these kinds of teachers by the total number of teachers.

School selectivity

As in previous surveys, school principals were asked about admittance policies at their school. Among these policies, principles were asked how much consideration was given to the following factors when students are admitted to the school, based on a scale with the categories "not considered", "considered", "high priority", and "pre-requisite": students' academic record (including placement tests) and the recommendation of feeder schools.

An index of school selectivity (*SELECT*) was computed by assigning schools to four different categories: (1) schools where none of these factors is considered for student admittance; (2) schools considering at least one of these factors; (3) schools giving high priority to at least one of these factors; and (4) schools where at least one of these factors is a pre-requisite for student admittance.

Ability grouping

School principals were asked to report the extent to which their school organises instruction differently for student with different abilities. PISA 2003 included a similar question with two additional items which focused on mathematics classes. In 2006, this has been reduced to two items which ask about subject

grouping in a more general sense. One item asked about the occurrence of ability grouping into different classes and the other regarding ability grouping within classes (with the response categories "For all subjects", "For some subjects" and "Not for any subject").

An index of ability grouping between or within classes (*ABGROUP*) was derived from the two items by assigning schools to three categories: (1) schools with no ability grouping for any subjects, (2) schools with at least one of these forms of ability grouping for some subjects and (3) schools with at least one of these two forms of ability grouping for all subjects.

School responsibility for resource allocation

An index of the relative level of responsibility of school staff in allocating resources (*RESPRES*) was derived from six items measuring the school principals' report on who has considerable responsibility for tasks regarding school management of resource allocation ("Selecting teachers for hire", "Firing teachers", "Establishing teachers' starting salaries", "Determining teachers' salaries increases", "Formulating the school budget", "Deciding on budget allocations within the school"). The index was calculated on the basis of the ratio of "yes" responses for principal or teachers to "yes" responses for central educational authority. Higher values on the scale indicate relatively higher levels of school responsibility in this area. The index was standardised to having an OECD mean of 0 and a standard deviation of 1 (for the pooled data with equally weighted country samples).[2]

School responsibility for curriculum and assessment

An index of the relative level of responsibility of school staff in issues relating to curriculum and assessment (*RESPCURR*) was computed from four items measuring the school principal's report concerning who had responsibility for curriculum and assessment ("Establishing student assessment policies", "Choosing which textbooks are used", "Determining course content", "Deciding which courses are offered"). The index was calculated on the basis of the ratio of "yes" responses for principal or teachers to "yes" responses for central education authorities. Higher values indicate relatively higher levels of school responsibility in this area. The index was standardised to having an OECD mean of zero and a standard deviation of one (for the pooled data with equally weighted country samples).[3]

Parent questionnaire indices

Educational level of parents

Administration of this instrument in PISA 2006 provided the opportunity to collect data on parental education directly from the parents in addition to the data provided by the student questionnaire. Similar to the student questionnaire data, parental education were classified using ISCED (OECD 1999). The question format differed from the one used in the student questionnaire as only four items were included with dichotomous response categories of Yes or No.

Indices were constructed by taking the highest level for father and mother and having the following categories: (0) None, (1) ISCED 3A (upper secondary) and/or ISCED 4 (non-tertiary post-secondary), (2) ISCED 5B (vocational tertiary), (3) ISCED 5A, 6 (theoretically oriented tertiary and post-graduate). Indices with these categories were computed for mother (*PQMISCED*) and father (*PQFISCED*). Highest Educational Level of Parents (*PQHISCED*) corresponds to the higher ISCED level of either parent.

Occupational status of parents

Occupational data for both the student's father and student's mother were obtained by asking open-ended questions in a manner similar to the questions asked of students. The responses were coded to four-digit

ISCO codes (ILO, 1990) and then mapped to the SEI index (Ganzeboom, de Graaf & Treiman, 1992). Three SEI indices were computed from these scores.

Recoding of ISCO codes into SEI gives scores for the Mother's occupational status (*PQBMMJ*) and Father's occupational status (*PQBFMJ*). The highest occupational level of parents (*PQHISEI*) is the higher SEI score of either parent or to the only available parent's SEI score. Higher scores of SEI will indicate higher level of occupational status.

Similar to the science-related career variables derived from the student questionnaire, three indicators were derived from the parent data: whether the mother (*SCIM3*) or father (*SCIF4*) is in a science-related career, and whether either or both of the parents is in a science-related career (*SCIH34*).

SCALING METHODOLOGY AND CONSTRUCT VALIDATION

Scaling procedures

Most questionnaire items were scaled using IRT scaling methodology. With the One-Parameter (Rasch) model (Rasch 1960) for dichotomous items, the probability of selecting category 1 instead of 0 is modelled as

16.2

$$P_i(\theta) = \frac{\exp(\theta_n - \delta_i)}{1 + \exp(\theta_n - \delta_i)}$$

where $P_i(\theta)$ is the probability of person *n* to score 1 on item *i*. θ_n is the estimated latent trait of person *n* and δ_i the estimated location of item *i* on this dimension. For each item, item responses are modelled as a function of the latent trait θ_n.

In the case of items with more than two (k) categories (as for example with Likert-type items) this model can be generalised to the Partial credit model (Masters and Wright, 1997), which takes the form of

16.3

$$P_{x_i}(\theta) = \frac{\exp \sum_{k=0}^{x} (\theta_n - \delta_i + \tau_{ij})}{\sum_{h=0}^{m_i} \exp \sum_{k=0}^{h} (\theta_n - \delta_i + \tau_{ik})} \quad x_i = 0, 1, \ldots, m_i$$

where $P_{xi}(\theta)$ denotes the probability of person *n* to score *x* on item *i*. θ_n denotes the person's latent trait, the item parameter δ_i gives the location of the item on the latent continuum and τ_{ij} denotes an additional step parameter.

Item fit was assessed using the weighted mean-square statistic (infit), which is a residual based fit statistic. Weighted infit statistics were reviewed both for item and step parameters. The ACER *ConQuest*® software (Wu, Adams and Wilson, 1997) was used for the estimation of item parameters and the analysis of item fit.

International item parameters were obtained from calibration samples consisting of randomly selected sub-samples:

- For the calibration of student item parameters, sub-samples of 500 students were randomly selected within each OECD country sample. As final student weights had not been available at the time the calibration sample was drawn, the random selection was based on preliminary student weights obtained from the ratio between sampled and enrolled student within explicit sampling strata. The final calibration sample included data from 15,000 students;

- For the calibration of school item parameters, 100 schools were randomly selected within each OECD country sample. The random selection was based on school level weights in order to ensure that a representative sample of schools was selected from each country. School data from Luxembourg were not included due to of the small number of schools. Data from France were not available because the school questionnaire was not administered in France. The final calibration sample included data from 2 800 school principals.

Once international item parameter had been estimated from the calibration sample, weighted likelihood estimation was used to obtain individual student scores. WLEs can be computed by minimising the equation

16.4

$$\sum_{i\in\Omega}\left[\left(r_x+\frac{J_n}{2I_n}\right)-\sum_{j=1}^{k}\frac{\exp\left(\sum_{j=0}^{x}\theta_n-\delta_i+\tau_{ij}\right)}{\sum_{h=0}^{m_i}\exp\sum_{k=1}^{k}\left(\theta_n-\delta_i+\tau_{ij}\right)}\right]=0$$

for each case n, where r_x is the sum score obtained from a set of k items with j categories. This can be achieved by applying the Newton-Raphson method. The term $J_n/2I_n$ (with I_n being the information function for student n and J_n being its derivative with respect to θ) is used as a weight function to account for the bias inherent to maximum likelihood estimation (*see* Warm, 1989). IRT scores were derived using ACER *ConQuest*® with pre-calibrated item parameters.

Table 16.3
OECD means and standard deviations of WL estimates

Student-level indices	Mean	Standard deviation
CARINFO	–0.14	2.12
CARPREP	1.33	2.19
CULTPOSS	0.30	1.64
ENVAWARE	0.30	1.39
ENVOPT	–0.92	1.39
ENVPERC	1.77	1.42
GENSCIE	1.65	1.65
HEDRES	2.67	1.52
HIGHCONF	1.33	1.36
HOMEPOS	1.57	1.11
INSTSCIE	0.65	3.19
INTCONF	2.52	1.29
INTSCIE	–0.09	1.35
INTUSE	0.24	0.88
JOYSCIE	0.42	3.29
PERSCIE	0.58	1.80
PRGUSE	–0.53	1.04
RESPDEV	1.52	1.45
SCAPPLY	–0.20	1.63
SCHANDS	–0.73	1.64
SCIEACT	–2.04	1.68
SCIEEFF	0.45	1.31
SCIEFUT	–1.52	3.16
SCINTACT	–0.07	1.56
SCINVEST	–1.58	1.64
SCSCIE	0.23	3.04
WEALTH	1.28	1.46
School-level indices		
ENVLRN	–1.87	1.54
SCIPROM	0.95	1.53
SCMATEDU	0.24	1.55
TCSHORT	0.62	1.40

Note: Means and standard deviations for equally weighted OECD data.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

WLEs were transformed to an international metric with an OECD average of zero and an OECD standard deviation of one. The transformation was achieved by applying the formula

16.5

$$\theta'_n = \frac{\theta_n - \bar{\theta}_{OECD}}{\sigma_{\theta(OECD)}}$$

where θ'_n are the scores in the international metric, θ_n the original WLE in logits, and $\bar{\theta}_{OECD}$ is the OECD mean of logit scores with equally weighted country sub-samples. $\sigma_{\theta(OECD)}$ is the corresponding OECD standard deviation of the original WL estimates. Means and standard deviations used for the transformation into the international metric are shown in Table 16.3.

Construct validation

As in previous PISA surveys, it was important to develop comparable measures of student background, attitudes and perceptions. There are different methodological approaches for validating questionnaire constructs, each with their advantages, limitations and problems. Cross-country validity of these constructs is of particular importance as measures derived from questionnaires are often used to explain differences in student performance within and across countries and are, thus, potential sources of policy-relevant information about ways of improving educational systems.

Cross-country validity of the constructs not only requires a thorough and closely monitored process of translation into different languages. It also makes assumptions about having measured similar characteristics, attitudes and perceptions in different national and cultural contexts. Psychometric techniques can be used to analyse the extent to which constructs have (1) consistent dimensionality and (2) consistent construct validity across participating countries. This means that, once the measurement stability for each scale is confirmed, the multidimensional relationship between these constructs should be reviewed as well (see Wilson, 1994; Schulz 2006a; Walker 2006). It should be noted, however, that between-country differences in the strength of relationships between constructs do not necessarily indicate a lack of consistency as they may be due to differences between national contexts (for example, different educational systems or learning practices).

Confirmatory factor analysis

Structural Equation Modelling (SEM) was used to confirm theoretically expected dimensions and, if necessary, to re-specify the dimensional structure (Kaplan, 2000). Using Confirmatory Factor Analysis (CFA) requires a theoretical model of item dimensionality, which can be tested using the collected data.

Fit indices measure the extent to which a model based on the a-priori structure as postulated by the researcher fits the data. In the PISA 2006 analysis, model fit was assessed using the root-mean square error of approximation (RMSEA), the root mean square residual (RMR), the comparative fit index (CFI) and the non-normed fit index (NNFI) (see Bollen and Long, 1993). RMSEA values over 0.10 are usually interpreted as a sign of unacceptable model fit whereas values below 0.05 indicate a close model fit. RMR values should be less than 0.05. Both CFI and NNFI are bound between 0 and 1 and values between 0.90 and 0.95 indicate an acceptable model fit, with values greater than 0.95 indicating a close model fit.

For the results presented in this chapter, maximum likelihood estimation and covariance matrices were used for the analyses of the (categorical) Likert-type items, that is, the items were treated as if they were continuous. Confirmatory factor analyses of student data were based on the international calibration sample in order to have comparable (sub-)sample sizes across OECD countries. For the comparative analysis of item dimensionality the use of random OECD sub-samples was deemed appropriate.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

The SAS® CALIS procedure and the *LISREL* program were used to estimate the models based on Likert-type items. In order to assess cross-country validity of item dimensionality and constructs models were estimated both for the pooled OECD calibration sample (with 500 students per country) and for each country calibration sub-sample separately. CFA were carried out only for the student questionnaire data.

In the case of dichotomous items, weighted least squares (WLS) estimation with polychoric correlations was used (see Jöreskog and Sörbom, 1993). As the unadjusted WLS estimator requires very large sample sizes, a mean- and variance- adjusted WLS estimator (WLSMV) was used, which is available in the *Mplus* software program, (see Muthén, du Toit, and Spisic, 1997). Confirmatory factor analyses for dichotomous student-level items were only estimated for the pooled international calibration sample.

Between-school variance of student-level indices

The structure of the national PISA samples includes students that are nested within schools. Consequently, the variation in variables collected from students can either be between or within schools. Analyses of cognitive data tend to show that depending on the structure of educational systems in some countries a considerable amount of variation is found between schools.

Table 16.4 shows the median, maximum and minimum percentages of between-school variance for student questionnaire indices. For most of the student-level indices the average proportion of between-school variance is below 10%. However, for some indices there is a considerable variance between schools. Notably, home background indices like *WEALTH*, or *CULTPOSS* have relatively high intra-class correlations in many countries.

Table 16.4
Median, minimum and maximum percentages of between-school variance for student-level indices across countries

	OECD countries			Partner countries and economies		
Index	Median	Max	Min	Median	Max	Min
CULTPOSS	11	24	3	11	21	5
WEALTH	10	40	5	20	44	0
ENVAWARE	8	16	2	10	19	4
SCINVEST	8	17	3	9	22	2
HEDRES	8	30	2	16	39	3
INTCONF	7	24	2	16	48	6
SCIEEFF	6	14	1	6	13	4
SCINTACT	6	10	2	6	11	2
SCHANDS	6	10	2	6	11	2
CARPREP	6	18	1	4	11	2
INTSCIE	5	13	2	5	14	2
JOYSCIE	5	13	2	6	16	2
SCAPPLY	5	13	1	6	13	0
INTUSE	5	12	1	7	22	0
INSTSCIE	5	14	1	6	12	0
SCIEFUT	5	12	0	6	18	1
SCSCIE	4	14	2	4	12	0
GENSCIE	4	9	2	4	8	2
SCIEACT	4	8	1	6	24	3
RESPDEV	4	11	2	4	8	1
PRGUSE	4	10	2	4	12	0
PERSCIE	3	8	1	4	9	1
CARINFO	3	12	0	4	9	1
HIGHCONF	3	9	2	8	35	0
ENVOPT	3	9	0	5	17	1
ENVPERC	2	11	0	3	9	0

Note: Results from multi-level analysis with random intercepts only.

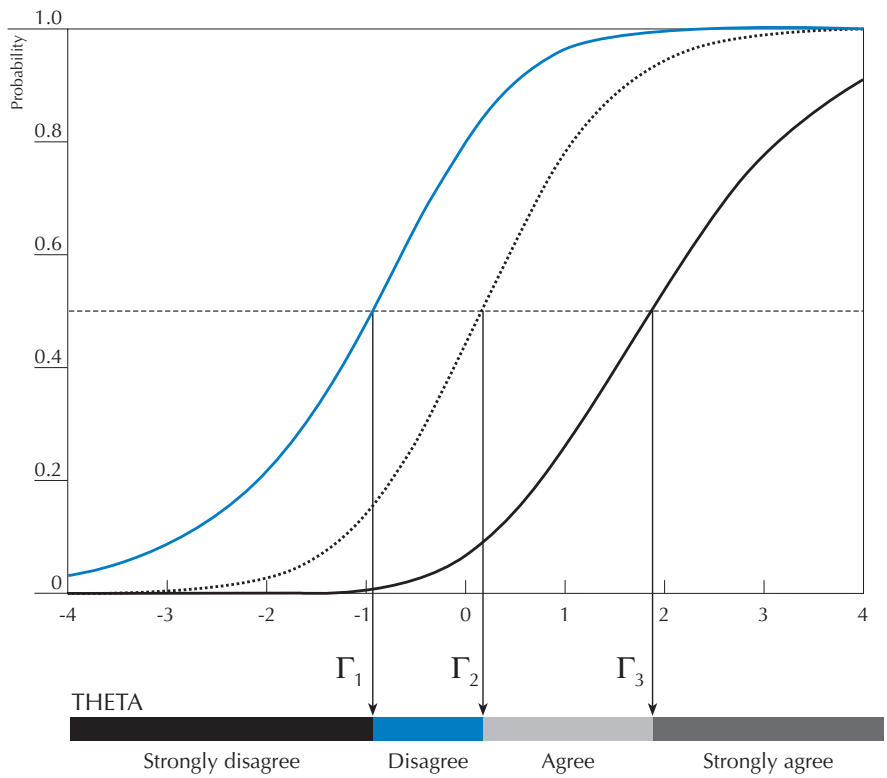
Describing questionnaire scale indices

As in previous PISA surveys, in PISA 2006 categorical items from the context questionnaires were scaled using IRT modelling. Weighted likelihood estimates (logits) for the latent dimensions were transformed to scales with an OECD average of 0 and a standard deviation of 1 (with equally weighted samples). It is possible to interpret these scores by comparing individual scores or group average scores to the OECD mean, but the individual scores do not reveal anything about the actual item responses and it is impossible to determine from scale score values to what extent respondents endorsed the items used for the measurement of the latent variable. However, the scaling model used to derive individual scores allows descriptions of these scales by mapping scale scores to (expected) item responses.[4]

Item characteristics can be described using the parameters of the partial credit model by summing for each category its probability of being chosen with the probabilities of all higher categories. This is equivalent to computing the odds of scoring higher than a particular category.

The results of plotting these cumulative probabilities against scale scores for a fictitious item are displayed in Figure 16.1. The three vertical lines denote those points on the latent continuum where it becomes more likely to score >0, >1 or >2. These locations Γ_k are Thurstonian thresholds that can be obtained through an iterative procedure that calculates summed probabilities for each category at each (decimal) point on the latent variable.

Figure 16.1
Summed category probabilities for fictitious item

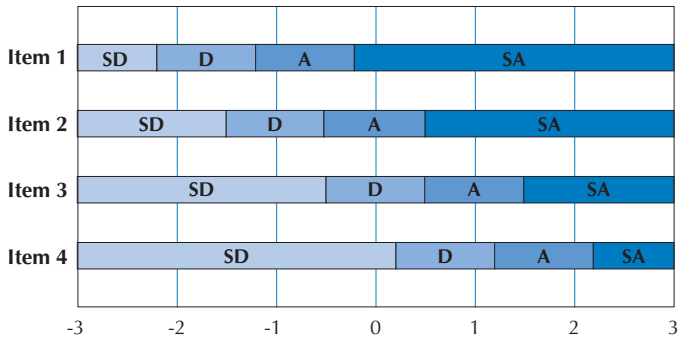


PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Summed probabilities are not identical with expected item scores and have to be understood in terms of the probability to score at least a particular category. Other ways of describing the item characteristics based on the partial credit model are item characteristic curves (by plotting the individual category probabilities) and expected item score curves (for a more detailed description see Masters and Wright, 1997).

Thurstonian thresholds can be used to indicate for each item category those points on a scale, at which respondents have a .5 probability to score this category or higher. For example, in the case of Likert-type items with categories "Strongly disagree" (SD), "Disagree" (D), "Agree" (A) and "Strongly agree" (SA) it is possible to determine at what point of a scale a respondent has 50% chance to agree with the item.

Figure 16.2
Fictitious example of an item map



The fictitious example in Figure 16.2 illustrates the interpretation of an item map for a fictitious scale with four different Likert-type items:

- Students with a score of –2 (that is, 2 standard deviations below the OECD average) have a 0.5 probability to disagree, agree or strongly agree (or not to disagree strongly with item 1), but they have more than a 50% chance to strongly disagree with the other three items;
- Students with a score of 1 (one standard deviation below the OECD average), have already more than 0.5 probability to agree with the first item, but they would still be expected to disagree with item 2 or even to strongly disagree with item 3 and 4;
- Likewise, students with a score 1 (one standard deviation above the OECD average) would have more than a 0.5 probability to strongly agree with the first two items, but still have less than 0.5 probability to agree with item 4.

Item maps can help to illustrate the relationship between scores and item responses. For example, even scores of one standard deviation below the OECD average on an attitudinal scale could still indicate affirmative responses. This would not be revealed by the international metric, which have to be interpreted relative to the OECD average, but can be concluded from the corresponding item map.

QUESTIONNAIRE SCALE INDICES

Student scale indices

Household possessions

Collecting household possessions as indicators of family wealth has received much attention in international studies in the field of education (Buchmann, 2000). Household assets are believed to capture wealth better than income because they reflect a more stable source of wealth.

In PISA 2006, students reported the availability of 13 different household items at home. In addition, countries added three specific household items that were seen as appropriate measures of family wealth within the country's context. Appendix 6 includes a list of the country-specific household items.

Four different indices were derived from these items: (i) family wealth possessions (*WEALTH*), (ii) cultural possessions (*CULTPOSS*), (iii) home educational resources (*HEDRES*) and (iiii) home possessions (*HOMEPOS*). The last index is a summary index of all household items and also included the variable indicating the number of books at home, but recoded into three categories: (0) 0-25 books, (1) 26-100 books, and (2) 101 or more books. *HOMEPOS* was also one of three components in the construction of the index on economic, social and cultural status (ESCS, see the section on ESCS index construction below). Table 16.5 shows the wording of items and their allocation to the four indices.

A confirmatory factor analysis using polychoric correlations with a WLSMV estimator showed a reasonable model fit for the international calibration sample of OECD countries (RMSEA = 0.080, CFI = 0.88, NNFI = 0.92). The estimated latent correlations between these constructs were 0.80 between *WEALTH* and *HEDRES*, 0.25 between *WEALTH* and *CULTPOSS*, and 0.52 between *CULTPOSS* and *HEDRES*.[5]

Analysis of differential item functioning (DIF) showed a considerable amount of between-country variation in the item parameters. It was decided to use nationally defined item parameters for scaling instead of using parameters estimated for the combined OECD sample (as done in previous cycles).

Table 16.5
Household possessions and home background indices

		Item is used to measure index			
Item		WEALTH	CULTPOSS	HEDRES	HOMEPOS
ST13	In your home, do you have:				
ST13Q01	A desk to study at			X	X
ST13Q02	A room of your own	X			X
ST13Q03	A quiet place to study			X	X
ST13Q04	A computer you can use for school work			X	
ST13Q05	Educational software			X	X
ST13Q06	A link to the Internet	X			X
ST13Q07	Your own calculator			X	X
ST13Q08	Classic literature (*e.g.* <Shakespeare>)		X		X
ST13Q09	Books of poetry		X		X
ST13Q10	Works of art (*e.g.* paintings)		X		X
ST13Q11	Books to help with your school work			X	X
ST13Q12	A dictionary			X	X
ST13Q13	A dishwasher (country-specific)	X			X
ST13Q14	A <DVD or VCR> player (country-specific)	X			X
ST13Q15	<Country-specific wealth item 1>	X			X
ST13Q16	<Country-specific wealth item 2>	X			X
ST13Q17	<Country-specific wealth item 3>	X			X
ST14	How many of these are there at your home?				
ST14Q01	Cellular phones	X			X
ST14Q02	Televisions	X			X
ST14Q03	Computers	X			X
ST14Q04	Cars	X			X
ST15	How many books are there in your home				X

Note: Item categories were "yes" (1) and "no" (2) for ST13, "None", "One", "Two" and "Three or more" for ST14, The categories for ST15 ("0-10 books", "11-25 books", "26-100 books", 101-200 books", "201-500 books" and "More than 500 books") were recoded into three categories ("0-25 books", 26-100 books" and "More than 100 books" ; Items in ST13 for were inverted for scaling and the first two categories of ST14Q01 and ST14Q02 were collapsed into one for scaling.

The *WEALTH* and *HOMEPOS* scales were constructed in two stages. A basket of common items was chosen (ST13Q02, ST13Q06, ST14Q01, ST14Q02, ST14Q03 and ST14Q04 for *WEALTH*, and in addition to these ST13Q01, ST13Q03, ST13Q05 to ST13Q12 and ST15Q01 for *HOMEPOS*) and item parameters were estimated for each country based on this item set. The sum of the set's item parameters was constrained to zero for each country. Next, these item parameters were anchored. The remaining country-specific items were added, and each country was scaled separately.

The other two scales derived from household possession items, *CULTPOSS* and *HEDRES*, were scaled in one step but the item parameters were allowed to vary by country.

Table 16.6 shows the scale reliabilities in OECD countries for all four scales, Table 16.7 those in partner countries. *HEDRES* has notably lower scale reliabilities when compared with the three indices. Similar results were already found for this index in PISA 2000 (see Schulz, 2002, p. 214) and PISA 2003 (see OECD, 2005, p. 284).

When comparing OECD and partner countries it appears that scale reliabilities for *WEALTH*, *HEDRES* and *HOMEPOS* are generally higher in partner countries. This may be due to the higher degree of accessibility of household items for larger proportions of the population in developed countries: In more developed countries there are very high percentages of students reporting the existence of many of the household items which makes them less appropriate as indicators of wealth.

Table 16.6
Scale reliabilities for home possession indices in OECD countries

	WEALTH	HEDRES	CULTPOSS	HOMEPOS
Australia	0.60	0.60	0.61	0.62
Austria	0.61	0.41	0.60	0.63
Belgium	0.62	0.47	0.62	0.61
Canada	0.61	0.53	0.63	0.64
Czech Republic	0.67	0.51	0.59	0.65
Denmark	0.60	0.43	0.64	0.58
Finland	0.59	0.44	0.67	0.61
France	0.65	0.46	0.64	0.64
Germany	0.64	0.47	0.61	0.62
Greece	0.66	0.42	0.53	0.65
Hungary	0.70	0.50	0.62	0.73
Iceland	0.55	0.44	0.61	0.59
Ireland	0.56	0.56	0.62	0.61
Italy	0.61	0.47	0.57	0.64
Japan	0.60	0.46	0.61	0.65
Korea	0.64	0.50	0.62	0.73
Luxembourg	0.61	0.49	0.64	0.64
Mexico	0.83	0.60	0.58	0.77
Netherlands	0.57	0.42	0.57	0.56
New Zealand	0.67	0.57	0.61	0.69
Norway	0.58	0.57	0.67	0.61
Poland	0.72	0.60	0.58	0.74
Portugal	0.73	0.51	0.67	0.75
Slovak Republic	0.68	0.62	0.62	0.71
Spain	0.64	0.45	0.58	0.64
Sweden	0.58	0.53	0.63	0.63
Switzerland	0.60	0.42	0.58	0.59
Turkey	0.78	0.66	0.54	0.76
United Kingdom	0.60	0.58	0.66	0.63
United States	0.70	0.65	0.66	0.74
Median	**0.62**	**0.49**	**0.61**	**0.64**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Table 16.7
Scale reliabilities for home possession indices in partner countries/economies

	WEALTH	HEDRES	CULTPOSS	HOMEPOS
Argentina	0.77	0.59	0.52	0.75
Azerbaijan	0.81	0.57	0.59	0.64
Brazil	0.80	0.58	0.46	0.72
Bulgaria	0.74	0.66	0.63	0.74
Chile	0.77	0.60	0.52	0.78
Colombia	0.77	0.64	0.56	0.74
Croatia	0.68	0.44	0.65	0.68
Estonia	0.69	0.45	0.57	0.69
Hong Kong-China	0.61	0.47	0.55	0.71
Indonesia[1]	0.78	0.55	0.48	0.65
Israel	0.75	0.62	0.68	0.65
Jordan	0.80	0.71	0.52	0.72
Kyrgyzstan	0.73	0.47	0.46	0.62
Latvia	0.69	0.51	0.57	0.70
Liechtenstein	0.59	0.33	0.66	0.57
Lithuania	0.70	0.50	0.70	0.73
Macao-China	0.70	0.50	0.56	0.72
Montenegro	0.72	0.52	0.58	0.66
Qatar	0.78	0.70	0.55	0.75
Romania	0.79	0.69	0.51	0.80
Russian Federation	0.68	0.56	0.46	0.72
Serbia	0.71	0.54	0.65	0.70
Slovenia	0.61	0.42	0.65	0.63
Chinese Taipei	0.56	0.55	0.68	0.68
Thailand	0.82	0.63	0.54	0.80
Tunisia	0.84	0.72	0.56	0.73
Uruguay	0.79	0.58	0.58	0.74
Median	**0.74**	**0.56**	**0.56**	**0.72**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.
1. Indonesia had omitted item ST13Q13 ("Dishwasher") from their national questionnaire and reliabilities for WEALTH and HOMEPOS were computed without this item.

Interest and enjoyment of science learning

Eight items are used to measure general interest in science learning in PISA 2006. While the interest items which are embedded in the test instrument provide data on interest in specific contexts, the items here will provide data on students' interest in more general terms. All items were inverted for scaling and positive scores indicate higher levels of interest in learning science. Item wording and model parameters are displayed in Table 16.8.

Table 16.8
Item parameters for interest in science learning (INTSCIE)

Item	How much interest do you have in learning about the following <broad science> topics?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST21Q01	a) Topics in physics	0.04	–1.52	–0.04	1.57
ST21Q02	b) Topics in chemistry	–0.05	–1.42	0.01	1.40
ST21Q03	c) The biology of plants	0.03	–1.62	0.06	1.55
ST21Q04	d) Human biology	–0.76	–1.35	–0.12	1.47
ST21Q05	e) Topics in astronomy	–0.2	–1.28	0.11	1.17
ST21Q06	f) Topics in geology	0.32	–1.7	0.08	1.62
ST21Q07	g) Ways scientists design experiments	0.11	–1.43	0.07	1.35
ST21Q08	h) What is required for scientific explanations	0.51	–1.60	0.05	1.55

Note: Item categories were "high interest", "medium interest", "low interest" and "no interest"; all items were inverted for scaling.

Four items are used to measure enjoyment of science learning in PISA 2006. All items were inverted for IRT scaling so that positive WLE scores on this new index for PISA 2006 indicate higher levels of enjoyment of science. Table 16.9 shows the item wording and the international item parameters for this scale.

Table 16.9
Item parameters for enjoyment of science (JOYSCIE)

Item	How much do you agree with the statements below?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST16Q01	a) I generally have fun when I am learning <broad science> topics	–0.43	–4.17	–0.4	4.57
ST16Q02	b) I like reading about <broad science>	0.51	–4.39	–0.12	4.51
ST16Q03	c) I am happy doing <broad science> problems	1.01	–4.6	0.02	4.57
ST16Q04	d) I enjoy acquiring new knowledge in <broad science>	–0.69	–3.91	–0.61	4.52
ST16Q05	e) I am interested in learning about <broad science>	–0.41	–3.84	–0.42	4.26

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

The fit for a two-factor model was not satisfactory for the pooled international sample and most of the country sub-samples (see Table 16.10). However, the lack of fit is mostly due to correlated error terms between interest items about similar topics (like biology of plants and human biology). The results also show high correlations (typically between 0.70 and 0.80) between the two constructs whose strength does not vary much across country sub-samples.

Table 16.10
Model fit and estimated latent correlations for interest in and enjoyment of science learning[1]

			Model fit			Latent correlations between:
		RMSEA	RMR	CFI	NNFI	INTSCIE/JOYSCIE
OECD	Australia	0.114	0.048	0.90	0.90	0.77
	Austria	0.110	0.061	0.88	0.88	0.76
	Belgium	0.113	0.056	0.88	0.88	0.81
	Canada	0.122	0.063	0.88	0.88	0.81
	Czech Republic	0.119	0.068	0.83	0.83	0.72
	Denmark	0.150	0.060	0.84	0.84	0.80
	Finland	0.138	0.055	0.85	0.85	0.72
	France	0.107	0.052	0.89	0.89	0.77
	Germany	0.106	0.056	0.89	0.89	0.79
	Greece	0.113	0.070	0.87	0.87	0.74
	Hungary	0.105	0.060	0.86	0.86	0.68
	Iceland	0.137	0.051	0.88	0.88	0.78
	Ireland	0.119	0.060	0.88	0.88	0.81
	Italy	0.102	0.043	0.88	0.88	0.72
	Japan	0.106	0.048	0.91	0.91	0.81
	Korea	0.115	0.057	0.86	0.86	0.81
	Luxembourg	0.100	0.053	0.90	0.90	0.71
	Mexico	0.121	0.048	0.81	0.81	0.59
	Netherlands	0.136	0.058	0.85	0.85	0.81
	New Zealand	0.106	0.050	0.90	0.90	0.83
	Norway	0.097	0.036	0.94	0.94	0.74
	Poland	0.126	0.062	0.85	0.85	0.73
	Portugal	0.114	0.047	0.86	0.86	0.67
	Slovak Republic	0.111	0.055	0.85	0.85	0.61
	Spain	0.139	0.068	0.84	0.84	0.77
	Sweden	0.105	0.039	0.93	0.93	0.81
	Switzerland	0.090	0.048	0.92	0.92	0.78
	Turkey	0.118	0.065	0.87	0.87	0.71
	United Kingdom	0.103	0.046	0.89	0.89	0.67
	United States	0.099	0.039	0.93	0.93	0.73
	OECD	**0.106**	**0.048**	**0.90**	**0.90**	**0.75**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.11 shows the scale reliabilities for both indices in OECD and partner countries. The internal consistency for both scales is very high and typically above 0.80 for *INTSCIE* and 0.90 for *JOYSCIE*.

Table 16.11
Scale reliabilities for interest in and enjoyment of science learning

		INTSCIE	JOYSCIE			INTSCIE	JOYSCIE
OECD	Australia	0.87	0.94	*Partners*	Argentina	0.83	0.88
	Austria	0.79	0.92		Azerbaijan	0.81	0.83
	Belgium	0.85	0.89		Brazil	0.85	0.83
	Canada	0.83	0.93		Bulgaria	0.82	0.85
	Czech Republic	0.76	0.88		Chile	0.82	0.87
	Denmark	0.87	0.93		Colombia	0.78	0.85
	Finland	0.85	0.93		Croatia	0.78	0.88
	France	0.83	0.90		Estonia	0.75	0.87
	Germany	0.80	0.92		Hong Kong-China	0.83	0.91
	Greece	0.81	0.92		Indonesia	0.76	0.83
	Hungary	0.75	0.89		Israel	0.88	0.94
	Iceland	0.89	0.95		Jordan	0.81	0.82
	Ireland	0.84	0.93		Kyrgyzstan	0.75	0.79
	Italy	0.80	0.88		Latvia	0.72	0.82
	Japan	0.86	0.93		Liechtenstein	0.87	0.94
	Korea	0.81	0.91		Lithuania	0.73	0.87
	Luxembourg	0.82	0.92		Macao-China	0.79	0.89
	Mexico	0.81	0.82		Montenegro	0.81	0.87
	Netherlands	0.85	0.92		Qatar	0.88	0.91
	New Zealand	0.85	0.93		Romania	0.81	0.83
	Norway	0.90	0.94		Russian Federation	0.76	0.84
	Poland	0.79	0.92		Serbia	0.77	0.85
	Portugal	0.83	0.87		Slovenia	0.79	0.90
	Slovak Republic	0.81	0.88		Chinese Taipei	0.87	0.91
	Spain	0.83	0.90		Thailand	0.84	0.82
	Sweden	0.88	0.95		Tunisia	0.71	0.77
	Switzerland	0.82	0.91		Uruguay	0.80	0.89
	Turkey	0.83	0.91				
	United Kingdom	0.85	0.91				
	United States	0.87	0.93				
	Median	**0.83**	**0.92**		**Median**	**0.81**	**0.87**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Motivation to learn science

Five items measuring the construct of instrumental motivation were included in the PISA 2006 main study. All items were inverted for IRT scaling: positive WLE scores on this new index for PISA 2006 indicate higher levels of instrumental motivation to learn science.

Table 16.12
Item parameters for instrumental motivation to learn science (INSTSCIE)

Item	How much do you agree with the statements below? (Strongly agree/Agree/Disagree/Strongly disagree)	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST35Q01	a) Making an effort in my <school science> subject(s) is worth it because this will help me in the work I want to do later on	–0.21	–3.46	–0.39	3.85
ST35Q02	b) What I learn in my <school science> subject(s) is important for me because I need this for what I want to study later on	0.24	–3.62	–0.17	3.79
ST35Q03	c) I study <school science> because I know it is useful for me	–0.37	–3.66	–0.67	4.33
ST35Q04	d) Studying my <school science> subject(s) is worthwhile for me because what I learn will improve my career prospects	0.00	–3.66	–0.45	4.11
ST35Q05	e) I will learn many things in my <school science> subject(s) that will help me get a job	0.34	–3.76	–0.29	4.05

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Expectations about tertiary science studies and working in science-related careers are another important aspect of student motivations to learning science. Four items measuring students' motivations to take up a science-related career were included in the student questionnaire. All items are reverse scored so that positive WLE scores on this index indicate higher levels of motivation to take up a science-related career.

Table 16.13
Item parameters for future-oriented science motivation (SCIEFUT)

Item	How much do you agree with the statements below?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST29Q01	a) I would like to work in a career involving <broad science>	–0.77	–3.58	0.34	3.24
ST29Q02	b) I would like to study <broad science> after <secondary school>	–0.27	–3.57	0.44	3.13
ST29Q03	c) I would like to spend my life doing advanced <broad science>	0.71	–3.81	0.57	3.23
ST29Q04	d) I would like to work on <broad science> projects as an adult	0.33	–3.78	0.30	3.48

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Table 16.14
Model fit and estimated latent correlations for motivation to learn science[1]

		Model fit			Latent correlations between:
		RMSEA	RMR	CFI	NNFI	INSTSCIE/SCIEFUT
OECD	Australia	0.130	0.028	0.95	0.95	0.79
	Austria	0.064	0.028	0.98	0.98	0.59
	Belgium	0.079	0.018	0.98	0.98	0.82
	Canada	0.092	0.023	0.98	0.98	0.78
	Czech Republic	0.089	0.019	0.97	0.97	0.65
	Denmark	0.065	0.019	0.99	0.99	0.71
	Finland	0.095	0.023	0.97	0.97	0.73
	France	0.132	0.038	0.94	0.94	0.80
	Germany	0.064	0.018	0.98	0.98	0.67
	Greece	0.089	0.022	0.97	0.97	0.72
	Hungary	0.071	0.018	0.98	0.98	0.67
	Iceland	0.070	0.016	0.99	0.99	0.77
	Ireland	0.112	0.027	0.96	0.96	0.79
	Italy	0.059	0.017	0.99	0.99	0.73
	Japan	0.106	0.019	0.97	0.97	0.74
	Korea	0.116	0.021	0.95	0.95	0.68
	Luxembourg	0.062	0.023	0.98	0.98	0.69
	Mexico	0.069	0.020	0.97	0.97	0.58
	Netherlands	0.080	0.018	0.98	0.98	0.60
	New Zealand	0.121	0.031	0.95	0.95	0.79
	Norway	0.061	0.019	0.99	0.99	0.66
	Poland	0.061	0.015	0.98	0.98	0.59
	Portugal	0.129	0.033	0.94	0.94	0.73
	Slovak Republic	0.105	0.018	0.96	0.96	0.71
	Spain	0.097	0.027	0.97	0.97	0.78
	Sweden	0.071	0.022	0.98	0.98	0.71
	Switzerland	0.078	0.031	0.97	0.97	0.70
	Turkey	0.100	0.022	0.96	0.96	0.63
	United Kingdom	0.117	0.026	0.95	0.95	0.73
	United States	0.078	0.021	0.98	0.98	0.67
	OECD	**0.086**	**0.020**	**0.97**	**0.96**	**0.72**

1. Model estimates based on international student calibration sample (500 students per OECD country).

The fit for the two-factor model was satisfactory for the pooled OECD sample (RMSEA = 0.086) and in most country sub-samples. The latent correlation between the two construct ranges is quite high and ranges between 0.59 and 0.82.

Table 16.15 shows that the reliabilities for both scales are highly satisfactory around 0.90 in most countries.

Table 16.15
Scale reliabilities for instrumental and future-oriented science motivation

		INTSCIE	SCIEFUT			INTSCIE	SCIEFUT
OECD	Australia	0.95	0.93	Partners	Argentina	0.88	0.90
	Austria	0.91	0.90		Azerbaijan	0.86	0.88
	Belgium	0.92	0.93		Brazil	0.86	0.89
	Canada	0.94	0.94		Bulgaria	0.87	0.89
	Czech Republic	0.89	0.92		Chile	0.91	0.91
	Denmark	0.92	0.95		Colombia	0.88	0.90
	Finland	0.92	0.91		Croatia	0.92	0.91
	France	0.91	0.92		Estonia	0.85	0.87
	Germany	0.90	0.91		Hong Kong-China	0.94	0.93
	Greece	0.89	0.93		Indonesia	0.86	0.83
	Hungary	0.88	0.91		Israel	0.92	0.93
	Iceland	0.95	0.94		Jordan	0.81	0.82
	Ireland	0.93	0.92		Kyrgyzstan	0.84	0.82
	Italy	0.88	0.90		Latvia	0.85	0.89
	Japan	0.94	0.94		Liechtenstein	0.91	0.90
	Korea	0.93	0.92		Lithuania	0.89	0.91
	Luxembourg	0.92	0.92		Macao-China	0.91	0.91
	Mexico	0.86	0.88		Montenegro	0.90	0.91
	Netherlands	0.93	0.93		Qatar	0.87	0.89
	New Zealand	0.94	0.93		Romania	0.86	0.87
	Norway	0.92	0.93		Russian Federation	0.88	0.86
	Poland	0.91	0.91		Serbia	0.89	0.90
	Portugal	0.94	0.91		Slovenia	0.91	0.90
	Slovak Republic	0.90	0.94		Chinese Taipei	0.92	0.94
	Spain	0.92	0.93		Thailand	0.84	0.87
	Sweden	0.93	0.93		Tunisia	0.82	0.80
	Switzerland	0.91	0.90		Uruguay	0.91	0.93
	Turkey	0.91	0.93				
	United Kingdom	0.92	0.92				
	United States	0.91	0.92				
	Median	**0.92**	**0.92**		**Median**	**0.88**	**0.90**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Table 16.16
Item parameters for science self-efficacy (SCIEEFF)

Item	How easy do you think it would be for you to perform the following tasks on your own?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST17Q01	a) Recognise the science question that underlies a newspaper report on a health issue	–0.28	–1.93	–0.49	2.41
ST17Q02	b) Explain why earthquakes occur more frequently in some areas than in others	–0.63	–1.49	–0.15	1.64
ST17Q03	c) Describe the role of antibiotics in the treatment of disease	0.17	–1.55	–0.12	1.68
ST17Q04	d) Identify the science question associated with the disposal of garbage	0.09	–1.80	–0.21	2.02
ST17Q05	e) Predict how changes to an environment will affect the survival of certain species	–0.05	–1.48	–0.19	1.67
ST17Q06	f) Interpret the scientific information provided on the labelling of food items	–0.05	–1.61	–0.17	1.78
ST17Q07	g) Discuss how new evidence can lead you to change your understanding about the possibility of life on Mars	0.49	–1.43	–0.14	1.57
ST17Q08	h) Identify the better of two explanations for the formation of acid rain	0.25	–1.46	–0.16	1.62

Note: Item categories were "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"; all items were inverted for scaling.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Self-related cognitions in science

Eight items measuring students' science self-efficacy (their confidence in performing science-related tasks) were included. These items cover important themes identified in the science literacy framework: identifying scientific questions, explaining phenomena scientifically and using scientific evidence. All items are reverse coded for IRT scaling so that positive WLE scores on this new index for PISA 2006 indicate higher levels of self-efficacy in science.

Six items on science self-concept were included in the student questionnaire. The items were inverted for scaling so that positive WLE scores on this new PISA 2006 index indicate a positive self-concept in science.

Table 16.17
Item parameters for science self-concept (SCSCIE)

Item	How much do you agree with the statements below? (Strongly agree/Agree/Disagree/Strongly disagree)	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST37Q01	a) Learning advanced <school science> topics would be easy for me	0.56	–4.23	–0.07	4.29
ST37Q02	b) I can usually give good answers to <test questions> on <school science> topics	–0.55	–4.35	–0.47	4.82
ST37Q03	c) I learn <school science> topics quickly	–0.19	–4.30	–0.09	4.38
ST37Q04	d) <School science> topics are easy for me	0.41	–4.35	0.13	4.23
ST37Q05	e) When I am being taught <school science>. I can understand the concepts very well	–0.22	–4.32	–0.3	4.62
ST37Q06	f) I can easily understand new ideas in <school science>	–0.01	–4.32	–0.16	4.49

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Table 16.18
Model fit and estimated latent correlations for science self-efficacy and science self-concept[1]

		RMSEA	Model fit		NNFI	Latent correlations between:
			RMR	CFI		SCIEEFF/SCSCIE
OECD	Australia	0.074	0.029	0.95	0.95	0.66
	Austria	0.068	0.034	0.94	0.94	0.58
	Belgium	0.060	0.032	0.96	0.96	0.58
	Canada	0.053	0.028	0.97	0.98	0.54
	Czech Republic	0.044	0.026	0.96	0.97	0.44
	Denmark	0.062	0.027	0.96	0.96	0.67
	Finland	0.039	0.020	0.98	0.98	0.61
	France	0.049	0.030	0.96	0.96	0.49
	Germany	0.035	0.024	0.98	0.98	0.65
	Greece	0.049	0.038	0.96	0.96	0.45
	Hungary	0.043	0.027	0.97	0.97	0.35
	Iceland	0.054	0.028	0.97	0.97	0.64
	Ireland	0.059	0.031	0.96	0.96	0.66
	Italy	0.046	0.028	0.96	0.97	0.45
	Japan	0.059	0.024	0.97	0.97	0.49
	Korea	0.066	0.027	0.95	0.95	0.46
	Luxembourg	0.054	0.029	0.96	0.96	0.59
	Mexico	0.057	0.026	0.95	0.95	0.39
	Netherlands	0.055	0.028	0.96	0.96	0.52
	New Zealand	0.049	0.021	0.97	0.97	0.63
	Norway	0.042	0.021	0.98	0.98	0.53
	Poland	0.035	0.020	0.98	0.98	0.43
	Portugal	0.054	0.024	0.96	0.96	0.34
	Slovak Republic	0.062	0.030	0.94	0.94	0.42
	Spain	0.055	0.032	0.97	0.97	0.46
	Sweden	0.056	0.027	0.97	0.97	0.57
	Switzerland	0.033	0.024	0.99	0.99	0.57
	Turkey	0.064	0.032	0.95	0.95	0.43
	United Kingdom	0.050	0.023	0.97	0.97	0.64
	United States	0.046	0.025	0.98	0.98	0.60
	OECD	**0.041**	**0.017**	**0.98**	**0.98**	**0.55**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.18 shows the results of confirmatory factor analyses (CFA) for a two-dimensional model of self-efficacy and self-concept items. The model fit is very well for the pooled OECD sample and also for all country sub-samples. The estimated latent correlation between the two constructs is moderately high and ranges between 0.35 and 0.67.

Table 16.19 shows internal consistencies for the two scales. Both constructs have high reliabilities across participating countries, for *SCIEEFF* the reliabilities are typically around 0.80 and for *SCSCIE* even higher (around 0.90).

Table 16.19
Scale reliabilities for science self-efficacy and science self-concept

		SCIEEFF	SCSCIE			SCIEEFF	SCSCIE
OECD	Australia	0.88	0.93	Partners	Argentina	0.76	0.89
	Austria	0.80	0.90		Azerbaijan	0.78	0.88
	Belgium	0.82	0.91		Brazil	0.79	0.86
	Canada	0.85	0.94		Bulgaria	0.81	0.87
	Czech Republic	0.78	0.88		Chile	0.81	0.89
	Denmark	0.84	0.94		Colombia	0.77	0.87
	Finland	0.83	0.92		Croatia	0.79	0.89
	France	0.79	0.91		Estonia	0.76	0.86
	Germany	0.82	0.90		Hong Kong-China	0.83	0.93
	Greece	0.77	0.90		Indonesia	0.73	0.86
	Hungary	0.76	0.88		Israel	0.84	0.92
	Iceland	0.88	0.94		Jordan	0.75	0.83
	Ireland	0.82	0.93		Kyrgyzstan	0.76	0.82
	Italy	0.75	0.89		Latvia	0.74	0.82
	Japan	0.85	0.93		Liechtenstein	0.85	0.93
	Korea	0.83	0.92		Lithuania	0.77	0.86
	Luxembourg	0.83	0.91		Macao-China	0.80	0.92
	Mexico	0.77	0.86		Montenegro	0.77	0.87
	Netherlands	0.84	0.91		Qatar	0.85	0.88
	New Zealand	0.87	0.92		Romania	0.79	0.84
	Norway	0.87	0.92		Russian Federation	0.79	0.84
	Poland	0.82	0.88		Serbia	0.78	0.90
	Portugal	0.84	0.91		Slovenia	0.80	0.90
	Slovak Republic	0.77	0.88		Chinese Taipei	0.85	0.93
	Spain	0.83	0.92		Thailand	0.79	0.87
	Sweden	0.87	0.93		Tunisia	0.66	0.82
	Switzerland	0.82	0.92		Uruguay	0.78	0.90
	Turkey	0.81	0.92				
	United Kingdom	0.85	0.91				
	United States	0.87	0.93				
	Median	**0.83**	**0.92**		**Median**	**0.79**	**0.87**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Table 16.20
Item parameters for general value of science (GENSCIE)

Item	How much do you agree with the statements below? (Strongly agree/Agree/Disagree/Strongly disagree)	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST18Q01	a) Advances in <broad science and technology> usually improve people's living conditions	–0.42	–1.68	–0.96	2.64
ST18Q02	b) <Broad science> is important for helping us to understand the natural world	–0.52	–1.71	–0.92	2.63
ST18Q04	d) Advances in <broad science and technology> usually help improve the economy	0.31	–2.37	–0.45	2.82
ST18Q06	f) <Broad science> is valuable to society	0.04	–1.93	–0.80	2.73
ST18Q09	i) Advances in <broad science and technology> usually bring social benefits	0.60	–2.43	–0.36	2.79

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Value of science

Five items measuring perceptions of the general value of science were included in the student questionnaire. The items are reverse coded for scaling so that positive WLE scores on this new PISA 2006 index indicate positive students' perceptions of the general value of science. Table 16.20 shows the item wording and international IRT parameters that were used for scaling.

Five items measuring perceptions of the personal value of science were included in the student questionnaire. The items were inverted for scaling so that positive WLE scores on this new PISA 2006 index indicate positive students' perceptions of the general value of science. Table 16.21 shows the item wording and international IRT parameters used for scaling.

Table 16.21
Item parameters for personal value of science (PERSCIE)

Item	How much do you agree with the statements below?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST18Q03	c) Some concepts in <broad science> help me see how I relate to other people	0.05	–2.97	–0.05	3.02
ST18Q05	e) I will use <broad science> in many ways when I am an adult	–0.02	–2.52	–0.21	2.74
ST18Q07	g) <Broad science> is very relevant to me	0.26	–2.36	–0.08	2.44
ST18Q08	h) I find that <broad science> helps me to understand the things around me	–0.52	–2.36	–0.45	2.81
ST18Q10	j) When I leave school there will be many opportunities for me to use <broad science>	0.24	–2.35	–0.18	2.53

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Table 16.22
Model fit and estimated latent correlations for general and personal value of science[1]

		Model fit				Latent correlations between:
		RMSEA	RMR	CFI	NNFI	GENSCIE/PERSCIE
OECD	Australia	0.090	0.029	0.93	0.94	0.75
	Austria	0.085	0.033	0.92	0.92	0.83
	Belgium	0.066	0.025	0.95	0.95	0.77
	Canada	0.083	0.027	0.94	0.94	0.82
	Czech Republic	0.077	0.025	0.92	0.92	0.74
	Denmark	0.062	0.022	0.96	0.96	0.71
	Finland	0.101	0.026	0.90	0.90	0.72
	France	0.070	0.026	0.95	0.95	0.72
	Germany	0.086	0.033	0.93	0.93	0.77
	Greece	0.059	0.026	0.92	0.93	0.71
	Hungary	0.083	0.031	0.91	0.91	0.74
	Iceland	0.107	0.039	0.92	0.92	0.79
	Ireland	0.078	0.030	0.94	0.95	0.75
	Italy	0.072	0.023	0.93	0.93	0.77
	Japan	0.092	0.031	0.92	0.92	0.79
	Korea	0.076	0.027	0.93	0.93	0.63
	Luxembourg	0.091	0.037	0.91	0.91	0.74
	Mexico	0.068	0.021	0.93	0.93	0.86
	Netherlands	0.062	0.018	0.96	0.96	0.74
	New Zealand	0.099	0.032	0.93	0.93	0.81
	Norway	0.077	0.025	0.95	0.95	0.79
	Poland	0.095	0.026	0.90	0.90	0.76
	Portugal	0.074	0.017	0.94	0.94	0.83
	Slovak Republic	0.054	0.018	0.96	0.96	0.69
	Spain	0.080	0.027	0.93	0.93	0.77
	Sweden	0.102	0.032	0.93	0.93	0.84
	Switzerland	0.064	0.026	0.96	0.96	0.77
	Turkey	0.090	0.029	0.91	0.92	0.75
	United Kingdom	0.084	0.027	0.94	0.94	0.77
	United States	0.098	0.030	0.93	0.93	0.77
	OECD	**0.076**	**0.023**	**0.94**	**0.94**	**0.78**

1. Model estimates based on international student calibration sample (500 students per OECD country).

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 16.23 shows the results of a CFA for general and personal value of science items. The model fit is satisfactory for the pooled sample and in all but three country sub-samples. Not unexpectedly, the estimated latent correlation between the two construct is quite high and ranges between 0.63 and 0.86.

Table 16.23 shows the scale reliabilities for general and personal value of science. For both constructs, the internal consistencies are high across participating countries. However, reliabilities for *GENSCIE* are somewhat lower in many partner countries.

Table 16.23
Scale reliabilities for general and personal value of science

		GENSCIE	PERSCIE			GENSCIE	PERSCIE
OECD	Australia	0.81	0.86	Partners	Argentina	0.69	0.77
	Austria	0.72	0.80		Azerbaijan	0.68	0.67
	Belgium	0.70	0.78		Brazil	0.67	0.75
	Canada	0.78	0.85		Bulgaria	0.73	0.76
	Czech Republic	0.71	0.79		Chile	0.72	0.78
	Denmark	0.70	0.85		Colombia	0.61	0.71
	Finland	0.76	0.83		Croatia	0.69	0.79
	France	0.68	0.80		Estonia	0.65	0.74
	Germany	0.75	0.81		Hong Kong-China	0.80	0.79
	Greece	0.66	0.74		Indonesia	0.62	0.66
	Hungary	0.67	0.77		Israel	0.79	0.83
	Iceland	0.80	0.87		Jordan	0.69	0.69
	Ireland	0.75	0.83		Kyrgyzstan	0.65	0.69
	Italy	0.68	0.73		Latvia	0.65	0.73
	Japan	0.80	0.76		Liechtenstein	0.79	0.84
	Korea	0.77	0.75		Lithuania	0.70	0.77
	Luxembourg	0.79	0.83		Macao-China	0.72	0.73
	Mexico	0.65	0.71		Montenegro	0.68	0.78
	Netherlands	0.78	0.78		Qatar	0.81	0.82
	New Zealand	0.79	0.85		Romania	0.69	0.71
	Norway	0.82	0.85		Russian Federation	0.64	0.77
	Poland	0.71	0.80		Serbia	0.68	0.76
	Portugal	0.74	0.79		Slovenia	0.74	0.81
	Slovak Republic	0.71	0.76		Chinese Taipei	0.82	0.79
	Spain	0.72	0.79		Thailand	0.72	0.72
	Sweden	0.82	0.85		Tunisia	0.64	0.62
	Switzerland	0.73	0.80		Uruguay	0.68	0.80
	Turkey	0.79	0.81				
	United Kingdom	0.78	0.83				
	United States	0.82	0.84				
	Median	**0.75**	**0.80**		**Median**	**0.69**	**0.76**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Science-related activities

Student participation in non-compulsory activities related to science or choice of course combinations with an emphasis on this subject are important indicators of engagement. Furthermore, out-of-school activities relating to science can contribute considerably to students' engagement and learning in science.

Six items measuring students' activities related to science were included in the student questionnaire. The items are reverse scored for scaling so that positive WLE scores on this new PISA 2006 index indicate higher frequencies of students' science activities. Table 16.24 shows the item wording and the international IRT parameters used for scaling.

Table 16.24
Item parameters for science activities (SCIEACT)

Item	How often do you do these things? (Very often/Regularly/Sometimes/Never or hardly ever)	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST19Q01	a) Watch TV programmes about <broad science>	–1.99	–2.50	0.82	1.68
ST19Q02	b) Borrow or buy books on <broad science> topics	0.31	–1.72	0.60	1.12
ST19Q03	c) Visit web sites about <broad science> topics	–0.17	–1.75	0.59	1.17
ST19Q04	d) Listen to radio programmes about advances in <broad science>	0.58	–1.45	0.43	1.03
ST19Q05	e) Read <broad science> magazines or science articles in newspapers	–0.68	–1.89	0.51	1.29
ST19Q06	f) Attend a <science club>	0.96	–0.21	–0.02	0.23

Note: Item categories were "very often", "regularly", "sometimes" and "never or hardly ever"; all items were inverted for scaling.

Table 16.25 shows the scale reliabilities across countries, which are satisfactory and range typically between 0.75 and 0.80 in a majority of countries.

Table 16.25
Scale reliabilities for the science activities index

		SCIEACT			SCIEACT
OECD	Australia	0.80	*Partners*	Argentina	0.77
	Austria	0.76		Azerbaijan	0.71
	Belgium	0.77		Brazil	0.80
	Canada	0.80		Bulgaria	0.75
	Czech Republic	0.77		Chile	0.81
	Denmark	0.79		Colombia	0.76
	Finland	0.76		Croatia	0.78
	France	0.75		Estonia	0.75
	Germany	0.77		Hong Kong-China	0.84
	Greece	0.82		Indonesia	0.71
	Hungary	0.77		Israel	0.88
	Iceland	0.81		Jordan	0.67
	Ireland	0.79		Kyrgyzstan	0.76
	Italy	0.76		Latvia	0.76
	Japan	0.80		Liechtenstein	0.78
	Korea	0.80		Lithuania	0.75
	Luxembourg	0.80		Macao-China	0.80
	Mexico	0.78		Montenegro	0.75
	Netherlands	0.78		Qatar	0.83
	New Zealand	0.78		Romania	0.76
	Norway	0.81		Russian Federation	0.76
	Poland	0.76		Serbia	0.73
	Portugal	0.80		Slovenia	0.81
	Slovak Republic	0.75		Chinese Taipei	0.84
	Spain	0.78		Thailand	0.77
	Sweden	0.79		Tunisia	0.60
	Switzerland	0.78		Uruguay	0.78
	Turkey	0.82			
	United Kingdom	0.78			
	United States	0.80			
	Median	**0.78**		**Median**	**0.76**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Scientific literacy and environment

Five items measuring students' awareness of environmental issues were included in the student questionnaire. Positive WLE scores on this index indicate higher levels of students' awareness of environmental issues. Table 16.26 shows the item wording and international IRT parameters for this scale.

Table 16.26
Item parameters for awareness of environmental issues (ENVAWARE)

		Parameter estimates			
Item	**How informed are you about the following environmental issues?**	**Delta**	**Tau(1)**	**Tau(2)**	**Tau(3)**
ST22Q01	a) The increase of greenhouse gases in the atmosphere	–0.05	–1.87	–0.01	1.88
ST22Q02	b) Use of genetically modified organisms (<GMO>)	0.88	–2.03	0.14	1.88
ST22Q03	c) Acid rain	–0.16	–2.13	0.07	2.07
ST22Q04	d) Nuclear waste	0.02	–2.47	0.23	2.25
ST22Q05	e) The consequences of clearing forests for other land use	–0.68	–1.56	–0.05	1.61

Six items measuring students' perception of environmental issues as a concern were included in the student questionnaire. The items were reverse scored for scaling so that positive WLE scores on this index indicate higher levels of students' concerns about environmental issues. Table 16.27 shows the item wording and the international IRT parameters for this scale.

Table 16.27
Item parameters for perception of environmental issues (ENVPERC)

Item	Do you see the environmental issues below as a serious concern for yourself and/or others?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST24Q01	a) Air pollution	–0.6	–0.81	–0.03	0.84
ST24Q02	b) Energy shortages	0.12	–1.53	0.13	1.40
ST24Q03	c) Extinction of plants and animals	0.11	–1.06	–0.02	1.08
ST24Q04	d) Clearing of forests for other land use	0.07	–1.5	0.21	1.29
ST24Q05	e) Water shortages	–0.06	–2.09	1.33	0.76
ST24Q06	f) Nuclear waste	0.35	–1.61	0.23	1.38

Students' optimism regarding environmental issues was measured by six items in the student questionnaire. The items were inverted for scaling so that positive WLE scores on this index indicate higher levels of students' optimism about environmental issues. Table 16.28 shows the item wording and the international IRT parameters for this scale.

Table 16.28
Item parameters for environmental optimism (ENVOPT)

Item	Do you think problems associated with the environmental issues below will improve or get worse over the next 20 years?	Parameter estimates		
		Delta	Tau(1)	Tau(2)
ST25Q01	a) Air pollution	0.20	0.05	–0.05
ST25Q02	b) Energy shortages	–0.45	–0.71	0.71
ST25Q03	c) Extinction of plants and animals	0.18	–0.57	0.57
ST25Q04	d) Clearing of forests for other land use	0.32	–0.37	0.37
ST25Q05	e) Water shortages	–0.25	–0.75	0.75
ST25Q06	f) Nuclear waste	0.00	–0.73	0.73

Seven items measuring students' responsibility for sustainable development were included in the student questionnaire. The items were reverse coded for scaling so that positive WLE scores on this new PISA 2006 index indicate higher levels of students' responsibility for sustainable development. Table 16.29 shows the item wording and the international IRT parameters for this scale.

Table 16.29
Item parameters for responsibility for sustainable development (RESPDEV)

Item	How much do you agree with the statements below?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST26Q01	a) It is important to carry out regular checks on the emissions from cars as a condition of their use	–0.42	–1.45	–0.86	2.32
ST26Q02	b) It disturbs me when energy is wasted through the unnecessary use of electrical appliances	0.61	–1.94	–0.05	1.99
ST26Q03	c) I am in favour of having laws that regulate factory emissions even if this would increase the price of products	0.65	–1.91	–0.06	1.97
ST26Q04	d) To reduce waste, the use of plastic packaging should be kept to a minimum	0.06	–1.79	–0.35	2.14
ST26Q05	e) Industries should be required to prove that they safely dispose of dangerous waste materials	–0.59	–1.21	–0.76	1.97
ST26Q06	f) I am in favour of having laws that protect the habitats of endangered species	–0.58	–1.12	–0.68	1.80
ST26Q07	g) Electricity should be produced from renewable sources as much as possible, even if this increases the cost	0.28	–1.70	–0.32	2.02

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Table 16.30 shows the model fit for a four-dimensional model for the environment-related items in PISA 2006. The model fit is satisfactory across participating countries and for the pooled OECD sample.

Table 16.30
Model fit environment-related constructs[1]

		Model fit			
		RMSEA	RMR	CFI	NNFI
OECD	Australia	0.056	0.027	0.90	0.90
OECD	Austria	0.043	0.030	0.90	0.90
OECD	Belgium	0.049	0.032	0.86	0.87
OECD	Canada	0.052	0.028	0.90	0.91
OECD	Czech Republic	0.051	0.030	0.86	0.86
OECD	Denmark	0.046	0.033	0.90	0.90
OECD	Finland	0.049	0.033	0.91	0.91
OECD	France	0.044	0.030	0.90	0.91
OECD	Germany	0.055	0.036	0.85	0.85
OECD	Greece	0.045	0.031	0.89	0.89
OECD	Hungary	0.041	0.025	0.90	0.90
OECD	Iceland	0.049	0.040	0.91	0.91
OECD	Ireland	0.046	0.032	0.91	0.92
OECD	Italy	0.055	0.032	0.83	0.83
OECD	Japan	0.043	0.023	0.94	0.94
OECD	Korea	0.041	0.020	0.93	0.93
OECD	Luxembourg	0.047	0.036	0.92	0.92
OECD	Mexico	0.049	0.025	0.89	0.90
OECD	Netherlands	0.050	0.030	0.87	0.87
OECD	New Zealand	0.057	0.032	0.89	0.89
OECD	Norway	0.057	0.041	0.89	0.89
OECD	Poland	0.051	0.028	0.89	0.89
OECD	Portugal	0.049	0.022	0.91	0.91
OECD	Slovak Republic	0.050	0.032	0.88	0.88
OECD	Spain	0.041	0.022	0.93	0.93
OECD	Sweden	0.047	0.036	0.92	0.92
OECD	Switzerland	0.051	0.035	0.87	0.87
OECD	Turkey	0.048	0.025	0.93	0.93
OECD	United Kingdom	0.049	0.029	0.92	0.92
OECD	United States	0.056	0.032	0.91	0.91
	OECD	**0.044**	**0.021**	**0.92**	**0.92**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.31
Estimated latent correlations for environment-related constructs[1]

		Latent correlations between					
		RESPDEV/ ENVAWARE	RESPDEV/ENVPERC	RESPDEV/ENVOPT	ENVAWARE/ ENVPERC	ENVAWARE/ ENVOPT	ENVPERC/ENVOPT
OECD	Australia	0.42	0.44	-0.12	0.23	-0.12	-0.15
OECD	Austria	0.29	0.26	-0.15	0.09	-0.15	-0.16
OECD	Belgium	0.30	0.42	-0.13	0.17	-0.14	-0.11
OECD	Canada	0.39	0.39	-0.15	0.18	-0.09	-0.15
OECD	Czech Republic	0.46	0.22	-0.08	-0.06	-0.06	-0.16
OECD	Denmark	0.37	0.31	-0.10	0.03	-0.04	-0.09
OECD	Finland	0.42	0.48	-0.32	0.18	-0.20	-0.13
OECD	France	0.39	0.43	-0.25	0.26	-0.25	-0.19
OECD	Germany	0.31	0.52	-0.06	0.19	-0.08	-0.17
OECD	Greece	0.34	0.66	-0.24	0.18	-0.28	-0.23
OECD	Hungary	0.37	0.52	-0.22	0.08	-0.22	-0.12
OECD	Iceland	0.38	0.20	-0.13	-0.13	-0.07	0.05
OECD	Ireland	0.55	0.37	-0.15	0.09	-0.07	-0.04
OECD	Italy	0.32	0.49	-0.24	0.16	-0.17	-0.26
OECD	Japan	0.45	0.60	-0.02	0.32	0.08	-0.07
OECD	Korea	0.31	0.43	-0.10	0.24	-0.06	-0.03
OECD	Luxembourg	0.27	0.47	-0.23	0.00	-0.16	-0.30
OECD	Mexico	0.28	0.43	-0.10	0.06	-0.06	-0.15
OECD	Netherlands	0.31	0.33	-0.09	0.04	-0.18	-0.17
OECD	New Zealand	0.48	0.43	-0.12	0.14	-0.15	-0.07
OECD	Norway	0.48	0.41	-0.05	0.19	-0.08	-0.01
OECD	Poland	0.35	0.33	-0.05	0.01	-0.13	-0.01
OECD	Portugal	0.44	0.23	-0.29	0.15	-0.34	-0.17
OECD	Slovak Republic	0.46	0.20	-0.17	-0.05	-0.26	-0.15
OECD	Spain	0.43	0.45	-0.18	0.21	-0.30	-0.17
OECD	Sweden	0.34	0.32	-0.18	0.10	-0.18	-0.17
OECD	Switzerland	0.50	0.35	-0.22	0.19	-0.15	-0.09
OECD	Turkey	0.21	0.34	-0.05	0.09	-0.27	-0.04
OECD	United Kingdom	0.40	0.33	-0.19	0.15	-0.10	-0.03
OECD	United States	0.31	0.34	-0.08	0.02	-0.05	0.02
	OECD	**0.34**	**0.44**	**-0.13**	**0.12**	**-0.14**	**-0.11**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.33 shows the estimated latent correlations for the four environment-related constructs. The highest correlations (0.44 for the pooled sample) are found for *RESPDEV* and *ENVPERC*. Environmental optimism has (weak) negative correlations with all other constructs.

Table 16.32
Scale reliabilities for environment-related scales in OECD countries

	ENVAWARE	ENVPERC	ENVOPT	RESPDEV
OECD				
Australia	0.79	0.85	0.79	0.80
Austria	0.76	0.77	0.68	0.75
Belgium	0.75	0.74	0.74	0.77
Canada	0.77	0.84	0.79	0.82
Czech Republic	0.73	0.74	0.73	0.72
Denmark	0.76	0.81	0.72	0.79
Finland	0.75	0.78	0.75	0.83
France	0.73	0.71	0.76	0.76
Germany	0.77	0.78	0.69	0.76
Greece	0.66	0.71	0.77	0.71
Hungary	0.68	0.72	0.76	0.74
Iceland	0.79	0.82	0.72	0.82
Ireland	0.76	0.82	0.73	0.76
Italy	0.73	0.68	0.74	0.70
Japan	0.79	0.84	0.79	0.81
Korea	0.75	0.81	0.78	0.78
Luxembourg	0.78	0.82	0.78	0.80
Mexico	0.74	0.76	0.85	0.70
Netherlands	0.74	0.75	0.74	0.76
New Zealand	0.79	0.84	0.79	0.79
Norway	0.78	0.85	0.79	0.84
Poland	0.77	0.81	0.79	0.79
Portugal	0.79	0.77	0.84	0.77
Slovakia	0.74	0.74	0.77	0.71
Spain	0.77	0.79	0.78	0.75
Sweden	0.78	0.85	0.76	0.82
Switzerland	0.75	0.76	0.74	0.79
Turkey	0.72	0.85	0.87	0.84
United Kingdom	0.79	0.84	0.80	0.81
United States	0.79	0.88	0.83	0.80
Median	**0.76**	**0.80**	**0.77**	**0.78**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Table 16.33
Scale reliabilities for environment-related scales in non-OECD countries

	ENVAWARE	ENVPERC	ENVOPT	RESPDEV
OECD				
Argentina	0.72	0.75	0.84	0.69
Azerbaijan	0.74	0.77	0.85	0.72
Brazil	0.77	0.80	0.88	0.68
Bulgaria	0.75	0.81	0.85	0.72
Chile	0.74	0.73	0.82	0.71
Colombia	0.74	0.79	0.87	0.64
Croatia	0.75	0.77	0.80	0.69
Estonia	0.70	0.72	0.76	0.72
Hong Kong-China	0.72	0.80	0.78	0.75
Indonesia	0.64	0.81	0.80	0.59
Israel	0.78	0.83	0.83	0.85
Jordan	0.66	0.78	0.83	0.73
Kyrgyzstan	0.71	0.83	0.82	0.68
Latvia	0.67	0.64	0.72	0.64
Liechtenstein	0.72	0.81	0.76	0.81
Lithuania	0.71	0.73	0.78	0.71
Macao-China	0.70	0.84	0.80	0.70
Montenegro	0.76	0.76	0.83	0.71
Qatar	0.77	0.83	0.82	0.81
Romania	0.71	0.81	0.80	0.69
Russian Federation	0.73	0.75	0.78	0.68
Serbia	0.75	0.77	0.84	0.73
Slovenia	0.73	0.79	0.77	0.76
Chinese Taipei	0.81	0.93	0.84	0.80
Thailand	0.73	0.80	0.86	0.72
Tunisia	0.55	0.64	0.73	0.64
Uruguay	0.73	0.75	0.84	0.72
Median	**0.73**	**0.79**	**0.82**	**0.71**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 16.32 shows the scale reliabilities for environment-related scale in OECD countries, Table 16.3 those for partner countries. For all four constructs the internal consistencies are generally satisfactory across participating countries. Only in few countries scale reliabilities are below 0.70.

Science career preparation

Four items measuring students' perceptions of the usefulness of schooling as preparation for science-related careers were included in the student questionnaire. All items were inverted so that positive WLE scores on this index indicate higher levels of agreement with usefulness of schooling for this purpose. Item wording and international IRT parameter are shown in Table 16.34.

Table 16.34
Item parameters for school preparation for science career (CARPREP)

Item	How much do you agree with the statements below?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST27Q01	a) The subjects available at my school provide students with the basic skills and knowledge for a <science-related career>	–0.38	–2.81	–0.76	3.57
ST27Q02	b) The <school science> subjects at my school provide students with the basic skills and knowledge for many different careers	–0.26	–2.96	–0.61	3.57
ST27Q03	c) The subjects I study provide me with the basic skills and knowledge for a <science-related career>	0.28	–2.86	–0.37	3.23
ST27Q04	d) My teachers equip me with the basic skills and knowledge I need for a <science-related career>	0.35	–2.65	–0.59	3.24

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Four items measuring students' perceptions of being informed about science-related careers are included in the student questionnaire. Items were reverse coded so that positive WLE scores on this index indicate higher levels of information about science-related careers. Table 16.35 shows the wording of items and the international IRT parameters used for scaling.

Table 16.35
Item parameters for student information on science careers (CARINFO)

Item	How informed are you about these topics?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST28Q01	a) <Science-related careers> that are available in the job market	–0.02	–3.34	0.06	3.28
ST28Q02	b) Where to find information about <science-related careers>	–0.35	–3.04	0.06	2.98
ST28Q03	c) The steps a student needs to take if they want a <science-related career>	–0.19	–2.82	0.07	2.75
ST28Q04	d) Employers or companies that hire people to work in <science-related careers>	0.57	–3.03	0.16	2.87

Note: Item categories were "Very well informed", "Fairly informed", "Not well informed" and "Not informed at all"; all items were inverted for scaling.

Table 16.36 shows the results of a CFA for the items related to science career preparation. The model fit is satisfactory for the pooled sample and in most country sub-samples. The estimated latent correlation between the two constructs is moderate to high: between 0.26 and 0.57.

Table 16.37 shows the scale reliabilities for *CARINFO* and *CARPREP* across participating countries. For both scales, the internal consistencies are high around 0.80.

Table 16.36
Model fit and estimated latent correlations for science career preparation indices[1]

		Model fit			Latent correlations between:
		RMSEA	RMR	CFI	NNFI	CARPREP/CARINFO
OECD	Australia	0.098	0.027	0.95	0.95	0.54
	Austria	0.060	0.023	0.98	0.98	0.53
	Belgium	0.088	0.025	0.95	0.95	0.44
	Canada	0.082	0.024	0.96	0.96	0.47
	Czech Republic	0.057	0.021	0.98	0.98	0.43
	Denmark	0.047	0.015	0.99	0.99	0.56
	Finland	0.025	0.010	1.00	1.00	0.43
	France	0.066	0.026	0.96	0.96	0.44
	Germany	0.047	0.018	0.98	0.98	0.35
	Greece	0.037	0.020	0.99	0.99	0.48
	Hungary	0.051	0.018	0.96	0.96	0.36
	Iceland	0.060	0.018	0.98	0.98	0.52
	Ireland	0.101	0.033	0.94	0.94	0.51
	Italy	0.045	0.019	0.98	0.98	0.33
	Japan	0.078	0.017	0.97	0.97	0.47
	Korea	0.050	0.015	0.98	0.98	0.26
	Luxembourg	0.060	0.022	0.98	0.98	0.47
	Mexico	0.050	0.024	0.98	0.98	0.43
	Netherlands	0.092	0.022	0.94	0.94	0.40
	New Zealand	0.114	0.028	0.92	0.92	0.44
	Norway	0.057	0.018	0.98	0.98	0.57
	Poland	0.053	0.014	0.98	0.98	0.39
	Portugal	0.108	0.023	0.93	0.93	0.40
	Slovak Republic	0.057	0.018	0.98	0.98	0.41
	Spain	0.078	0.021	0.95	0.95	0.45
	Sweden	0.047	0.019	0.99	0.99	0.47
	Switzerland	0.031	0.014	0.99	0.99	0.47
	Turkey	0.086	0.023	0.96	0.96	0.32
	United Kingdom	0.053	0.020	0.98	0.98	0.48
	United States	0.078	0.021	0.97	0.97	0.45
	OECD	**0.054**	**0.012**	**0.98**	**0.98**	**0.45**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.37
Scale reliabilities for science career preparation indices

		CARPREP	CARINFO			CARPREP	CARINFO
OECD	Australia	0.81	0.86	Partners	Argentina	0.79	0.80
	Austria	0.83	0.79		Azerbaijan	0.74	0.76
	Belgium	0.81	0.78		Brazil	0.78	0.79
	Canada	0.83	0.84		Bulgaria	0.79	0.80
	Czech Republic	0.81	0.78		Chile	0.80	0.81
	Denmark	0.81	0.84		Colombia	0.79	0.74
	Finland	0.86	0.80		Croatia	0.83	0.78
	France	0.81	0.76		Estonia	0.78	0.76
	Germany	0.83	0.78		Hong Kong-China	0.79	0.77
	Greece	0.74	0.77		Indonesia	0.72	0.76
	Hungary	0.75	0.69		Israel	0.84	0.82
	Iceland	0.84	0.84		Jordan	0.71	0.68
	Ireland	0.79	0.83		Kyrgyzstan	0.70	0.71
	Italy	0.79	0.72		Latvia	0.76	0.75
	Japan	0.83	0.87		Liechtenstein	0.87	0.82
	Korea	0.80	0.78		Lithuania	0.80	0.72
	Luxembourg	0.82	0.81		Macao-China	0.80	0.77
	Mexico	0.75	0.81		Montenegro	0.76	0.80
	Netherlands	0.72	0.82		Qatar	0.81	0.79
	New Zealand	0.81	0.84		Romania	0.75	0.74
	Norway	0.82	0.87		Russian Federation	0.76	0.73
	Poland	0.80	0.82		Serbia	0.79	0.79
	Portugal	0.77	0.82		Slovenia	0.79	0.79
	Slovak Republic	0.81	0.80		Chinese Taipei	0.84	0.77
	Spain	0.78	0.80		Thailand	0.78	0.76
	Sweden	0.85	0.85		Tunisia	0.68	0.67
	Switzerland	0.82	0.78		Uruguay	0.75	0.79
	Turkey	0.88	0.83				
	United Kingdom	0.83	0.85				
	United States	0.82	0.85				
	Median	**0.81**	**0.82**		**Median**	**0.79**	**0.77**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Science learning and teaching

Four items measuring students' reports on the frequency of interactive teaching in science lessons were included in the student questionnaire. Items were inverted such that positive WLE scores on this index indicate higher frequencies of interactive science teaching. Table 16.38 shows the item wording and international IRT parameters used for scaling.

Table 16.38
Item parameters for science teaching: interaction (SCINTACT)

Item	When learning <school science> topics at school, how often do the following activities occur?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST34Q01	a) Students are given opportunities to explain their ideas	–0.64	–1.72	0.11	1.61
ST34Q05	e) The lessons involve students' opinions about the topics	–0.07	–1.72	0.09	1.63
ST34Q09	i) There is a class debate or discussion	0.48	–1.70	0.26	1.44
ST34Q13	m) The students have discussions about the topics	0.23	–1.71	0.08	1.63

Four items measuring students' reports on the frequency of hands-on activities in science lessons are included in the main study. These were reverse scored so that positive WLE scores on this index indicate higher frequencies of this type of science teaching. Table 16.39 shows the item wording and the international item parameters used for scaling.

Table 16.39
Item parameters for science teaching: hands-on activities (SCHANDS)

Item	When learning <school science> topics at school, how often do the following activities occur?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST34Q02	b) Students spend time in the laboratory doing practical experiments	0.57	–2.17	0.38	1.8
ST34Q03	c) Students are required to design how a <school science> question could be investigated in the laboratory	0.64	–1.77	0.15	1.62
ST34Q06	f) Students are asked to draw conclusions from an experiment they have conducted	–0.74	–1.82	–0.02	1.84
ST34Q14	n) Students do experiments by following the instructions of the teacher	–0.47	–1.71	0.05	1.67

Three items measuring students' reports on the frequency of student investigations in science lessons were included in the student questionnaire. Responses were inverted so that positive WLE scores on this index indicate perceived higher frequencies of this type of science teaching. Table 16.40 shows the item wording and the international IRT parameters for this scale.

Table 16.40
Item parameters for science teaching: student investigations (SCINVEST)

Item	When learning <school science> topics at school, how often do the following activities occur?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST34Q08	h) Students are allowed to design their own experiments	0.16	–1.36	0.08	1.28
ST34Q11	k) Students are given the chance to choose their own investigations	0.12	–1.78	0.24	1.53
ST34Q16	p) Students are asked to do an investigation to test out their own ideas	–0.28	–1.88	0.16	1.72

Note: Item categories were "In all lessons", "In most lessons", "In some lessons" and "Never or hardly ever"; all items were inverted for scaling.

Five items measuring students' reports on the frequency of teaching in science lessons with a focus on applications are included in the student questionnaire. All items were reverse scored so that positive WLE scores on this index indicate higher frequencies of this type of science teaching. Table 16.41 shows the item wording and the international IRT parameters for this scale.

Table 16.41
Item parameters for science teaching: focus on models or applications (SCAPPLY)

Item	When learning <school science> topics at school, how often do the following activities occur?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
ST34Q07	g) The teacher explains how a <school science> idea can be applied to a number of different phenomena (*e.g.* the movement of objects, substances with similar properties)	–0.64	–1.85	–0.02	1.87
ST34Q12	l) The teacher uses science to help students understand the world outside school	0.3	–1.99	0.13	1.87
ST34Q15	o) The teacher clearly explains the relevance of <broad science> concepts to our lives	–0.11	–2.05	0.18	1.87
ST34Q17	q) The teacher uses examples of technological application to show how <school science> is relevant to society	0.45	–1.95	0.15	1.8

Note: Item categories were "In all lessons", "In most lessons", "In some lessons" and "Never or hardly ever"; all items were inverted for scaling.

Table 16.42
Model fit for CFA with science teaching and learning[1]

		Model fit			
		RMSEA	RMR	CFI	NNFI
OECD	Australia	0.079	0.039	0.91	0.91
	Austria	0.067	0.039	0.94	0.94
	Belgium	0.077	0.045	0.91	0.91
	Canada	0.078	0.046	0.92	0.92
	Czech Republic	0.089	0.040	0.88	0.88
	Denmark	0.070	0.036	0.91	0.92
	Finland	0.094	0.037	0.83	0.84
	France	0.079	0.054	0.86	0.87
	Germany	0.071	0.042	0.91	0.91
	Greece	0.076	0.050	0.91	0.91
	Hungary	0.069	0.048	0.92	0.92
	Iceland	0.077	0.046	0.90	0.90
	Ireland	0.079	0.042	0.90	0.90
	Italy	0.069	0.046	0.93	0.93
	Japan	0.100	0.048	0.87	0.87
	Korea	0.067	0.031	0.93	0.93
	Luxembourg	0.079	0.047	0.91	0.91
	Mexico	0.094	0.057	0.86	0.87
	Netherlands	0.067	0.041	0.92	0.92
	New Zealand	0.072	0.039	0.92	0.92
	Norway	0.081	0.038	0.90	0.90
	Poland	0.084	0.040	0.90	0.90
	Portugal	0.075	0.040	0.91	0.91
	Slovak Republic	0.074	0.037	0.91	0.91
	Spain	0.090	0.051	0.88	0.88
	Sweden	0.100	0.049	0.85	0.85
	Switzerland	0.088	0.049	0.88	0.88
	Turkey	0.092	0.044	0.90	0.90
	United Kingdom	0.083	0.037	0.89	0.89
	United States	0.094	0.049	0.89	0.89
	OECD	**0.071**	**0.035**	**0.93**	**0.93**

1. Model estimates based on international student calibration sample (500 students per OECD country).

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 16.43 shows the model fit for a four-dimensional model for the science teaching and learning items in PISA 2006. The model fit is satisfactory for the pooled OECD sample and in all but two OECD countries.

Table 16.43
Estimated latent correlations for constructs related to science teaching and learning[1]

		Latent correlations between					
		SCINTACT/ SCHANDS	SCINTACT/ SCINVEST	SCINTACT/ SCAPPLY	SCHANDS/ SCINVEST	SCHANDS/ SCAPPLY	SCINVEST/ SCAPPLY
OECD	Australia	0.57	0.55	0.76	0.67	0.58	0.68
	Austria	0.57	0.60	0.58	0.83	0.63	0.59
	Belgium	0.49	0.64	0.56	0.70	0.81	0.60
	Canada	0.59	0.62	0.78	0.67	0.57	0.54
	Czech Republic	0.65	0.64	0.77	0.74	0.77	0.70
	Denmark	0.52	0.57	0.76	0.42	0.71	0.45
	Finland	0.64	0.62	0.68	0.74	0.73	0.60
	France	0.45	0.78	0.64	0.44	0.70	0.67
	Germany	0.47	0.63	0.59	0.74	0.52	0.61
	Greece	0.67	0.60	0.80	0.94	0.73	0.68
	Hungary	0.48	0.48	0.75	1.00	0.55	0.59
	Iceland	0.47	0.35	0.69	0.89	0.49	0.41
	Ireland	0.70	0.67	0.71	0.76	0.73	0.62
	Italy	0.43	0.64	0.68	0.73	0.59	0.80
	Japan	0.63	0.92	0.77	0.89	0.61	0.90
	Korea	0.71	0.92	0.55	0.80	0.49	0.66
	Luxembourg	0.65	0.68	0.83	0.76	0.70	0.67
	Mexico	0.75	0.75	0.80	0.78	0.86	0.82
	Netherlands	0.45	0.62	0.57	0.51	0.67	0.57
	New Zealand	0.70	0.51	0.69	0.65	0.65	0.51
	Norway	0.58	0.54	0.81	0.66	0.66	0.63
	Poland	0.82	0.72	0.86	0.94	0.85	0.74
	Portugal	0.64	0.76	0.77	0.82	0.80	0.71
	Slovak Republic	0.57	0.70	0.69	0.81	0.82	0.85
	Spain	0.61	0.70	0.66	0.78	0.69	0.68
	Sweden	0.70	0.80	0.73	0.67	0.64	0.71
	Switzerland	0.49	0.69	0.71	0.67	0.70	0.61
	Turkey	0.72	0.85	0.75	0.91	0.81	0.84
	United Kingdom	0.69	0.68	0.74	0.72	0.74	0.70
	United States	0.77	0.76	0.81	0.66	0.67	0.73
	OECD	**0.55**	**0.66**	**0.74**	**0.71**	**0.66**	**0.67**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.44 shows the estimated latent correlations for the four environment-related constructs. All four constructs are positively correlated with each other, the highest correlations are found between *SCINTACT* and *SCINVEST* and between *SCHANDS* and *SCINVEST*.

Table 16.45 shows the scale reliabilities for the indices related to science teaching and learning. The internal consistency of all four scales is satisfactory across countries and is typically between 0.70 and 0.80. Similar reliabilities are found in partner countries (see Table 16.44).

Table 16.44
Scale reliabilities for scales to science teaching and learning in OECD countries

	SCINTACT	SCHANDS	SCINVEST	SCAPPLY
Australia	0.80	0.71	0.75	0.81
Austria	0.82	0.78	0.79	0.73
Belgium	0.74	0.72	0.77	0.76
Canada	0.79	0.72	0.77	0.81
Czech Republic	0.79	0.72	0.76	0.76
Denmark	0.79	0.73	0.76	0.74
Finland	0.69	0.70	0.64	0.74
France	0.71	0.69	0.68	0.72
Germany	0.75	0.72	0.77	0.73
Greece	0.77	0.76	0.74	0.74
Hungary	0.75	0.74	0.74	0.74
Iceland	0.67	0.75	0.72	0.77
Ireland	0.78	0.69	0.72	0.78
Italy	0.74	0.79	0.75	0.71
Japan	0.70	0.78	0.65	0.83
Korea	0.76	0.74	0.73	0.78
Luxembourg	0.76	0.76	0.78	0.75
Mexico	0.69	0.75	0.69	0.76
Netherlands	0.73	0.75	0.75	0.75
New Zealand	0.79	0.70	0.74	0.81
Norway	0.80	0.76	0.78	0.76
Poland	0.73	0.65	0.78	0.76
Portugal	0.79	0.72	0.68	0.79
Slovakia	0.77	0.71	0.74	0.72
Spain	0.77	0.74	0.75	0.77
Sweden	0.76	0.75	0.71	0.78
Switzerland	0.72	0.76	0.75	0.75
Turkey	0.76	0.81	0.76	0.81
United Kingdom	0.77	0.69	0.75	0.77
United States	0.77	0.75	0.79	0.80
Median	**0.76**	**0.74**	**0.75**	**0.76**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

Table 16.45
Scale reliabilities for scales to science teaching and learning in partner countries/economies

	SCINTACT	SCHANDS	SCINVEST	SCAPPLY
Argentina	0.70	0.75	0.73	0.74
Azerbaijan	0.64	0.76	0.71	0.71
Brazil	0.75	0.74	0.73	0.75
Bulgaria	0.74	0.79	0.74	0.78
Chile	0.74	0.76	0.70	0.78
Colombia	0.64	0.68	0.68	0.71
Croatia	0.82	0.77	0.75	0.80
Estonia	0.72	0.70	0.67	0.72
Hong Kong-China	0.78	0.80	0.72	0.82
Indonesia	0.72	0.72	0.73	0.77
Israel	0.77	0.81	0.78	0.80
Jordan	0.70	0.72	0.67	0.71
Kyrgyzstan	0.63	0.69	0.71	0.69
Latvia	0.72	0.64	0.68	0.67
Liechtenstein	0.77	0.73	0.79	0.79
Lithuania	0.71	0.65	0.69	0.73
Macao-China	0.77	0.73	0.73	0.75
Montenegro	0.71	0.79	0.80	0.80
Qatar	0.78	0.81	0.77	0.79
Romania	0.66	0.73	0.74	0.74
Russian Federation	0.70	0.68	0.74	0.73
Serbia	0.72	0.79	0.76	0.78
Slovenia	0.71	0.73	0.71	0.78
Chinese Taipei	0.78	0.78	0.79	0.83
Thailand	0.70	0.72	0.68	0.76
Tunisia	0.64	0.58	0.64	0.62
Uruguay	0.74	0.73	0.71	0.74
Median	**0.72**	**0.73**	**0.73**	**0.75**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

ICT familiarity

The ICT familiarity questionnaire was an optional instrument administered which was administered in 40 of the participating countries in PISA 2006, for which four scaled indices were computed.

As in PISA 2003, six items measuring the frequency of ICT use related to Internet and entertainment were included in the PISA 2006 student questionnaire. All items are reverse scored so that positive WLE scores on this index indicate high frequencies of ICT use. Table 16.46 shows the item wording and the international parameters used for scaling.

Table 16.46
Item parameters for ICT Internet/entertainment use (INTUSE)

		Parameter estimates				
Item	How often do you use computers for the following reasons?	Delta	Tau(1)	Tau(2)	Tau(3)	Tau(4)
IC04Q01	a) Browse the Internet for information about people, things, or ideas	–0.29	–0.52	–0.54	–0.13	1.18
IC04Q02	b) Play games	–0.05	–0.24	–0.10	–0.19	0.52
IC04Q04	d) Use the Internet to collaborate with a group or team	0.42	0.03	–0.44	–0.07	0.48
IC04Q06	f) Download software from the Internet to (including games)	0.31	0.09	–0.3	–0.22	0.43
IC04Q09	i) Download music from the Internet	–0.05	0.57	–0.45	–0.35	0.24
IC04Q11	k) For communication (*e.g.* e-mail or "chat rooms")	–0.34	0.65	–0.08	–0.43	–0.14

Note: Item categories were "Almost every day", "Once or twice a week", "A few times a month", "Once a month or less" and "Never"; all items were inverted for scaling.

As in PISA 2003, six items measuring the frequency of ICT use related to programming and software packages are included in the PISA 2006 student questionnaire. All items are reverse coded so that positive WLE scores on this index indicate high frequencies of ICT use. Table 16.47 shows the item wording and the international parameters used for scaling.

Table 16.47
Item parameters for ICT program/software use (PRGUSE)

		Parameter estimates				
Item	How often do you use computers for the following reasons?	Delta	Tau(1)	Tau(2)	Tau(3)	Tau(4)
IC04Q03	c) Write documents (*e.g.* with <Word® or WordPerfect®>)	–0.79	–1.04	–0.86	0.16	1.75
IC04Q05	e) Use spreadsheets (*e.g.* <Lotus 1 2 3® or Microsoft Excel®>)	0.21	–0.77	–0.53	0.02	1.27
IC04Q07	g) Drawing, painting or using graphics programs	–0.19	–0.71	–0.27	0.04	0.94
IC04Q08	h) Use educational software such as Mathematics programs	0.46	–0.47	–0.45	0.00	0.92
IC04Q10	j) Writing computer programs	0.31	0.15	–0.39	–0.16	0.40

Note: Item categories were "Almost every day", "Once or twice a week", "A few times a month", "Once a month or less" and "Never"; all items were inverted for scaling.

As in PISA 2003, items measuring students' confidence in doing ICT Internet tasks were included. However, a modified set of six items was used in the PISA 2006 student questionnaire where three items were already included in the previous cycle. All items were inverted for IRT scaling and positive WLE scores on this index indicate high self-confidence. Table 16.48 shows the item wording and the international parameters used for scaling.

Table 16.48
Item parameters for ICT self-confidence in Internet tasks (INTCONF)

		Parameter estimates			
Item	How often do you use computers for the following reasons?	Delta	Tau(1)	Tau(2)	Tau(3)
IC05Q01	a) Chat online	0.01	–1.24	0.73	0.50
IC05Q07	g) Search the Internet for information	–0.71	–0.55	0.45	0.10
IC05Q08	h) Download files or programs from the Internet	0.13	–1.39	0.21	1.18
IC05Q09	i) Attach a file to an e-mail message	0.55	–1.26	0.19	1.07
IC05Q13	m) Download music from the Internet	0.19	–1.54	0.42	1.13
IC05Q15	o) Write and send e-mails	–0.18	–1.13	0.48	0.65

Note: Item categories were " I can do this very well by myself", "I can do this with help from someone", "I know what this means but I cannot do it" and "I don't know what this means"; all items were inverted for scaling.

As in PISA 2003, items measuring student's confidence in doing ICT high-level tasks were included in the PISA 2006 student questionnaire. The set of eight items used in the PISA 2006 main study is modified somewhat from the 2003 item set. Items are inverted for IRT scaling and positive WLE scores on this index indicate high self-confidence. Item wording and international IRT parameters for scaling are shown in Table 16.49.

Table 16.49
Item parameters for ICT self-confidence in high-level ICT tasks (HIGHCONF)

Item	How often do you use computers for the following reasons?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
IC05Q02	b) Use software to find and get rid of computer viruses	0.09	–1.59	0.47	1.11
IC05Q03	c) Edit digital photographs or other graphic images	–0.40	–1.31	0.30	1.01
IC05Q04	d) Create a database (*e.g.* using <Microsoft Access®>)	1.10	–1.05	–0.10	1.15
IC05Q10	j) Use a word processor (*e.g.* to write an essay for school)	–0.96	–0.30	0.26	0.04
IC05Q11	k) Use a spreadsheet to plot a graph	0.08	–0.84	–0.17	1.01
IC05Q12	l) Create a presentation (*e.g.* using <Microsoft PowerPoint®>)	–0.16	–0.73	0.01	0.72
IC05Q14	n) Create a multi-media presentation (with sound, pictures, video)	–0.07	–1.55	0.10	1.46
IC05Q16	p) Construct a web page	0.33	–1.9	0.19	1.71

Note: Item categories were " I can do this very well by myself", "I can do this with help from someone", "I know what this means but I cannot do it" and "I don't know what this means"; all items were inverted for scaling.

Table 16.50 shows the model fit for a four-dimensional model for the ICT familiarity items in PISA 2006. The model fit is satisfactory for the pooled OECD sample and in all but two OECD countries.

Table 16.50
Model fit for CFA with ICT familiarity items[1]

		RMSEA	Model fit		NNFI
			RMR	CFI	
OECD	Australia	0.088	0.073	0.72	0.73
	Austria	0.081	0.079	0.79	0.79
	Belgium	0.080	0.080	0.78	0.78
	Canada	0.097	0.083	0.75	0.75
	Czech Republic	0.084	0.076	0.84	0.84
	Denmark	0.099	0.084	0.69	0.70
	Finland	0.108	0.088	0.69	0.70
	Germany	0.089	0.084	0.76	0.76
	Greece	0.084	0.097	0.84	0.84
	Hungary	0.087	0.083	0.81	0.81
	Iceland	0.089	0.078	0.71	0.72
	Ireland	0.090	0.093	0.79	0.79
	Italy	0.082	0.106	0.84	0.84
	Japan	0.086	0.071	0.83	0.83
	Korea	0.077	0.060	0.79	0.80
	Netherlands	0.079	0.061	0.72	0.72
	New Zealand	0.081	0.086	0.81	0.81
	Norway	0.100	0.082	0.76	0.76
	Poland	0.091	0.099	0.84	0.84
	Portugal	0.096	0.082	0.77	0.77
	Slovak Republic	0.084	0.090	0.83	0.83
	Spain	0.091	0.117	0.78	0.78
	Sweden	0.095	0.091	0.72	0.72
	Switzerland	0.084	0.080	0.76	0.76
	Turkey	0.084	0.079	0.84	0.84
	OECD	**0.084**	**0.082**	**0.81**	**0.81**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.51 shows the estimated latent correlations for the four environment-related constructs. All four constructs are positively correlated with each other, the highest correlations are found between the two constructs reflecting self-confidence in ICT tasks.

Table 16.51
Estimated latent correlations for constructs related to ICT familiarity[1]

		Latent correlations between					
		INTUSE/PRGUSE	INTUSE/INTCONF	INTUSE/ HIGHCONF	PRGUSE/INTCONF	PRGUSE/ HIGHCONF	INTCONF/ HIGHCONF
OECD	Australia	0.53	0.58	0.52	0.16	0.49	0.71
	Austria	0.62	0.53	0.40	0.29	0.53	0.78
	Belgium	0.52	0.64	0.61	0.16	0.58	0.61
	Canada	0.61	0.61	0.56	0.12	0.46	0.67
	Czech Republic	0.54	0.62	0.56	0.37	0.62	0.79
	Denmark	0.70	0.42	0.57	0.17	0.59	0.73
	Finland	0.57	0.73	0.55	0.37	0.69	0.78
	Germany	0.59	0.64	0.53	0.32	0.57	0.78
	Greece	0.78	0.72	0.64	0.44	0.59	0.89
	Hungary	0.66	0.59	0.60	0.21	0.49	0.81
	Iceland	0.65	0.59	0.61	0.17	0.56	0.69
	Ireland	0.61	0.76	0.60	0.31	0.57	0.79
	Italy	0.55	0.73	0.51	0.37	0.66	0.80
	Japan	0.67	0.65	0.55	0.26	0.47	0.84
	Korea	0.76	0.15	0.29	0.01	0.50	0.40
	Netherlands	0.62	0.43	0.60	0.17	0.53	0.47
	New Zealand	0.47	0.54	0.42	0.18	0.44	0.71
	Norway	0.67	0.51	0.66	0.14	0.52	0.68
	Poland	0.53	0.64	0.53	0.27	0.50	0.84
	Portugal	0.57	0.67	0.48	0.24	0.39	0.80
	Slovak Republic	0.60	0.70	0.62	0.39	0.59	0.82
	Spain	0.55	0.76	0.55	0.32	0.66	0.68
	Sweden	0.56	0.58	0.58	0.16	0.57	0.64
	Switzerland	0.59	0.57	0.64	0.10	0.45	0.72
	Turkey	0.77	0.68	0.64	0.37	0.60	0.87
	OECD	**0.61**	**0.65**	**0.60**	**0.21**	**0.54**	**0.76**

1. Model estimates based on international student calibration sample (500 students per OECD country).

Table 16.52
Scale reliabilities for ICT familiarity scales

		INTUSE	PRGUSE	INTCONF	HIGHCONF
OECD	Australia	0.75	0.71	0.83	0.80
	Austria	0.71	0.68	0.80	0.80
	Belgium	0.71	0.72	0.82	0.78
	Canada	0.71	0.74	0.83	0.79
	Czech Republic	0.83	0.78	0.85	0.85
	Denmark	0.66	0.73	0.76	0.80
	Finland	0.72	0.73	0.76	0.83
	Germany	0.71	0.72	0.80	0.82
	Greece	0.82	0.79	0.82	0.84
	Hungary	0.79	0.71	0.86	0.83
	Iceland	0.69	0.75	0.74	0.77
	Ireland	0.78	0.75	0.84	0.84
	Italy	0.82	0.73	0.86	0.83
	Japan	0.80	0.75	0.88	0.89
	Korea	0.66	0.71	0.81	0.82
	Netherlands	0.63	0.73	0.80	0.76
	New Zealand	0.76	0.75	0.83	0.82
	Norway	0.73	0.75	0.84	0.79
	Poland	0.86	0.79	0.90	0.86
	Portugal	0.79	0.75	0.84	0.84
	Slovak Republic	0.78	0.79	0.87	0.86
	Spain	0.77	0.72	0.84	0.82
	Sweden	0.71	0.75	0.77	0.83
	Switzerland	0.72	0.74	0.83	0.82
	Turkey	0.85	0.83	0.89	0.89
	Median	**0.75**	**0.74**	**0.83**	**0.82**
Partners	Bulgaria	0.81	0.79	0.90	0.85
	Chile	0.81	0.80	0.87	0.83
	Colombia	0.83	0.77	0.88	0.85
	Croatia	0.80	0.79	0.86	0.85
	Jordan	0.84	0.80	0.87	0.86
	Latvia	0.78	0.74	0.82	0.77
	Liechtenstein	0.67	0.71	0.81	0.83
	Lithuania	0.82	0.78	0.86	0.83
	Macao-China	0.73	0.73	0.83	0.80
	Qatar	0.84	0.83	0.87	0.87
	Russian Federation	0.89	0.81	0.93	0.88
	Serbia	0.85	0.75	0.92	0.87
	Slovenia	0.78	0.78	0.86	0.82
	Thailand	0.85	0.78	0.89	0.88
	Uruguay	0.82	0.84	0.89	0.85
	Median	**0.82**	**0.78**	**0.87**	**0.85**

Table 16.52 shows the scale reliabilities for the countries that administered the ICT familiarity questionnaire. The internal consistencies are high most countries; in only very few countries there are reliabilities 0.7 for *INTUSE* and *PRGUSE.*

School questionnaire scale indices

The Index on Teacher Shortage (*TCSHORT*) was derived from four items measuring the school principal's perceptions of potential factors hindering instruction at school. Similar items were used in PISA 2000 and 2003. The items were not inverted for scaling such that higher WLE scores indicate higher rates of teacher shortage at a school. Table 16.53 shows the item wording and the international parameters used for IRT scaling.

Table 16.53
Item parameters for teacher shortage (TCSHORT)

Item	Is your school's capacity to provide instruction hindered by any of the following?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
SC14Q01	a) A lack of qualified science teachers	0.10	–1.24	–0.53	1.76
SC14Q02	b) A lack of qualified mathematics teachers	–0.05	–0.92	–0.21	1.12
SC14Q03	c) A lack of qualified <test language> teachers	0.25	–0.82	–0.18	1.00
SC14Q04	d) A lack of qualified teachers of other subjects	–0.30	–1.79	–0.31	2.10

Note: Categories were "not at all", "very little", "to some extent" and "a lot".

The index on the school's educational resources (*SCMATEDU*) was computed on the basis of seven items measuring the school principal's perceptions of potential factors hindering instruction at school. Similar items were used in PISA 2000 and 2003 but question format and item wording were modified for PISA 2006. All items were inverted for IRT scaling and positive WLE scores indicate better quality of educational resources. Table 16.54 shows the item wording and the international parameters used for IRT scaling.

Table 16.54
Item parameters for quality of educational resources (SCMATEDU)

Item	Is your school's capacity to provide instruction hindered by any of the following?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
SC14Q07	g) Shortage or inadequacy of science laboratory equipment	0.40	–1.47	0.25	1.22
SC14Q08	h) Shortage or inadequacy of instructional materials (*e.g.* textbooks)	–0.43	–1.85	0.28	1.57
SC14Q09	i) Shortage or inadequacy of computers for instruction	0.05	–1.49	0.18	1.31
SC14Q10	j) Lack or inadequacy of internet connectivity	–0.50	–0.81	0.04	0.78
SC14Q11	k) Shortage or inadequacy of computer software for instruction	0.12	–1.64	0.13	1.50
SC14Q12	l) Shortage or inadequacy of library materials	0.06	–1.92	0.04	1.88
SC14Q13	m) Shortage or inadequacy of audio-visual resources	0.31	–1.64	–0.02	1.66

Note: Categories were "not at all", "very little", "to some extent" and "a lot"; all items were inverted for scaling.

School principals are asked to report what activities to promote students' learning of science occur at their school. Items were coded (Yes=1, No=0) so that positive WLE scores indicate higher levels of school activities in this area. Table 16.55 shows the item wording and the international parameters used for IRT scaling.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 16.55
Item parameters for school activities to promote the learning of science (SCIPROM)

Item	Is your school involved in any of the following activities to promote engagement with science among students in <national modal grade for 15-year-olds>? (Yes/No)	Parameter estimates
		Delta
SC20Q01	a) Science clubs	0.90
SC20Q02	b) Science fairs	0.76
SC20Q03	c) Science competitions	0.23
SC20Q04	d) Extracurricular science projects (including research)	0.24
SC20Q05	e) Excursions and field trips	–2.13

Note: Categories were "Yes" and "No"; all items were inverted for scaling.

School principals are asked to report what activities to promote students' learning of environmental topics occur at their school. Items will be coded (Yes=1, No=0) so that positive WLE scores indicate higher levels of school activities in this area. Table 1656 shows the item wording and the international parameters used for IRT scaling.

Table 16.56
Item parameters for school activities for learning environmental topics (ENVLEARN)

Item	Does your school organise any of the following activities to provide opportunities to students in <national modal grade for 15-year-olds> to learn about environmental topics?	Parameter estimates
		Delta
SC22Q01	a) <Outdoor education>	–0.37
SC22Q02	b) Trips to museums	–0.77
SC22Q03	c) Trips to science and/or technology centres	–0.09
SC22Q04	d) Extracurricular environmental projects (including research)	0.76
SC22Q05	e) Lectures and/or seminars (*e.g.* guest speakers)	0.46

Note: Categories were "Yes" and "No"; all items were inverted for scaling.

Table 16.57
Scale reliabilities for school-level scales in OECD countries

	TCSHORT	SCMATEDU	SCIPROM	ENVLEARN
Australia	0.87	0.90	0.35	0.60
Austria	0.71	0.87	0.65	0.58
Belgium	0.87	0.84	0.43	0.48
Canada	0.85	0.87	0.59	0.63
Czech Republic	0.72	0.79	0.63	0.46
Denmark	0.71	0.84	0.45	0.57
Finland	0.64	0.86	0.26	0.51
Germany	0.78	0.86	0.63	0.49
Greece	0.92	0.81	0.49	0.34
Hungary	0.67	0.81	0.49	0.53
Iceland	0.82	0.76	0.49	0.50
Ireland	0.75	0.84	0.62	0.74
Italy	0.86	0.86	0.53	0.60
Japan	0.79	0.86	0.62	0.69
Korea	0.87	0.85	0.59	0.60
Luxembourg	0.87	0.86	0.66	0.64
Mexico	0.89	0.90	0.62	0.67
Netherlands	0.75	0.82	0.64	0.62
New Zealand	0.71	0.88	0.59	0.73
Norway	0.75	0.78	0.42	0.49
Poland	0.55	0.85	0.37	0.40
Portugal	0.52	0.83	0.41	0.44
Slovakia	0.67	0.79	0.59	0.32
Spain	0.84	0.85	0.50	0.51
Sweden	0.77	0.82	0.49	0.42
Switzerland	0.76	0.83	0.38	0.42
Turkey	0.93	0.85	0.67	0.62
United Kingdom	0.81	0.89	0.58	0.71
United States	0.84	0.86	0.61	0.67
Median	**0.78**	**0.85**	**0.58**	**0.57**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table 16.57 shows the scale reliabilities for school-level indices in OECD countries. Both *TCSHORT* and *SCMATEDU* have high reliabilities across countries. The internal consistencies for the scales on school activities to learn science and environmental issues are rather low (in some countries even below 0.5).

Table 16.58 shows the scale reliabilities for partner countries. Again, high reliabilities can be observed for the two indices related to school resources but the internal consistencies for the two indices on school activities are low and even very low in some of the countries.

Table 16.58
Scale reliabilities for environment-related scales in partner countries/economies

	TCSHORT	SCMATEDU	SCIPROM	ENVLEARN
Argentina	0.85	0.87	0.66	0.62
Azerbaijan	0.85	0.75	0.57	0.49
Brazil	0.86	0.90	0.41	0.57
Bulgaria1	0.42	0.69	0.45	0.41
Chile	0.83	0.89	0.71	0.66
Colombia	0.87	0.89	0.49	0.49
Croatia	0.71	0.82	0.72	0.59
Estonia	0.71	0.76	0.08	0.49
Hong Kong-China	0.88	0.89	0.30	0.59
Indonesia	0.84	0.89	0.71	0.57
Israel	0.78	0.90	0.69	0.71
Jordan	0.93	0.84	0.60	0.52
Kyrgyzstan	0.85	0.85	0.48	0.57
Latvia	0.72	0.81	0.33	0.55
Liechtenstein	0.88	0.86	0.53	0.28
Lithuania	0.84	0.81	0.43	0.44
Macao-China	0.93	0.84	0.40	0.71
Montenegro	0.72	0.86	0.65	0.65
Qatar	0.95	0.86	0.64	0.39
Romania	0.64	0.83	0.58	0.44
Russian Federation	0.88	0.81	0.41	0.45
Serbia	0.67	0.77	0.66	0.50
Slovenia	0.62	0.80	0.62	0.57
Chinese Taipei	0.93	0.93	0.72	0.68
Thailand	0.74	0.92	0.62	0.61
Tunisia	0.70	0.74	0.59	0.52
Uruguay	0.82	0.90	0.65	0.62
Median	**0.84**	**0.85**	**0.59**	**0.57**

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.
1. Reliability for SCIPROM in Bulgaria was calculated without the item SC20Q01 ("science clubs").

Parent questionnaire scale indices

Parent questionnaire indices are only available for the 16 countries which chose to administer the optional parent questionnaire.

Six items measuring students' activities related to science at age 10 were included in the parent questionnaire. The items were inverted for scaling so that positive WLE scores on this index indicate higher frequencies of students' science activities. The item wording and international parameters for IRT scaling are shown in Table 16.59.

Seven items measuring parents' perceptions of the quality of school learning were included in the parent questionnaire. The items were reverse scored prior to scaling so that positive WLE scores on this index indicate positive evaluations of the school's quality. Table 16.60 shows then item wording and the international parameters used for IRT scaling.

Table 16.59
Item parameters for science activities at age 10 (PQSCIACT)

Item	Thinking back to when your child was about 10 years old, how often would your child have done these things?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
PA02Q01	a) Watched TV programmes about science	–0.89	–2.40	1.08	1.31
PA02Q02	b) Read books on scientific discoveries	–0.05	–1.85	0.89	0.97
PA02Q03	c) Watched, read or listened to science fiction	–0.79	–1.81	0.69	1.12
PA02Q04	d) Visited web sites about science topics	0.63	–0.97	0.60	0.38
PA02Q05	e) Attended a science club	1.09	–0.36	0.11	0.24

Note: Categories were "very often", "regularly", "sometimes" and "never"; all items were inverted for scaling.

Table 16.60
Item parameters for parent's perception of school quality (PQSCHOOL)

Item	How much do you agree with the following statements?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
PA03Q01	a) Most of my child's school teachers seem competent and dedicated	–0.4	–2.8	–1.11	3.91
PA03Q02	b) Standards of achievement are high in my child's school	0.11	–3.35	–0.3	3.64
PA03Q03	c) I am happy with the content taught and the instructional methods used in my child's school	0.01	–3.02	–0.75	3.77
PA03Q04	d) I am satisfied with the disciplinary atmosphere in my child's school	0.13	–2.38	–0.71	3.09
PA03Q05	e) My child's progress is carefully monitored by the school	0.19	–2.87	–0.56	3.43
PA03Q06	f) My child's school provides regular and useful information on my child's progress	0.35	–2.49	–0.46	2.95
PA03Q07	g) My child's school does a good job in educating students	–0.37	–2.69	–0.83	3.52

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Four items measuring parents' views on the importance of science were included in the PISA 2006 parent questionnaire. The items were inverted for scaling so that positive WLE scores on this index will indicate positive evaluations of the school's quality. Table 16.61 shows then item wording and the international parameters used for IRT scaling.

Table 16.61
Item parameters for parent's views on importance of science (PQSCIMP)

Item	We are interested in what you think about the need for science skills in the job market today. How much do you agree with the following statements?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
PA04Q01	a) It is important to have good scientific knowledge and skills in order to get any good job in today's world	–0.40	–3.73	0.02	3.71
PA04Q02	b) Employers generally appreciate strong scientific knowledge and skills among their employees	0.48	–4.2	0.12	4.08
PA04Q03	c) Most jobs today require some scientific knowledge and skills	0.33	–4.27	0.04	4.22
PA04Q04	d) It is an advantage in the job market to have good scientific knowledge and skills	–0.41	–3.55	–0.32	3.87

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Four items measuring parents' reports on science career motivation for their child were included in the PISA 2006 parent questionnaire. The items were e inverted for scaling so that positive WLE scores on this index indicate higher levels of science career motivation. One item in this set (*PA05Q01* "Does anybody in your family (including you) work in a <science-related career>?') was not included in the scale since it is unrelated to the construct of career motivation of parents for their child. Item wording and international IRT parameters are shown in Table 16.62.

Table 16.62
Item parameters for parent's reports on science career motivation (PQSCCAR)

Item	Please answer the questions below (Yes/No)	Parameter estimates
PA05Q02	b) Does your child show an interest to work in a <science-related career>?	–0.42
PA05Q03	c) Do you expect your child will go into a <science-related career>?	–0.44
PA05Q04	d) Has your child shown interest in studying science after completing <secondary school>?	0.03
PA05Q05	e) Do you expect your child will study science after completing <secondary school>?	–0.29

Note: Categories were "Yes" and "No"; all items were inverted for scaling.
Five items measuring parents' perceptions of the general value of science were included in the PISA 2006 parent questionnaire; similar items were also included in the student questionnaire. As with the student scale, the items are reverse scored for scaling so that positive WLE scores on this new PISA 2006 index indicate positive parents' perceptions of the general value of science. Table 55 shows the item wording and international parameters used for scaling.

Five items measuring parents' perceptions of the general value of science were included in the PISA 2006 parent questionnaire; similar items were also included in the student questionnaire. As with the student scale, the items are reverse scored for scaling so that positive WLE scores on this new PISA 2006 index indicate positive parents' perceptions of the general value of science. Table 16.63 shows the item wording and international parameters used for scaling.

Table 16.63
Item parameters for parents' view on general value of science (PQGENSCI)

Item	The following question asks about your views towards science. How much do you agree with the following statements?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
PA06Q01	a) Advances in <broad science and technology> usually improve people's living conditions	–0.29	–2.45	–1.10	3.56
PA06Q02	b) <Broad science> is important for helping us to understand the natural world	–0.49	–2.34	–1.32	3.66
PA06Q04	d) Advances in <broad science and technology> usually help improve the economy	0.30	–2.86	–0.67	3.54
PA06Q06	f) <Broad science> is valuable to society	–0.09	–2.42	–1.14	3.56
PA06Q09	i) Advances in <broad science and technology> usually bring social benefits	0.56	–2.82	–0.68	3.50

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Four items measuring parents' perceptions of the personal value of science are included in the PISA 2006 parent questionnaire; similar items are included in the student questionnaire. The items were inverted for scaling so that positive WLE scores indicate positive students' perceptions of the general value of science. Table 16.64 shows the item wording and international parameters used for scaling.

Table 16.64
Item parameters for parent's view on personal value of science (PQPERSCI)

Item	The following question asks about your views towards science. How much do you agree with the following statements?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
PA06Q03	c) Some concepts in <broad science> help me to see how I relate to other people	0.10	–3.51	–0.2	3.71
PA06Q05	e) There are many opportunities for me to use <broad science> in my everyday life	0.51	–3.18	–0.05	3.23
PA06Q07	g) <Broad science> is very relevant to me	–0.03	–2.74	–0.22	2.96
PA06Q08	h) I find that <broad science> helps me to understand the things around me	–0.57	–2.87	–0.63	3.49

Note: Item categories were "strongly agree", "agree", "disagree" and "strongly disagree"; all items were inverted for scaling.

Six items measuring perception of environmental issues as a concern were included in the PISA 2006 parent questionnaire; similar items were also included in the student questionnaire. The items were reverse scored for scaling so that positive WLE scores on this index indicate higher levels of parents' concerns about environmental issues. Table 16.65 shows the item wording and international parameters used for scaling.

Table 16.65
Item parameters for parent's perception of environmental issues (PQENPERC)

Item	Do you see the environmental issues below as a serious concern for yourself and/or others?	Parameter estimates			
		Delta	Tau(1)	Tau(2)	Tau(3)
PA07Q01	a) Air pollution	–0.79	–0.74	0.20	0.54
PA07Q02	b) Energy shortages	–0.06	–1.24	0.44	0.80
PA07Q03	c) Extinction of plants and animals	0.30	–1.3	0.35	0.95
PA07Q04	d) Clearing of forests for other land use	0.27	–1.73	0.8	0.92
PA07Q05	e) Water shortages	–0.1	–1.87	1.37	0.51
PA07Q06	f) Nuclear waste	0.37	–1.79	1.04	0.75

Note: Item categories were "This is a serious concern for me personally as well as others" , "This is a serious concern for other people in my country but not me personally", "This is a serious concern for people in other countries" and "This is not a serious concern to anyone"; all items were inverted for scaling.

Six items measuring parents' optimism regarding environmental issues were included in the PISA 2006 parent questionnaire similar to items on the student questionnaire. These were inverted for scaling so that positive WLE scores on the index indicate higher levels of parents' optimism about environmental issues. Table 16.66 shows the item wording and international parameters used for scaling.

Table 16.66
Item parameters for parent's environmental optimism (PQENVOPT)

Item	Do you think problems associated with the environmental issues below will improve or get worse over the next 20 years?	Parameter estimates		
		Delta	Tau(1)	Tau(2)
PA08Q01	a) Air pollution	–0.04	–0.14	0.14
PA08Q02	b) Energy shortages	–0.33	–0.64	0.64
PA08Q03	c) Extinction of plants and animals	0.14	–0.64	0.64
PA08Q04	d) Clearing of forests for other land use	0.17	–0.44	0.44
PA08Q05	e) Water shortages	0.04	–0.64	0.64
PA08Q06	f) Nuclear waste	0.01	–0.64	0.64

Note: Item categories were "Improve", "Stay about the same" and "Get worse"; all items were inverted for scaling.

Table 16.67 shows the reliabilities for the scale indices derived from the parent questionnaire. Most indices have high reliabilities across countries, only the index PQSCIEACT has somewhat lower internal consistency but it is still satisfactory in most country sub-samples.

Table 16.67
Scale reliabilities for parent questionnaire scales

		PQSCIEACT	PQSCHOOL	PQSCIMP	PQSCCAR	PQGENSCI	PQPERSCI	PQENPERC	PQENVOPT
OECD	Denmark	0.63	0.90	0.88	0.93	0.81	0.83	0.82	0.75
	Germany	0.50	0.84	0.86	0.80	0.77	0.78	0.80	0.76
	Iceland	0.72	0.87	0.84	0.96	0.83	0.82	0.85	0.78
	Italy	0.65	0.82	0.83	0.93	0.77	0.72	0.75	0.82
	Korea	0.78	0.84	0.76	0.82	0.83	0.77	0.84	0.87
	Luxembourg	0.60	0.84	0.86	0.86	0.80	0.79	0.81	0.85
	New Zealand	0.67	0.88	0.86	0.94	0.83	0.83	0.82	0.84
	Poland[1]		0.84						
	Portugal	0.73	0.83	0.83	0.95	0.80	0.76	0.81	0.86
	Turkey	0.67	0.80	0.72	0.85	0.70	0.69	0.77	0.83
Partners	Bulgaria[2]	0.78	0.84	0.73	0.88	0.78	0.72	0.81	0.88
	Colombia	0.67	0.84	0.71	0.81	0.77	0.73	0.71	0.91
	Croatia	0.67	0.78	0.82	0.88	0.81	0.80	0.76	0.85
	Hong Kong-China	0.76	0.80	0.85	0.79	0.81	0.76	0.82	0.80
	Macao-China	0.74	0.82	0.82	0.79	0.80	0.77	0.85	0.86
	Qatar	0.72	0.87	0.75	0.87	0.81	0.78	0.81	0.86

Note: Reliabilities (Cronbach's alpha) computed with weighted national samples.
1. Poland did not submit results for any items in PQSCIEACT, PQSCIMP, PQSCCAR, PQGENSCI, PQENPERC, PQENVOPT.
2. Reliability for the index of PQSCIEACT in Bulgaria was calculated with the omission of PA02Q05.

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

The PISA index of economic, social and cultural status (ESCS)

Computation of ESCS

The index of *ESCS* was used first in the PISA 2000 analysis and at that time was derived from five indices: highest occupational status of parents (*HISEI*), highest educational level of parents (in years of education according to ISCED), family wealth, cultural possessions and home educational resources (all three WLE estimates based on student reports on home possessions).

The *ESCS* for PISA 2003 was derived from three variables related to family background: highest parental education (in number of years of education according to ISCED classification), highest parental occupation (*HISEI* scores), and number of home possessions including books in the home.[6] The rationale for using these three components is that socio-economic status is usually seen as based on education, occupational status and income. As no direct income measure is available from the PISA data, the existence of household items is used as proxy for family wealth.

The *ESCS* has been slightly modified because: (i) there were more indicators available in the recent survey; and (ii) a consultation with countries regarding the mapping of ISCED levels to years of schooling led to minor changes in the indicator of parental education.

As in PISA 2003, the components comprising *ESCS* for 2006 are home possessions, *HOMEPOS* (which comprises all items on the *WEALTH*, *CULTPOS* and *HEDRES* scales (except ST14Q04), as well as books in the home (ST15Q01) recoded into a three-level categorical variable (less than 25 books, 25-100 books, more than 100 books), the higher parental occupation (*HISEI*) and the higher parental education expressed as years of schooling (*PARED*).

Missing values for students with missing data for only one component were imputed with predicted values plus a random component based on a regression on the other two variables. Variables with imputed values were then used for a principal component analysis with an OECD senate weight.

The *ESCS* scores were obtained as component scores for the first principal component with zero being the score of an average OECD student and one the standard deviation across equally weighted OECD countries. For partner countries, *ESCS* scores were obtained as

16.6

$$ESCS = \frac{\beta_1 HISEI' + \beta_2 PARED' + \beta_3 HOMEPOS'}{\varepsilon_f}$$

where β_1, β_2 and β_3 are the OECD factor loadings, *HISEI'*, *PARED'* and *HOMEPOS'* the "OECD-standardised" variables and ε_f is the eigenvalue of the first principal component.[7]

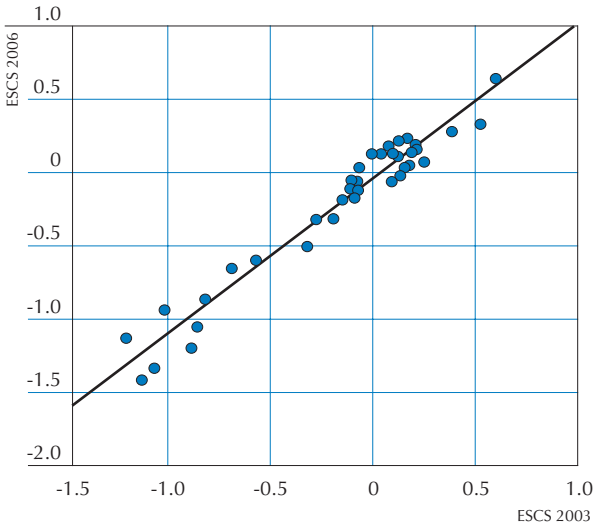
Consistency across cycles

Results for similar *ESCS* indices in 2003 and 2000 showed quite a high degree of consistency (see Schulz, 2006b). Comparing *ESCS* mean scores per country shows that in spite of these differences there is a very high correlation of 0.98 between ESCS 2003 and ESCS 2006 country means (*see* Figure 16.3).

Consistency across countries

Using principal component analysis (PCA) to derive factor loading for each participating country provides insight into the extent to which there are similar relationships between the three components. Table 16.68 shows the PCA results for the OECD countries and Table 16.69 those for partner countries. The tables also include the scale reliabilities for the *z*-standardised variables (Cronbach's Alpha).

Figure 16.3
Scatterplot of country means for ESCS 2003 and ESCS 2006



Note: Weighted averages for OECD and partner countries and economies participating in both cycles.

Table 16.68
Factor loadings and internal consistency of ESCS 2006 in OECD countries

	Factor loadings			
	HISEI	PARED	HOMEPOS	Reliability[1]
Australia	0.80	0.78	0.67	0.59
Austria	0.81	0.78	0.71	0.64
Belgium	0.83	0.80	0.71	0.68
Canada	0.79	0.78	0.67	0.60
Czech Republic	0.84	0.78	0.70	0.65
Denmark	0.79	0.78	0.70	0.63
Finland	0.77	0.75	0.63	0.52
France	0.82	0.79	0.73	0.67
Germany	0.81	0.76	0.72	0.64
Greece	0.84	0.82	0.72	0.71
Hungary	0.83	0.85	0.77	0.74
Iceland	0.80	0.80	0.59	0.57
Ireland	0.81	0.79	0.74	0.67
Italy	0.84	0.81	0.73	0.71
Japan	0.72	0.77	0.68	0.53
Korea	0.76	0.81	0.75	0.66
Luxembourg	0.83	0.81	0.73	0.69
Mexico	0.85	0.86	0.82	0.80
Netherlands	0.82	0.78	0.75	0.68
New Zealand	0.79	0.76	0.69	0.59
Norway	0.78	0.77	0.66	0.55
Poland	0.87	0.86	0.74	0.73
Portugal	0.86	0.85	0.80	0.77
Slovakia	0.85	0.82	0.74	0.72
Spain	0.84	0.82	0.70	0.69
Sweden	0.77	0.73	0.70	0.57
Switzerland	0.80	0.78	0.68	0.62
Turkey	0.80	0.83	0.79	0.72
United Kingdom	0.78	0.75	0.71	0.60
United States	0.80	0.81	0.74	0.67
Median	**0.81**	**0.79**	**0.72**	**0.67**

1. Reliabilities (Standardised Cronbach's alpha) computed with weighted national samples.

Comparing results from within-country PCA reveals that patterns of factor loadings are generally similar across countries. Only in a few countries somehow distinct patterns emerge, however, all three components contribute more or less equally to this index with factor loadings ranging from 0.55 to 0.87. Internal consistency ranges between 0.52 and 0.80, the median scale reliability for the pooled OECD countries is 0.67.

Table 16.69
Factor loadings and internal consistency of ESCS 2006 in partner countries/economies

	Factor loadings			
	HISEI	PARED	HOMEPOS	Reliability[1]
Argentina	0.81	0.78	0.79	0.69
Azerbaijan	0.83	0.83	0.73	0.70
Brazil	0.82	0.83	0.80	0.73
Bulgaria	0.84	0.83	0.77	0.74
Chile	0.86	0.85	0.83	0.80
Colombia	0.82	0.82	0.79	0.73
Croatia	0.83	0.81	0.73	0.69
Estonia	0.81	0.77	0.72	0.63
Hong Kong-China	0.83	0.82	0.77	0.72
Indonesia	0.81	0.83	0.78	0.73
Israel	0.78	0.75	0.73	0.60
Jordan	0.83	0.83	0.75	0.73
Kyrgyzstan	0.76	0.76	0.71	0.57
Latvia	0.81	0.78	0.74	0.66
Liechtenstein	0.83	0.81	0.62	0.63
Lithuania	0.81	0.79	0.76	0.68
Macao-China	0.79	0.77	0.75	0.65
Montenegro	0.80	0.80	0.73	0.66
Qatar	0.82	0.86	0.55	0.60
Romania	0.82	0.75	0.80	0.69
Russian Federation	0.81	0.79	0.69	0.59
Serbia	0.84	0.84	0.72	0.71
Slovenia	0.84	0.84	0.71	0.71
Chinese Taipei	0.77	0.79	0.70	0.61
Thailand	0.85	0.84	0.82	0.78
Tunisia	0.86	0.85	0.83	0.79
Uruguay	0.83	0.81	0.81	0.74
Median	**0.82**	**0.81**	**0.75**	**0.69**

1. Reliabilities (Cronbach's alpha) computed with weighted national samples.

Notes

1. Data on public/private school ownership in Australia are not included in the PISA 2003 database. In Austria, the question on funding was omitted and only for private schools information on government funding was provided to construct this index.

2. The raw index was transformed as (*RESPRES*_raw-2.57)/2.2.

3. The raw index was transformed as (*RESPCURR*_raw-2.72)/1.8.

4. A similar approach was used in the IEA Civic Education Study (*see* Schulz, 2004).

5. This analysis did not include the country-specific items.

6. Here, home possessions only included items from ST17, as well as books in the home (*ST19Q01*) which was recoded into a dichotomous item (0 = "Less than 100 books', 1 = "100 books or more") (see OECD, 2004, p. 283).

7. Only one principal component with an eigenvalue greater than 1 was identified in each of the participating countries.

Reader's Guide

Country codes – the following country codes are used in this report:

OECD countries

AUS	Australia
AUT	Austria
BEL	Belgium
BEF	Belgium (French Community)
BEN	Belgium (Flemish Community)
CAN	Canada
CAE	Canada (English Community)
CAF	Canada (French Community)
CZE	Czech Republic
DNK	Denmark
FIN	Finland
FRA	France
DEU	Germany
GRC	Greece
HUN	Hungary
ISL	Iceland
IRL	Ireland
ITA	Italy
JPN	Japan
KOR	Korea
LUX	Luxembourg
LXF	Luxembourg (French Community)
LXG	Luxembourg (German Community)
MEX	Mexico
NLD	Netherlands
NZL	New Zealand
NOR	Norway
POL	Poland
PRT	Portugal
SVK	Slovak Republic
ESP	Spain
ESB	Spain (Basque Community)
ESC	Spain (Catalonian Community)
ESS	Spain (Castillian Community)
SWE	Sweden
CHE	Switzerland
CHF	Switzerland (French Community)
CHG	Switzerland (German Community)
CHI	Switzerland (Italian Community)
TUR	Turkey
GBR	United Kingdom
IRL	Ireland
SCO	Scotland
USA	United States

Partner countries and economies

ARG	Argentina
AZE	Azerbaijan
BGR	Bulgaria
BRA	Brazil
CHL	Chile
COL	Colombia
EST	Estonia
HKG	Hong Kong-China
HRV	Croatia
IDN	Indonesia
JOR	Jordan
KGZ	Kyrgyztan
LIE	Liechtenstein
LTU	Lithuania
LVA	Latvia
LVL	Latvia (Latvian Community)
LVR	Latvia (Russian Community)
MAC	Macao-China
MNE	Montenegro
QAT	Qatar
ROU	Romania
RUS	Russian Federation
SRB	Serbia
SVN	Slovenia
TAP	Chinese Taipei
THA	Thailand
TUN	Tunisia
URY	Uruguay

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

References

Adams, R.J., Wilson, M. & **Wang, W.C.** (1997), The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement,* No. 21, pp. 1-23.

Adams, R.J., Wilson, M. R. & **Wu, M.L.** (1997), Multilevel item response models: An approach to errors in variables regression, *Journal of Educational and Behavioural Statistics,* No. 22 (1), pp. 46-75.

Adams, R.J. & **Wu, M.L.** (2002), *PISA 2000 Technical Report,* OECD, Paris.

Bollen, K.A. & **Long, S.J.** (1993) (eds.), *Testing Structural Equation Models*, Newbury Park: London.

Beaton, A.E. (1987), Implementing the new design: The NAEP 1983-84 technical report (Rep. No. 15-TR-20), Princeton, NJ: Educational Testing Service.

Buchmann, C. (2000), Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school? *Soc. Forces,* No. 78, pp. 1349-79.

Buchmann, C. (2002), Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges, in Porter, A.C. and Gamoran, A. (eds.). *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150-97), Washington, DC: National Academy Press.

Creemers, B.P.M. (1994), *The Effective Classroom*, London: Cassell.

Cochran, W.G. (1977), *Sampling techniques,* third edition, New York, NY: John Wiley and Sons.

Ganzeboom, H.B.G., de Graaf, P.M. & **Treiman, D.J.** (1992), A standard international socio-economic index of occupational status, *Social Science Research*, No. 21, pp. 1-56.

Ganzeboom H.B. & **Treiman, D.J.** (1996), Internationally comparable measures of occupational status for the 1988 international standard classification of occupations, *Social Science Research*, No. 25, pp. 201-239.

Grisay, A. (2003), Translation procedures in OECD/PISA 2000 international assessment, *Language Testing,* No. 20 (2), pp. 225-240.

Hambleton, R.K., Swaminathan, H. & **Rogers, H.J.** (1991), *Fundamentals of item response theory*, Newbury Park, London, New Delhi: SAGE Publications.

Hambleton, R.K., Merenda, P.F. & **Spielberger, C.D.** (2005), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment,* IEA Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.

Harkness, J.A., Van de Vijver, F.J.R. & **Mohler, P.Ph** (2003), *Cross-Cultural Survey Methods,* Wiley-Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey.

Harvey-Beavis, A. (2002), Student and School Questionnaire Development, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report,* (pp. 33-38), OECD, Paris.

International Labour Organisation (ILO) (1990), *International Standard Classification of Occupations: ISCO-88.* Geneva: International Labour Office.

Jöreskog, K.G. & **Sörbom, Dag** (1993), *LISREL 8 User's Reference Guide*, Chicago: SSI.

Judkins, D.R. (1990), Fay's Method of Variance Estimation, *Journal of Official Statistics*, No. 6 (3), pp. 223-239.

Kaplan, D. (2000), *Structural equation modeling: Foundation and extensions,* Thousand Oaks: SAGE Publications.

Keyfitz, N. (1951), Sampling with probabilities proportionate to science: Adjustment for changes in probabilities, *Journal of the American Statistical Association,* No. 46, American Statistical Association, Alexandria, pp. 105-109.

Kish, L. (1992), Weighting for Unequal, *Pi. Journal of Official Statistics*, No. 8 (2), pp. 183-200.

LISREL (1993), K.G. Jöreskog & D. Sörbom, [computer software], Lincolnwood, IL: Scientific Software International, Inc.

Lohr, S.L. (1999), *Sampling: Design and Analysis*, Duxberry: Pacific Grove.

Macaskill, G., Adams, R.J. & **Wu, M.L.** (1998), Scaling methodology and procedures for the mathematics and science literacy, advanced mathematics and physics scale, in M. Martin and D.L. Kelly, Editors, *Third International Mathematics and Science Study, technical report Volume 3: Implementation and analysis*, Boston College, Chestnut Hill, MA.

Masters, G.N. & **Wright, B.D.** (1997), The Partial Credit Model, in W.J. van der Linden, & R.K. Hambleton (eds.), *Handbook of Modern Item Response Theory* (pp. 101-122), New York/Berlin/Heidelberg: Springer.

Mislevy, R.J. (1991), Randomization-based inference about latent variables from complex samples, *Psychometrika,* No. 56, pp. 177-196.

Mislevy, R.J., Beaton, A., Kaplan, B.A. & **Sheehan, K.** (1992), Estimating population characteristics from sparse matrix samples of item responses, *Journal of Educational Measurement,* No. 29 (2), pp. 133-161.

Mislevy, R.J. & **Sheehan, K.M.** (1987), Marginal estimation procedures, in Beaton, A.E., Editor, 1987. *The NAEP 1983-84 technical report,* National Assessment of Educational Progress, Educational Testing Service, Princeton, pp. 293-360.

Mislevy, R.J. & **Sheehan, K.M.** (1989), Information matrices in latent-variable models, *Journal of Educational Statistics,* No. 14, pp. 335-350.

Mislevy, R.J. & **Sheehan, K.M.** (1989), The role of collateral information about examinees in item parameter estimation, *Psychometrika,* No. 54, pp. 661-679.

Monseur, C. & **Berezner, A.** (2007), The Computation of Equating Errors in International Surveys in Education, *Journal of Applied Measurement,* No. 8 (3), 2007, pp. 323-335.

Monseur, C. (2005), An exploratory alternative approach for student non response weight adjustment, *Studies in Educational Evaluation,* No. 31 (2-3), pp. 129-144.

Muthen, B. & **L. Muthen** (1998), [computer software], *Mplus* Los Angeles, CA: Muthen & Muthen.

Muthen, B., du Toit, S.H.C. & **Spisic, D.** (1997), *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes,* unpublished manuscript.

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

OECD (2003), *Literacy Skills for the World of Tomorrow: Further results from PISA 2000,* OECD, Paris.

OECD (2004), *Learning for Tomorrow's World – First Results from PISA 2003,* OECD, Paris.

OECD (2005), *Technical Report for the OECD Programme for International Student Assessment 2003,* OECD, Paris.

OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006,* OECD, Paris.

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World,* OECD, Paris.

PISA Consortium (2006), *PISA 2006 Main Study Data Management Manual, https://mypisa.acer.edu.au/images/mypisadoc/opmanual/pisa2006_data_management_manual.pdf*

Rasch, G. (1960), Probabilistic models for some intelligence and attainment tests, Copenhagen: Nielsen & Lydiche.

Routitski A. & **Berezner, A.** (2006), Issues influencing the validity of cross-national comparisons of student performance. Data Entry Quality and Parameter Estimation. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in San Francisco, 7-11 April, *https://mypisa.acer.edu.au/images/mypisadoc/aera06routitsky_berezner.pdf*

Rust, K. (1985), Variance Estimation for Complex Estimators in Sample Surveys, *Journal of Official Statistics,* No. 1, pp. 381-397.

Rust, K.F. & **Rao, J.N.K.** (1996), Variance Estimation for Complex Surveys Using Replication Techniques, *Survey Methods in Medical Research,* No. 5, pp. 283-310.

Shao, J. (1996), Resampling Methods in Sample Surveys (with Discussion), *Statistics,* No. 27, pp. 203-254.

Särndal, C.-E., Swensson, B. & **Wretman, J.** (1992), *Model Assisted Survey Sampling,* New York: Springer-Verlag.

SAS® CALIS (1992), W. Hartmann [computer software], Cary, NC: SAS Institute Inc.

Scheerens, J. (1990), School effectiveness and the development of process indicators of school functioning, *School effectiveness and school improvement,* No. 1, pp. 61-80.

Scheerens, J. & **Bosker, R.J.** (1997), *The Foundations of School Effectiveness,* Oxford: Pergamon.

Schulz, W. (2002), Constructing and Validating the Questionnaire composites, in R.J. Adams and M.L. Wu (eds.), *PISA 2000 Technical Report,* OECD, Paris.

Schulz, W. (2004), Mapping Student Scores to Item Responses, in W. Schulz and H. Sibberns (eds.), *IEA Civic Education Study, Technical Report* (pp. 127-132), Amsterdam: IEA.

Schulz, W. (2006a), *Testing Parameter Invariance for Questionnaire Indices using Confirmatory Factor Analysis and Item Response Theory,* Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Schulz, W. (2006b), *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003,* Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Thorndike, R.L. (1973), *Reading comprehension in fifteen countries,* New York, Wiley: and Stockholm: Almqvist & Wiksell.

Travers, K.J. & **Westbury, I.** (1989), *The IEA Study of Mathematics I: Analysis of Mathematics Curricula,* Oxford: Pergamon Press.

Travers, K.J., Garden R.A. & **Rosier, M.** (1989), Introduction to the Study, in Robitaille, D. A. and Garden, R. A. (eds), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics Curricula,* Oxford: Pergamon Press.

Verhelst, N. (2002), Coder and Marker Reliabilaity Studies, in R.J. Adams & M.L. Wu (eds.), *PISA 2000 Technical Report*. OECD, Paris.

Walberg, H.J. (1984), Improving the productivity of American schools, *Educational Leadership,* No. 41, pp. 19-27.

Walberg, H. (1986), Synthesis of research on teaching, in M. Wittrock (ed.), *Handbook of research on teaching* (pp. 214-229), New York: Macmillan.

Walker, M. (2006), *The choice of Likert or dichotomous items to measure attitudes across culturally distinct countries in international comparative educational research*. Paper presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.

Walker, M. (2007), Ameliorating Culturally-Based Extreme Response Tendencies To Attitude items, *Journal of Applied Measurement,* No. 8, pp. 267-278.

Warm, T.A. (1989), Weighted Likelihood Estimation of Ability in Item Response Theory, *Psychometrika*, No. 54 (3), pp. 427-450.

Westat (2007), *WesVar® 5.1* Computer software and manual, Rockville, MD: Author (also see *http://www.westat.com/wesvar/*).

Wilson, M. (1994), Comparing Attitude Across Different Cultures: Two Quantitative Approaches to Construct Validity, in M. Wilson (ed.), *Objective measurement II: Theory into practice* (pp. 271-292), Norwood, NJ: Ablex.

Wolter, K.M. (2007), *Introduction to Variance Estimation*. Second edition, Springer: New York.

Wu, M.L., Adams, R.J. & **Wilson, M.R.** (1997), *ConQuest®: Multi-Aspect Test Software* [computer program manual], Camberwell, Vic.: Australian Council for Educational Research.

List of abbreviations – the following abbreviations are used in this report:

ACER	Australian Council for Educational Research
AGFI	Adjusted Goodness-of-Fit Index
BRR	Balanced Repeated Replication
CBAS	Computer Based Assessment of Science
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CITO	National Institute for Educational Measurement, The Netherlands
CIVED	Civic Education Study
DIF	Differential Item Functioning
ENR	Enrolment of 15-year-olds
ESCS	PISA Index of Economic, Social and Cultural Status
ETS	Educational Testing Service
IAEP	International Assessment of Educational Progress
I	Sampling Interval
ICR	Inter-Country Coder Reliability Study
ICT	Information Communication Technology
IEA	International Association for the Evaluation of Educational Achievement
INES	OECD Indicators of Education Systems
IRT	Item Response Theory
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISEI	International Socio-Economic Index
MENR	Enrolment for moderately small school
MOS	Measure of size
NCQM	National Centre Quality Monitor
NDP	National Desired Population
NEP	National Enrolled Population
NFI	Normed Fit Index
NIER	National Institute for Educational Research, Japan
NNFI	Non-Normed Fit Index
NPM	National Project Manager
OECD	Organisation for Economic Cooperation and Development
PISA	Programme for International Student Assessment
PPS	Probability Proportional to Size
PGB	PISA Governing Board
PQM	PISA Quality Monitor
PSU	Primary Sampling Units
QAS	Questionnaire Adaptations Spreadsheet
RMSEA	Root Mean Square Error of Approximation
RN	Random Number
SC	School Co-ordinator
SE	Standard Error
SD	Standard Deviation
SEM	Structural Equation Modelling
SMEG	Subject Matter Expert Group
SPT	Study Programme Table
TA	Test Administrator
TAG	Technical Advisory Group
TCS	Target Cluster Size
TIMSS	Third International Mathematics and Science Study
TIMSS-R	Third International Mathematics and Science Study – Repeat
VENR	Enrolment for very small schools
WLE	Weighted Likelihood Estimates

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

Table of contents

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

LIST OF BOXES

LIST OF FIGURES

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

LIST OF TABLES

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009

PISA 2006 TECHNICAL REPORT – ISBN 978-92-64-04808-9 – © OECD 2009