Chapter 16: SCALING PROCEDURES AND CONSTRUCT VALIDATION OF CONTEXT QUESTIONNAIRE DATA

INTRODUCTION

The PISA-D context questionnaires are based on the questionnaire framework described in Chapter 3 of this report. Many of the questionnaire items were designed to be combined in some way to measure latent constructs that cannot be observed directly, such as students' levels of depression. For these items, scaling procedures were applied to construct meaningful indices.

Three types of scales were developed for PISA-D:

- Basic services at the school include factors such as potable water, sewage services, bathrooms, electricity, and telephones.
- Scales identical to PISA 2015. These scales used sets of items identical to those used in PISA 2015. They enable the PISA-D countries to make direct comparisons of their results to those of the countries that participated in PISA 2015.
- Scales that extended those of PISA 2015. These scales used a subset of items from scales used in PISA 2015 as well as new items relevant to PISA-D countries.
- Scales unique to PISA-D. PISA-D included some new scales that were used for the first time in a PISA study.

This chapter describes the methodology used for scaling and construct validation of the derived variables that underlies these three types of scales.

SCALING METHODOLOGY AND CONSTRUCT VALIDATION

Scaling procedures

As in PISA 2015 and in previous cycles of PISA, the derived variables for PISA-D were constructed using IRT *(*item response theory) (OECD, 2017). The IRT models used in PISA-D are subsets of the generalised partial credit model (Masters and Wright, 1997).

The responses for each item are modelled as a function of the latent construct, θ_j. With a one-parameter model, called the Rasch model (Rasch, 1960), for dichotomous items, the probability of person *j* selecting category 1 instead of 0 is modelled as:

(1)
$$P\left(X_{ji} = 1 \middle| \theta_j, \beta_i\right) = \frac{\exp(\theta_j - \beta_i)}{1 + \exp(\theta_j - \beta_i)}$$

where $P(X_{ji} = 1)$ is the probability of person *j* to score 1 on item *i*; θ_j is the estimated latent trait

of person *j* and β_i is the estimated location or difficulty of item *i* on this dimension. In the case of items with more than two (*m*) categories (e.g., Likert-type items), this model can be generalised to the partial credit model, which takes the form of:

(2)
$$P\big(X_{ji} = k \big| \theta_j, \beta_i, d_i\big) = \frac{\exp(\sum_{r=0}^{k} \theta_j - (\beta_i + d_{ir}))}{\sum_{u=0}^{m_j} exp(\sum_{r=0}^{u} \theta_j - (\beta_i + d_{ir}))}$$

where $P\big(X_{ji} = k\big)$ denotes the probability of person *j* to score *k* on item *i* out of the m_i possible scores (e.g., 1 … 5) on the item. θ_j denotes the person's latent trait, the item parameter β_i gives the general location of the item on the latent continuum, and d_{ir} denote additional step parameters. This model has been used throughout previous cycles of PISA for scaling derived variables of the context questionnaires. However, research literature (especially, Glas and Jehangir, 2014) suggests that a generalisation of this model, the generalised partial credit model (GPCM) (Muraki, 1992), is more appropriate in the context of PISA since it allows for the item discrimination to vary between items within any given scale. This model takes the form of:

(3)
$$P\big(X_{ji} = k \big| \theta_j, \beta_i, d_i\big) = \frac{\exp(\sum_{r=0}^{k} \alpha_i(\theta_j - (\beta_i + d_{ir})))}{\sum_{u=0}^{m_j} exp(\sum_{r=0}^{u} \alpha_i(\theta_j - (\beta_i + d_{ir})))}$$

in which the additional discrimination parameter α_i allows for the items of a scale to contribute with different weights to the measurement of the latent construct.

Following Wu and Adams (2007), "the delta (δ) or d parameters do not reflect the difficulty of achieving a score point in a partial credit item. For partial credit items, to achieve a score of 2, students would generally need to accomplish more tasks than for achieving a score of 1. To reflect this "cumulative achievement", the Thurstonian thresholds are sometimes used as indicators of "score difficulties". The Thurstonian threshold for a score category is defined as the ability at which the probability of achieving *that score or higher* reaches 0.50. (p. 50). Therefore, throughout the report, the Thurstonian thresholds (Wu and Adams, 2007) are provided for each of the scales in which IRT was used.

Values for the β_i, α_i, and d_{ir} parameters are presented in the Appendix of this chapter (Tables 16.A1 – 16.A15).

Construct validation
We assessed the cross-country validity of measures of student background, practices, attitudes, and perceptions in PISA-D following two approaches implemented for context questionnaires in PISA 2015.

Internal consistency. Cronbach's alpha assesses the internal consistency of each scale within the countries and compares it between the countries. The alpha coefficient ranges between 0 and 1, with higher values indicating higher internal consistency. Commonly accepted cut-off values are 0.9 to signify excellent, 0.8 for good, and 0.7 for acceptable internal consistency. Following the approach used in PISA 2015, in PISA-D Cronbach's alpha was used to assess internal consistency

for each scale by country.

Evaluating cross-country comparability of latent constructs. Cross-country validity of the constructs assumes that the same constructs can be measured consistently in different national and cultural contexts. All of the scales and indicators in PISA-D are based on persons' self-reports. Such measures can suffer from various measurement errors, stemming from retrospective reports of behaviour and cultural differences in respondents' beliefs, behaviours, and attitudes (Bempechat, Jimenez, and Boulay, 2002). The literature consistently shows that response biases, such as social desirability, acquiescence, and extreme response choice, are more common in contexts with lower socioeconomic development and socioeconomic status, and that response styles also differ between genders (Buckley, 2009).

Following the approach pioneered in PISA 2015, we estimated international item and person parameters based on all examinees across all seven countries (OECD, 2017). This estimation produced the root mean square deviance (RMSD) item-fit statistic for each country and item as:

(4) $$RMSD = \sqrt{\int \left(P_o(e) - P_e(e)\right)^2 f(e)\,de}$$

quantifying the difference between the *observed item characteristic curve* (*ICC, $P_o(q)$*) with the *model-based ICC* (*$P_e(q)$*)) (OECD, 2017, p. 269).

This statistic indicates the extent of the discrepancy between the observed item characteristic curve (ICC) and the model-based ICC. The RMSD is sensitive to the group-specific deviations of both the item difficulty parameters and item slope parameters from the international parameters. Values close to zero indicate good item fit, meaning that the model with international item parameters describes the responses in this group very well. The theoretical minimum (RMSD=0) indicates perfect fit of the international item parameters for this group. A value of RMSD=0.3 was set as a criterion, with larger values indicating that the international item parameters were not appropriate for this group. The item RMSDs within each country were generally consistent with the average country RMSD. Therefore, we constrained the parameters to be the same across countries. The final distribution of RMSD values across countries for each scale are provided in the Appendix of this chapter (Table 16A.15).

Scales and indices identical to PISA 2015

For the scales identical to PISA 2015, the international item and person parameters were originally obtained from a calibration process based on a GPCM for a single analysis based on data from all persons in all countries. The development of the indices also followed the same process as was used for PISA 2015.

The list of scales and indices is shown in Table 16.1.

As in PISA 2015, in PISA-D, for each scale, only persons with a minimum number of three valid responses were included; students and teachers from the seven countries of PISA-D were unweighted. The calibration of the PISA-D item and person parameters anchored the values of

the 2015 parameters using the TAM R package (Robitzsch, Kiefer, and Wu, 2018). The TAM package produced weighted likelihood estimates (WLEs; Warm, 1989) as individual participant scores. The WLEs obtained with the TAM package were rescaled with a linear transformation to link them to the PISA 2015 scale by subtracting the PISA 2015 unweighted WLE mean from the PISA-D original WLE scores and dividing the difference by the PISA 2015 unweighted WLE standard deviation.

Table 16.1 **Scales identical to PISA 2015**

Derived Variable	Description	Questionnaire	Indices and Scales
MISCED	Mother's education (ISCED)	Student	Index
HISCED	Highest education of parents (ISCED)	Student	Index
FISCED	Father's education (ISCED)	Student	Index
PARED	Highest education of parents in years	Student	Index
BFMJ2	ISEI of father	Student	Index
BMMJ1	ISEI of mother	Student	Index
HISEI	Index highest parental occupational status	Student	Index
BELONG	Sense of belonging to school	Student	IRT Scale
DISCLI	Classroom disciplinary climate class	Student	IRT Scale
SATJOB	Satisfaction with the current job	Teacher	IRT Scale
SATTEACH	Satisfaction with teaching profession	Teacher	IRT Scale
TCLEAD	Teachers' views on school leadership	Teacher	IRT Scale

Educational level of parents (MISCED and FISCED; PARED)

Students' responses on questions ST031Q01TA, ST034Q01TA, ST038Q01TA, and ST041Q01TA regarding parental education were classified using ISCED 1997 (OECD, 1999). Three indices on parental education resulted from the recoding of educational qualifications into the following categories: (0) None, (1) ISCED 1 (primary education), (2) ISCED (International Standard Classification of Education) 2 (lower secondary), (3) ISCED Level 3B or 3C (vocational/pre-vocational upper secondary), (4) ISCED 3A (general upper secondary) and in some cases ISCED 4 (non-tertiary post-secondary), (5) ISCED 5B (vocational tertiary) and (6) ISCED 5A, and in some cases ISCED 6 (theoretically oriented tertiary and post-graduate). The index MISCED indicates the educational level of the mother; FISCED indicates that of the father; and the index of highest educational level of parents (HISCED) corresponds to the higher ISCED level of either parent. PARED is the index of the estimated number of years of education generated from HISCED. In PISA-D, the mapping of ISCED levels to years of schooling (PARED) was done in consultation with

the seven countries, considering the structure of their education systems.

Highest occupational status of parents

In PISA-D, students were asked with open-ended questions about the occupations of their mothers and fathers. The responses were coded to four-digit International Standard Classification of Occupations (ISCO) codes and then mapped to the International Socioeconomic Index of occupational status (ISEI) (Ganzeboom and Treiman, 2003). In PISA-D we adopted the new ISCO and ISEI in their 2008 version. Based on this information, we computed three indexes: father's occupational status (BFMJ2); mother's occupational status (BMMJ1); and the highest occupational status of parents (HISEI), which corresponds to the higher ISEI score of either parent or to the only available parent's ISEI score. For all three indices, higher ISEI scores indicate higher levels of occupational status.

Sense of belonging

In the PISA-D student questionnaire, students were asked about their sense of belonging at school (ST068Q01TA to ST068Q06TA) using six trend items previously used in PISA 2015 and 2012 (Tables 16.2 and 16.3). The response format was a four-point Likert scale with the response categories "strongly agree", "agree", "disagree", and "strongly disagree". The derived IRT scale is named BELONG. The items were coded such that a higher score indicated a positive sense of belonging for all items. The levels of reliability were less than 0.7, as they were for several PISA 2015 countries. The index can be used to describe the average levels of sense of belonging at the national level; however, due to low reliability of this scale, analyses that use this construct in models as a student-level outcome or covariate may have low statistical power.

Table 16.2 **Reliability estimates for sense of belonging, by country**

	Cronbach's Alpha
Cambodia	0.548
Ecuador	0.680
Guatemala	0.646
Honduras	0.683
Paraguay	0.683
Senegal	0.508
Zambia	0.538

Table 16.3 **Item thresholds for sense of belonging**

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST068Q01TA	I feel like an outsider (or left out of things) at school.	-0.852	-0.131	0.956
ST068Q02TA	I make friends easily at school.	-1.795	-0.396	2.193
ST068Q03TA	I feel like I belong at school.	-2.188	-0.309	3.241
ST068Q04TA	I feel awkward and out of place in my school.	-1.021	-0.138	1.144
ST068Q05TA	Other students seem to like me.	-2.497	-0.751	3.453
ST068Q06TA	I feel lonely at school.	-0.669	-0.124	0.636

Disciplinary climate

PISA 2015 focused on science learning in school by including several questions about the learning environment in science classes (Tables 16.4 and 16.5). Students were asked how often specific activities happened in the school science course. The questions included the disciplinary climate in science classes (DISCLISCI). In PISA-D, the same questions were framed generally, rather than focused on science learning. Analyses of the Field Trial data indicated that the subject-specific scale of 2015 and the general scale used in PISA-D yielded comparable results. Therefore, the PISA-D DISCI was scaled using the IRT scaling model with the fixed parameters from PISA 2015.

Table 16.4 **Reliability estimates for classroom disciplinary climate, by country**

	Cronbach's Alpha
Cambodia	0.854
Ecuador	0.804
Guatemala	0.824
Honduras	0.845
Paraguay	0.844
Senegal	0.711
Zambia	0.764

Table 16.5 Item thresholds for classroom disciplinary climate

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST074Q01TA	Students don't listen to what the teacher says.	-1.448	-0.123	2.165
ST074Q02TA	There is noise and disorder.	-1.014	-0.005	1.463
ST074Q03TA	The teacher has to wait a long time for students to quiet down.	-1.212	-0.154	1.335
ST074Q04TA	Students cannot work well.	-2.217	-0.760	1.689
ST074Q05TA	Students don't start working for a long time after the lesson begins.	-1.883	-0.475	1.646

Job satisfaction

The teacher questionnaires used one question (TC033) to ask about teachers' job satisfaction (Tables 16.6 and 16.7). The four-point Likert scale included four response categories: "strongly agree", "agree", "disagree", and "strongly disagree". The derived variable "satisfaction with the current job environment" (SATJOB) was scaled using items TC033Q04TA, TC033Q06TA, TC033Q07TA, and TC033Q08TA.

Table 16.6 Reliability estimates for teacher satisfaction with current job, by country

	Cronbach's Alpha
Cambodia	0.852
Ecuador	0.827
Guatemala	0.791
Honduras	0.790
Paraguay	0.848
Senegal	0.702
Zambia	0.742

Table 16.7 **Item thresholds for teacher satisfaction with current job**

Item	Description	Threshold 1	Threshold 2	Threshold 3
TC033Q04TA	I enjoy working at this school.	-1.416	-0.227	1.843
TC033Q06TA	I would recommend my school as a good place to work.	-1.361	0.059	2.273
TC033Q07TA	I am satisfied with my performance in this school.	-3.325	-1.388	3.336
TC033Q08TA	All in all, I am satisfied with my job.	-2.764	-1.033	2.682

Satisfaction with the teaching profession

The derived variable "satisfaction with teaching profession" (SATTEACH) was scaled using items TC033Q01TA, TC033Q02TA, TC033Q03TA, and TC033Q05TA (Tables 16.8 and 16.9). The four-point Likert scale included four response categories: "strongly agree", "agree", "disagree", and "strongly disagree".

Table 16.8 **Reliability estimates for teacher satisfaction with the profession**, by country

	Cronbach's Alpha
Cambodia	0.517
Ecuador	0.739
Guatemala	0.632
Honduras	0.724
Paraguay	0.722
Senegal	0.630
Zambia	0.691

Table 16.9 **Item thresholds for teacher satisfaction with the profession**

Item	Description	Threshold 1	Threshold 2	Threshold 3
TC033Q01TA	The advantages of being a teacher clearly outweigh the disadvantages.	-3.044	-0.641	3.727
TC033Q02TA	If I could decide again, I would still choose to work as a teacher.	-1.006	-0.041	1.227
TC033Q03TA	I regret that I decided to become a teacher.	-1.304	-0.427	0.913
TC033Q05TA	I wonder whether it would have been better to choose another profession.	-1.675	0.594	2.413

Teachers' views on school leadership

TC032 asked about teachers' views on school leadership (TCLEAD) (Tables 16.10 and 16.11). The derived variable "teachers' views on school leadership" (TCLEAD) was scaled using items TC032Q01TA, TC032Q02TA, TC032Q03TA, TC032Q04TA, and TC033Q05TA. The four-point Likert scale included four response categories: "strongly agree", "agree", "disagree", and "strongly disagree".

Table 16.10 **Reliability estimates for teachers' views on school leadership, by country**

	Cronbach's Alpha
Cambodia	0.800
Ecuador	0.893
Guatemala	0.906
Honduras	0.905
Paraguay	0.876
Senegal	0.841
Zambia	0.831

Table 16.11 **Item thresholds for teachers' views on school leadership**

Item	Description	Threshold 1	Threshold 2	Threshold 3
TC032Q01TA	The principal tries to achieve consensus with all staff when defining priorities and goals in school.	-2.068	-0.441	2.212
TC032Q02TA	The principal is aware of my needs.	-2.055	-0.160	2.597
TC032Q03TA	The principal inspires new ideas for my professional learning.	-2.777	-0.019	3.825
TC032Q04TA	The principal treats teaching staff as professionals.	-2.696	-1.214	1.891
TC032Q05TA	The principal ensures our involvement in decision making.	-1.488	0.052	2.132

Scales that extended those of PISA 2015

Two of the scales administered in PISA 2015 were extended with additional items in PISA-D. They are household possessions, HOMEPOS15, and Educational, Social, and Cultural Status, ESCS. The scaled scores for these constructs in PISA-D allow for linking the PISA-D scores with those of PISA 2015 because of a common calibration linking procedure, which consists of two phases: joint calibration and linking transformation.

The joint calibration phase produced international item and person parameters using a generalised partial credit model (see equation 2) in a single analysis and based on joint data, comprised of a 5% simple random sample of unweighted persons in all countries from PISA 2015 and all unweighted persons in all countries from PISA-D. We fixed the values of the common items to the PISA 2015 parameters and let the TAM package generate the parameters of the new items. For each scale, only persons with a minimum number of three valid responses were included. We conducted additional analyses on the invariance of item parameters across PISA-D countries and considered assigning unique parameters if necessary (see the section on "Cross-country Comparability" in this chapter). From this concurrent calibration, we derived WLEs for the 5% sample from PISA 2015 and for PISA-D.

In the linking phase, the PISA-D WLEs obtained in the calibration phase were initially standardised and then linked to the 2015 metric by a linear transformation, subtracting the PISA 2015 mean and dividing by the PISA 2015 standard deviation, both obtained from data from all persons in 2015.

Household possessions

In PISA 2015, students reported the availability of 16 household items at home (ST011), including three country-specific household items that were seen as appropriate measures of family wealth within the country's context. In addition, students reported the amount of possessions and books at home (ST012, ST013). PISA-D included 14 items from PISA 2015. In addition, the scale was

extended to include 4 new items. Further, the question pertaining to books in the home was replaced with a question that was more appropriate for PISA-D countries (ST066Q01NA). The resulting measure, HOMEPOS15, included 19 items, which are shown in Table 16.12.

The HOMEPOS15 scale was constructed in two steps. In the first step, international item parameters for PISA 2015 items administered also in PISA-D were obtained from a concurrent calibration of the 2015 and PISA-D data. This step is identical with the regular scaling of HOMEPOS in PISA 2015. In the second step, items from PISA-D were scaled with the parameters fixed for all items administered also in 2015 and for which no unique (i.e., country-specific) item parameters were necessary. Item parameters for all other items (except national items) were freely estimated but constrained to be equal across countries. Once this process was finished, we estimated WLEs for all students from PISA-D. By restricting the largest subset of items to be equal across PISA-D and PISA 2015, the HOMEPOS15 scores can be regarded to be on a joint scale, allowing for comparisons of countries across PISA-D and PISA 2015. This also facilitates the calculation of a measure of ESCS for PISA-D that can be compared with the ESCS measure used in PISA 2015. For more information, see Tables 16.13 and 16.14.

Table 16.12 **Measures of household possessions**

Item	Description	HOMEPOS15	HOMEPOS
ST062Q01TA	A desk to study at	✓	✓
ST062Q02TA	A room of your own	✓	✓
ST062Q03TA	A quiet place to study	✓	✓
ST062Q04TA	A computer you can use for school work	✓	✓
ST062Q05TA	Educational software	✓	✓
ST062Q06TA	A link to the Internet	✓	✓
ST062Q10TA	Books of to help you with your school work	✓	✓
ST062Q12TA	A dictionary	✓	✓
ST064Q01NA	A table to have meals	✓	
ST064Q03NA	A washer	✓	
ST064Q04NA	A refrigerator or freezer	✓	
ST064Q06NA	A stove or burner for cooking	✓	
ST063Q01TA	Televisions	✓	✓
ST063Q02TA	Cars	✓	✓
ST063Q03TA	Rooms with a bath or shower	✓	✓
ST063Q04TA	Cellphones with internet access (smartphones)	✓	✓
ST063Q05TA	Computers	✓	✓
ST063Q06TA	Musical instruments	✓	✓
ST066Q01NA	Number of books in the home	✓	
	Number of books in the home (PISA 2015 version)		✓

Table 16.13 **Reliability estimates for household possessions, by country**

	Cronbach's Alpha
Cambodia	0.853
Ecuador	0.820
Guatemala	0.841
Honduras	0.860
Paraguay	0.821
Senegal	0.851
Zambia	0.848

Table 16.14 **Item thresholds for household possessions**

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST062Q01TA	A desk to study at	-1.170		
ST062Q02TA	A room of your own	-1.160		
ST062Q03TA	A quiet place to study	-1.932		
ST062Q04TA	A computer you can use for school work	-0.326		
ST062Q05TA	Educational software	0.392		
ST062Q06TA	A link to the Internet	-0.569		
ST062Q10TA	Books of to help you with your school work	-2.346		
ST062Q12TA	A dictionary	-1.951		
ST063Q01TA	Televisions	-2.386	-0.307	0.750
ST063Q02TA	Cars	-0.276	0.719	1.675
ST063Q03TA	Rooms with a bath or shower	-1.461	0.588	1.692
ST063Q05TA	Computers	-1.307	-0.499	-0.132
ST063Q06TA	Musical instruments	-0.394	0.388	0.952
ST064Q01NA	A table to have meals	-1.587		
ST064Q03NA	A washer	-0.366		
ST064Q04NA	A refrigerator or freezer	-1.058		
ST064Q06NA	A stove or burner for cooking	-1.533		
ST066Q01NA	Number of books in the home	-2.915	0.122	2.000

Economic, social, and cultural status

In PISA-D the PISA index of economic, social, and cultural status (ESCS) was computed with the same procedure applied to the other PISA cycles.

ESCS is a composite score derived from a principal component analysis (PCA) of three derived measures: parental education (PARED), highest parental occupation (HISEI), and home possessions (HOMEPOS15). These three variables represent the traditional components of socioeconomic status: education, occupational status, and income. In the absence of a direct measure of household income, the presence of material possessions and assets in the household are used as a proxy for family wealth (Willms and Tramonte, in press).

Computation of ESCS in PISA-D

For students with missing data on one out of the three components, a regression imputation based on data for the other two variables were used to predict the scores for variable with missing data. The prediction included the addition of a random component to the predicted value. Missing data on more than one component resulted in a missing value for the ESCS score. As the goal of the computation of ESCS was to link it to the index in PISA 2015, after imputation, all three components were standardised for PISA-D countries based on the OECD means and standard deviations from the PISA 2015 dataset.

In PISA 2015, standardised variables, including imputed values, were entered in the PCA to obtain ESCS values. As in previous cycles, ESCS was the component score for the first principal component. The PCA ran across equally weighted countries, including OECD as well as partner countries and economies. Thus, all countries and economies contribute equally to the estimation of ESCS scores. The ESCS scale was scaled to have a mean of zero and a standard deviation of one at the student level for all OECD countries, weighted equally. In PISA-D, the factor loadings and first eigenvalue from the PCA analysis of PISA 2015 were used to calculate ESCS. Figure 1 displays the components used in the calculation. The elements comprising the figure for the parents' education and occupation are described above.

■ Figure 16.1 ■

Computation of ESCS in PISA-D
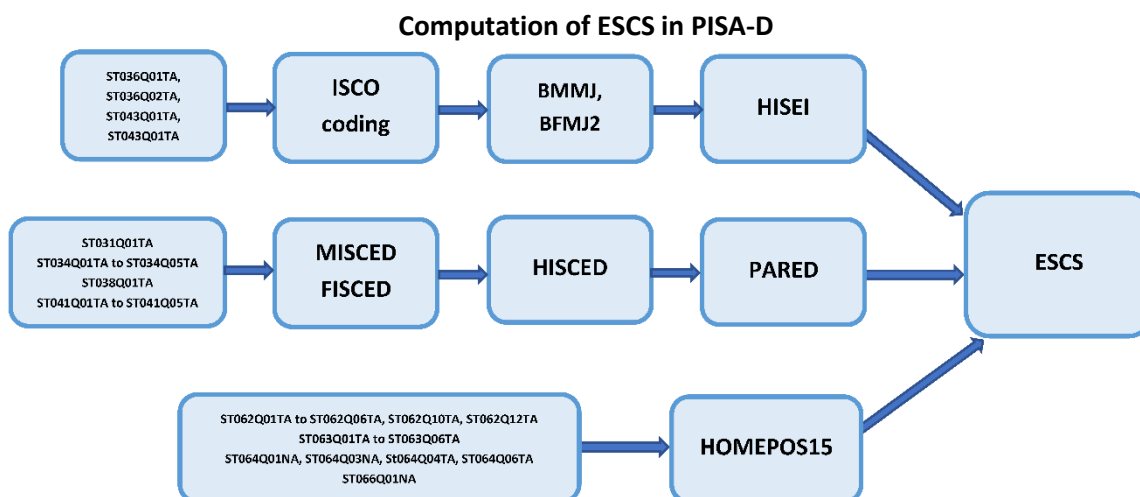
Scales and indices unique to PISA-D

PISA-D included 13 scales that were unique to PISA-D. These are shown in Table 16.15. For the scales unique to PISA-D, international item and person parameters were obtained from a GPCM (see equation 3) in a single analysis based on data from all students in the seven PISA-D countries using the TAM package in R. For each scale, only persons with a minimum number of three valid responses were included. Respondents were left unweighted, and all countries contributed equally to the estimation. Additional analyses were conducted to assess the invariance of item parameters across countries and evaluate the possibility to assign unique parameters in cases of severe misfit (see the section on "Cross-country Comparability" in this chapter).

The WLE obtained from this process represented individual participant scores. We transformed the WLEs into an international PISA-D metric ranging from 0 to 10, where 0 indicates that all the item responses were negative (or corresponding to the least endorsed option), and 10 indicates that all the item responses were positive (or corresponding to the most endorsed option). In four cases, categorical variables based on critical cut-offs on the WLE continuum were calculated.

Table 16.15 **Scales unique to PISA-D**

Derived Variable	Description	Questionnaire	Indices and Scales
DEPRESSION	Depression	Student	IRT Scale
DEPRECAT	Levels of depression	Student	Index
FAMRES	Family resources	Student	IRT Scale
POVERTY	Household poverty index	Student	Index
ATSCH	Attitudes towards school	Student	IRT Scale
ATTAINMENT	Student attainment	Student	Index
STTCHREL	Supportive student teacher relationships	Student	IRT Scale
TCEXPSUC	Teacher expectations for success	Student	IRT Scale
STRLSMAT	Structured lessons in mathematics	Student	IRT Scale
INSTRRES	Instructional resources	Student	IRT Scale
INSTRRESCAT	Levels of instructional resources	Teacher	Index
SCHMATRES	Basic school infrastructure	Teacher	YES
SCHRESOURCES	Levels of school resources	Teacher	Index

Depression

In the PISA-D student questionnaire, students were asked six questions about their mental health using six items from The Learning Bar's OurSCHOOL Survey. Students were asked to consider their feelings at home and at school, and how often they occurred, with the responses including four categories: "never of almost never", "about once a week", "2 to 3 times a week", and "almost every day". Higher WLEs and higher difficulty correspond to higher levels of depression on all items. Once transformed on a 0-10 metric, the variable DEPRESSION indicated that students who scored 0 did not report any signs of depression, while those who scored 10 had reported the most signs of depression (Tables 16.16 and 16.17). An ordinal variable, DEPRCAT, with three categories was calculated from the continuous variable. Its values are: 0 - not depressed (DEPRESSION scores lower than 4.31); 1 - moderately depressed (DEPRESSION scores ranging from 4.31 to 5.8); and 3 – depressed (DEPRESSION scores greater than 5.8).

Table 16.16 **Reliability estimates for depression, by country**

	Cronbach's Alpha
Cambodia	0.754
Ecuador	0.732
Guatemala	0.741
Honduras	0.754
Paraguay	0.770
Senegal	0.669
Zambia	0.743

Table 16.17 **Item thresholds for depression**

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST017Q06NA	I cry without a good reason.	1.511	1.903	2.382
ST017Q07NA	I feel lonely.	0.711	1.229	1.671
ST017Q08NA	Other students seem to have more fun than me.	0.784	1.458	1.983
ST017Q09NA	I feel sad or depressed.	0.526	1.226	1.826
ST017Q10NA	I have trouble falling asleep at night.	0.746	1.526	2.346
ST017Q11NA	A lot of things seem to bother me	-0.050	0.904	1.611

Family resources and poverty

The PISA-D student questionnaire included several items pertaining to personal and material possessions in the home as well items about the infrastructure of the home (Tables 16.18

and 16.19). Other than the trend items and the new items that constitute HOMEPOS15, students were also asked about several home possessions that were more closely related to living in poverty. These included: whether they shared a toilet facility with other people who were not members of their household; whether they had a flush toilet; what the material of the floor was in their home; whether any of the household members had a bank account; and whether the student experienced hunger in the previous month. The measure of family resources, FAMRES, was based on the WLE of the index of family resources. It was recoded on a 0-10 metric, with 0 corresponding to a complete lack of resources in the home and 10 indicating the presence of all family resources.

A measure of household poverty, POVERTY, was derived from FAMRES. It includes four categories: "extremely poor", "severely poor", "poor", and "not poor". The classification was based on three cut-off points on the WLE scale after ordering the items by their difficulty score. Students with a WLE score below -3.5 were considered to be in "extreme poverty". These students would likely have rudimentary flooring in their home and were sharing a toilet facility with others who were not members of their family. Students with a WLE score at or above -3.5, but less than -1.8, were considered to be in "poverty". These students would likely have indicated that they had been hungry in the past 30 days and did not have a flush toilet in their home. Students with a WLE score at or above -1.8, but less than -0.75, were considered to be "poor". These students would likely have indicated that they had a flush toilet in their home but did not have a washing machine. Students with a WLE score at or above -0.75 were considered "not poor". They would likely have a washing machine in the home, indicating that they had running water and electricity. The majority of these students would also have a computer they could use for school work. Further details are provided in Tramonte and Willms (in press; see also Willms, Tramonte, Duarte, and Bos, 2012).

Table 16.18 **Reliability estimates for family resources, by country**

	Cronbach's Alpha
Cambodia	0.851
Ecuador	0.834
Guatemala	0.844
Honduras	0.860
Paraguay	0.832
Senegal	0.861
Zambia	0.862

Table 16.19 **Item thresholds for family resources (sorted by Threshold 1)**

Item	Description	Threshold 1	Threshold 2	Threshold 3	
ST066Q01	Number of books in the home	-4.436	0.097	2.898	Extreme Poverty
ST049Q01	Shared toilet facility	-3.630			
ST051Q01	Floor composition	-3.467	-0.398		
ST062Q10	Books to help with your school work	-3.418			Severe Poverty
ST063Q01	Televisions	-3.170	-0.504	0.775	
ST062Q03	A quiet place to study	-3.089			
ST059Q01	Hungry	-2.995			
ST062Q12	A dictionary	-2.583			
ST064Q01	A table to have meals	-2.289			
ST064Q06	A stove or burner for cooking	-2.288			
ST063Q03	Rooms with a bath or shower	-2.038	0.448	1.644	
ST063Q04	<Cell phones> with internet access (e.g., smartphones)	-1.965	-0.767	-0.244	
ST062Q01	A desk to study at	-1.783			Poor
ST062Q02	A room of your own	-1.753			
ST048Q01	Flush toilet	-1.699			
ST064Q04	A refrigerator or freezer	-1.571			
ST057Q01	Bank account	-1.429			
ST062Q06	A link to the internet	-0.776			
ST064Q03	A washer	-0.603			Not Poor
ST063Q05	Computers (desktop computer, portable laptop, or notebook)	-0.458	0.683	1.468	
ST062Q04	A computer you can use for school work	-0.423			
ST063Q02	Cars, vans or trucks	-0.217	1.116	2.182	
ST062Q05	Educational software	0.384			

Attitudes towards school

In the student questionnaire of PISA-D, students were asked about the impact that school had on their lives (Tables 16.20 and 16.21). In particular, they reported their level of agreement of the importance of trying hard at school, on how school helped them with their confidence in making decisions, on whether the things learned at school would be useful in a prospective job, and on whether schooling gave them a better chance to get a good job or be accepted at a good university. The response categories were a four-point Likert scale, ranging from strongly disagree to strongly agree. The WLE scores for the construct were recoded on a 0-10 continuum, where 0 corresponds to a very negative attitude towards valuing schooling outcomes, and 10 corresponds to full endorsement of all the items related to valuing school outcomes. The derived variable is called ATSCH.

Table 16.20 **Reliability estimates for attitudes towards school, by country**

	Cronbach's Alpha
Cambodia	0.831
Ecuador	0.782
Guatemala	0.798
Honduras	0.839
Paraguay	0.791
Senegal	0.754
Zambia	0.871

Table 16.21 **Item thresholds for attitudes towards school**

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST067Q03TA	School has helped give me confidence to make decisions.	-3.377	-2.604	0.137
ST067Q04TA	School has taught me things which could be useful in a job.	-2.786	-2.367	-0.706
ST067Q05TA	Trying hard at school will help me get a good job.	-2.575	-2.160	-0.718
ST067Q06TA	Trying hard at school will help me get into a good <university>.	-2.588	-2.190	-0.750
ST067Q07TA	I enjoy receiving good <grades>.	-2.825	-2.551	-0.864
ST067Q08TA	Trying hard at school is important.	-2.543	-2.292	-0.761

Attainment

The measure of attainment was derived from data collected with the student tracking form on the age and grade of the student. A student was considered to be "on track" if he or she was in

the modal grade or in the grade above the modal grade, corresponding to his or her age as of 31 December 2016. For example, depending on a country's rules for entry into lower primary, students who were 15 years, 4 months would be "on track" if they were in grade 8, while students who were 15 years, 5 months would be in grade 9 if they were on track. A categorical variable was constructed separately for each country with the following categories: "On track"; "One year behind track"; and "Two or more years behind track". The analysis accounted for the two different school-entry dates for Ecuador. This approach proves to be much more accurate than estimates based on students' reports of their grade and whether or not they had repeated a grade.

Supportive student teacher relationships and teacher expectations for success

Students responded to a series of questions related to their perception of the teachers at their school. The response categories were a four-point Likert scale, ranging from strongly disagree to strongly agree. Supportive student-teacher relationships is a scale produced by the relative agreement of students to a series of questions about their interpersonal relationships with their teachers: on their perception of teacher interest in students' well-being, whether teachers to listen to students and respect their opinions, willingness to help them if in need,; fairness in treatment, and interest in students' learning (Tables 16.22 and 16.23).

Teacher expectations for success derives from a set of questions asked of students concerning their perceptions of teachers' expectations for student success, and in particular teachers' expectations for students to work hard, do their best, and complete homework on time.

Both scales, STTCHREL and TCEXPSUC, were derived from WLE scores and scaled on a 0 to 10 continuum, with 0 indicating reports of very poor relationships and low expectations, and 10 indicating very high positive teacher-student relations and high expectations from their teachers (Tables 16.24 and 16.25).

Table 16.22 **Reliability estimates for teacher-student relationships, by country**

	Cronbach's Alpha
Cambodia	0.833
Ecuador	0.866
Guatemala	0.875
Honduras	0.888
Paraguay	0.874
Senegal	0.746
Zambia	0.790

Table 16.23 **Item thresholds for teacher-student relationships**

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST072Q01NA	I get along well with most of my teachers.	-2.497	-1.901	0.054
ST072Q02NA	Most of my teachers are interested in my well-being.	-2.313	-1.530	0.362
ST072Q03NA	Most of my teachers listen to what I have to say.	-2.230	-1.399	0.332
ST072Q04NA	If I need extra help, I will receive it from my teachers.	-2.259	-1.466	0.284
ST072Q05NA	Most of my teachers treat me fairly.	-2.234	-1.517	0.280
ST072Q06NA	The teachers show an interest in every student's learning.	-2.165	-1.668	-0.183
ST072Q07NA	The teachers give students an opportunity to express opinions.	-2.214	-1.714	-0.180

Table 16.24 **Reliability estimates for expectations for success, by country**

	Cronbach's Alpha
Cambodia	0.814
Ecuador	0.784
Guatemala	0.787
Honduras	0.828
Paraguay	0.799
Senegal	0.682
Zambia	0.790

Table 16.25 **Item thresholds for expectations for success**

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST072Q08NA	Our teachers expect us to work hard.	-2.367	-2.026	-0.436
ST072Q09NA	Our teachers encourage students to do their best work.	-2.418	-2.043	-0.512
ST072Q10NA	Our teachers expect us to do our homework on time.	-2.318	-1.952	-0.391
ST072Q11NA	Students understand what is expected of them for their <courses>.	-2.607	-2.038	0.102

Structured lessons in mathematics

The question on structured lessons in mathematics is based on Anderson's (2004) model of teacher effectiveness and a broader literature emphasising the importance of delivering structured lessons. In the student questionnaire of PISA-D, students were asked to think about their lessons in mathematics and report on how often teachers' practices were evident in their lessons. The response categories were: "every lesson", "most lessons", "some lessons", and "none or hardly ever". The question considered the key elements of a structured lesson at the beginning, during, and after the lesson. Students were asked how often teachers explained the purpose of the lesson and offered a review of what had been done previously at the beginning of the lesson. For the section concerning the body of the lesson, students were asked to report on the frequency of typical practices, such as offering examples of problem solving and successful work, giving clear answers to questions, giving a formal lecture on a topic, explaining mathematical concepts, giving work to do at their desk, and talking to students about their work. For the end of the lesson, students were asked how often their teachers formally closed a lesson, with a summary of the lesson and with homework to practice what was learned in class (Tables 16.26 and 16.27).

The WLEs were rescaled on a 0-10 continuum, with 0 corresponding to patterns of ordered items in which students systematically reported that their teachers never followed any of the steps of a structured lesson, while a score of 10 corresponding to patterns of ordered items in which teachers consistently used the practices associated with a structured lesson.

Table 16.26 **Reliability estimates for structured lessons in mathematics, by country**

	Cronbach's Alpha
Cambodia	0.890
Ecuador	0.886
Guatemala	0.884
Honduras	0.884
Paraguay	0.862
Senegal	0.798
Zambia	0.851

Table 16.27 **Item thresholds for structured lessons in mathematics**

Item	Description	Threshold 1	Threshold 2	Threshold 3
ST075Q01NA	The teacher explains the purpose of the lesson.	-2.369	-1.365	-0.760
ST075Q02NA	The teacher reviews what we learned in previous lessons.	-2.276	-1.092	-0.259
ST075Q03NA	The teacher shows us how to solve problems.	-2.269	-1.459	-0.832
ST075Q04NA	The teacher provides examples of successful work.	-2.071	-1.284	-0.610
ST075Q05NA	The teacher gives clear answers to students' questions.	-2.100	-1.267	-0.583
ST075Q06NA	The teacher gives a formal lecture on the topic.	-2.107	-1.248	-0.429
ST075Q07NA	The teacher explains mathematical concepts.	-2.248	-1.317	-0.619
ST075Q08NA	The teacher gives us work to do at our desk.	-2.576	-1.434	-0.580
ST075Q09NA	The teacher talks with students about their work.	-2.125	-1.116	-0.283
ST075Q11NA	The teacher summarises what we have done that day.	-2.012	-0.816	0.038
ST075Q12NA	The teacher gives us homework to practise what we have learned.	-3.062	-1.402	-0.382

Instructional resources and levels of instructional resources

PISA-D used a schema set out by Murillo and Román (2011) that distinguishes between basic services, didactic facilities, and didactic resources:

- Basic services at the school include factors such as potable water, sewage services, bathrooms, electricity, and telephones.

- Didactic facilities refer to places other than the classroom for teaching and learning. These include, for example, school libraries, gymnasiums, art and music rooms, science laboratories, computer rooms, and sports fields.

- Didactic resources can include very basic materials such as textbooks and blackboards as well as computers in the school, laptop computers for students and teachers, and quality books in the library.

The questions in PISA-D consider both the availability of instructional resources and teachers' use of them. Data from the teacher questionnaire regarding the availability of resources were used to construct a continuous measure of instructional resources, INSTRRES, with the WLEs rescaled onto a 0-10 continuum. In addition, an ordinal measure with five categories, INSTRRESCAT, was constructed based on specific cut points on the WLE continuum. A very low level of instructional resources, Level 1, corresponds to a pattern or ranked items that shows the presence of only very

basic resources, with computers for administrators (WLE scores lower than -1.35) being unlikely. A low level of instructional resources, or Level 2, indicates schools in which basic resources are available to the teacher except for a library (WLE score between -1.35 and -0.80). An adequate level of instructional resources, or Level 3, corresponds to a pattern of resources in which teachers endorse up to having access to a photocopier. A moderately high level of instructional resources, Level 4 (WLE greater than -.40 and lower than 0.24), corresponds to having basic and more sophisticated didactic resources, including a science lab. Finally, a high level of instructional resources (WLE greater than 0.24), Level 5, corresponds to having more complex and expensive didactic resources in the school (Tables 16.28 and 16.29). Further details are provided in Willms and Tramonte (in press; see also Willms, Tramonte, Duarte and Bos, 2012).

Table 16.28 **Reliability estimates for instructional resources, by country**

	Cronbach's Alpha
Cambodia	0.923
Ecuador	0.922
Guatemala	0.950
Honduras	0.951
Paraguay	0.940
Senegal	0.882
Zambia	0.919

Table 16.29 **Item thresholds for instructional resources (sorted by Threshold 2)**

Item	Description	Threshold 1	Threshold 2	Threshold 3	
TC017Q04NA	Chalk (or other markers)	-3.188	-2.657	-1.818	Level 1
TC017Q03NA	Writing board (black, white, green)	-4.111	-2.453	-0.533	Level 1
TC017Q01NA	Chairs for students	-2.227	-1.587	0.322	Level 1
TC017Q10NA	Reading, mathematics, or science textbooks	-2.069	-1.537	-0.806	Level 1
TC017Q02NA	Desks for students	-2.198	-1.534	0.485	Level 1
TC035Q15NA	School administrative office	-1.960	-1.494	-0.610	Level 1
TC035Q05NA	Computers for administrative use	-1.806	-1.320	-0.284	Level 2
TC017Q12NA	Teacher's guide	-1.466	-1.177	-0.596	Level 2
TC017Q17NA	Teacher table and chair	-1.942	-0.969	0.331	Level 2
TC017Q11NA	Reference books for teachers	-1.289	-0.955	-0.245	Level 2

Item	Description	Threshold 1	Threshold 2	Threshold 3	
TC017Q09NA	Dictionary	-1.261	-0.847	-0.116	
TC017Q05NA	A wall chart, map or diagram	-1.383	-0.773	0.752	
TC017Q13NA	School library	-1.213	-0.737	0.076	
TC017Q08NA	Work sheets	-0.735	-0.605	-0.143	
TC017Q07NA	Workbooks	-0.756	-0.561	0.013	Level 3
TC035Q06NA	Computer room	-0.816	-0.524	0.217	
TC035Q16NA	Storage room	-1.126	-0.498	0.314	
TC035Q14NA	Teacher staff room	-0.823	-0.425	0.339	
TC035Q01NA	Computers for students	-0.740	-0.315	0.828	
TC035Q08NA	Photocopier	-0.743	-0.284	0.722	
TC035Q09NA	Overhead or slide projector	-0.499	-0.228	0.419	
TC017Q18NA	Room for student guidance or counselling	-0.187	-0.015	0.563	Level 4
TC035Q13NA	Telephone line	-0.046	0.100	0.412	
TC017Q06NA	One or more bookshelves	-0.244	0.212	1.680	
TC035Q10NA	Audio or video disk players (e.g., CD. DVD, or VCD)	0.102	0.309	0.911	
TC035Q07NA	Science lab	-0.017	0.336	1.180	
TC017Q19NA	<Education resource centre>	0.324	0.455	0.887	
TC035Q12NA	TV or screens	0.173	0.484	1.163	
TC019Q01NA	Reading textbooks	0.365	0.505	0.942	
TC035Q04NA	Internet connection for teachers	0.278	0.596	1.203	
TC035Q03NA	Computers for teachers	0.386	0.637	1.467	
TC035Q02NA	Internet connection for students	0.499	0.735	1.234	Level 5
TC017Q20NA	<Area for productive projects>	0.541	0.832	1.548	
TC017Q14NA	Gym	0.668	0.872	1.680	
TC019Q01NA	Mathematics textbooks	0.606	0.887	1.515	
TC035Q11NA	Radio	0.944	1.158	1.787	
TC017Q15NA	Music room	1.400	1.498	1.908	
TC017Q16NA	Art room	1.522	1.638	2.110	

Basic school infrastructure and levels of school resources

Following Murillo and Román (2011), in PISA-D, school administrators and principals were asked to report on the presence and condition of infrastructural resources. In the school questionnaire, they were asked whether basic and advanced infrastructural features and facilities were available and in good condition. For example, they responded to an array of questions pertaining to the condition of the school roof, walls, floors, entrance doors, windows, hallways, classrooms, toilets, kitchen, and whether there was drinking water, running water, electricity, indoor plumbing, a first aid room, a healthcare room, a cafeteria, a sports area, a fence, an access ramp, fans, lighting, gender-specific toilets, staff toilets, and textbooks. The response categories were: "no, not available", "yes, but in poor condition", "yes, but in need of minor repairs", and "yes, in good condition".

The WLEs for the construct were rescaled onto a 0 to 10 scale. The variable, SCHMATRES, indicates the level of school resources. An ordinal variable, SCHRESOURCES, was constructed based on the patterns of responses to the ranked items. It has five levels: Level 1 (WLE lower than -1.44) corresponds to very low levels of basic resources, from bare minimum up to access to running water in the school; Level 2 (WLE ranging from -1.44 to -1.054) indicates low infrastructural resources, up to having flush toilets in the school; Level 3 (WLE ranging from -1.053 to -0.620) indicates adequate resources, up to having a cafeteria; Level 4 (WLE ranging from -0.619 to 0.270) indicates moderately high levels of resources, such as a school having a cafeteria but not a kitchen; and Level 5 (WLEs greater than 0.270) indicates a high level of infrastructural resources (Tables 16.30 and 16.31). Further details are provided in Willms and Tramonte (in press; see also Willms, Tramonte, Duarte and Bos, 2012).

Table 16.30 **Reliability estimates for school infrastructure, by country**

	Cronbach's Alpha
Cambodia	0.883
Ecuador	0.935
Guatemala	0.934
Honduras	0.953
Paraguay	0.909
Senegal	0.906
Zambia	0.907

Table 16.31 **Item thresholds for school infrastructure (sorted by Threshold 2)**

Item	Description	Threshold 1	Threshold 2	Threshold 3	
SC011Q01NA	Roof	-3.430	-2.118	-0.566	Level 1
SC011Q02NA	Walls	-2.769	-2.113	-0.889	
SC011Q08NA	Classrooms	-2.399	-1.900	-0.368	
SC011Q04NA	Building entrance	-2.411	-1.845	-0.654	
SC011Q05NA	Doors	-2.729	-1.814	-0.292	
SC011Q07NA	Hallways	-2.075	-1.790	-0.737	
SC011Q06NA	Windows	-2.807	-1.784	-0.279	
SC011Q03NA	Floors	-2.290	-1.740	-0.373	
SC012Q06NA	Electricity	-2.148	-1.615	-0.760	
SC015Q01NA	Math textbooks	-4.784	-1.608	0.014	
SC013Q01NA	Separate toilets for girls and boys	-2.474	-1.561	-0.220	
SC013Q02NA	Separate toilets for school staff and students	-1.931	-1.462	-0.057	
SC012Q05NA	Running water	-1.710	-1.447	-0.729	Level 2
SC012Q11NA	Sports area or playground	-2.387	-1.382	0.150	
SC014Q01NA	Reading textbook	-4.101	-1.352	-0.101	
SC012Q04NA	Place with drinkable water	-1.727	-1.344	-0.353	
SC012Q15NA	<Lighting>	-1.745	-1.280	-0.233	
SC012Q12NA	Fence or hedge on the school borders	-1.780	-1.119	-0.077	
SC012Q01NA	Flush toilets	-1.396	-0.991	0.232	Level 3
SC012Q07NA	Indoor plumbing	-1.203	-0.906	0.233	
SC012Q14NA	<Fans>	-0.622	-0.300	0.602	
SC012Q10NA	Cafeteria	-0.164	0.014	0.690	Level 4
SC012Q13NA	<Access ramp>	-0.166	0.052	1.003	
SC012Q03NA	Kitchen	0.199	0.416	1.282	Level 5
SC012Q08NA	<First aid room>	0.771	0.850	1.425	
SC012Q09NA	Immunisation or health care room	1.443	1.492	1.823	

References

Anderson, L. W. (2004), *Increasing Teacher Effectiveness (2nd ed.),* UNESCO International Institute for Educational Planning, Paris.

Bempechat, J., N. V. Jimenez and **B. A. Boulay** (2002), "Cultural-cognitive issues in academic achievement: New directions for cross-national research", in A. C. Porter and A. Gamoran (eds.), *Methodological Advances in Cross-National Surveys of Educational Achievement, National Academic Press*, Washington.

Buckley, J. (2009), *Cross-National Response Styles in International Educational Assessments: Evidence from PISA 2006*, https://edsurveys.rti.org/PISA/documents/ Buckley_PISAresponsestyle.pdf (accessed on 2 November 2016).

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

Glas, C. A. W. and **K. Jehangir** (2014), "Modeling country-specific differential item functioning", in L. Rutkowski, M. von Davier and D. Rutkowski (eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues and Methods of Data Analysis,* pp. 97-115, Springer, New York.

Masters, G. N. and **B. D. Wright** (1997), "The partial credit model", in W. J. van der Linden and R. K. Hambleton (eds.), *Handbook of Modern Item Response Theory*, Springer, New York, NY.

Murillo, F. J., and M. Román (2011), "School infrastructure and resources do matter: Analysis of the incidence of school resources on the performance of Latin American students", *School Effectiveness and School Improvement*, Vol. 22, pp. 29-50.

Muraki, E. (1992), *A Generalized Partial Credit Model: Application of an EM Algorithm*, ETS Research Report Series, Vol. 1992/1, pp. i-30.

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

OECD (2017), *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264281820-en.

Rasch, G. (1960), *Probabilistic Models for Some Intelligence and Attainment Tests*, Copenhagen, Denmark: Nielsen and Lydiche (Expanded edition, Chicago, University of Chicago Press, 1980).

Robitzsch, A., T. Kiefer and M. Wu (2018), TAM: Test analysis modules. R package version

2.12-18. https://CRAN.R-project.org/package=TAM.

Tramonte, L. and J. D. Willms (in press). "A measure of poverty for educational studies in low- and middle-income countries", *Research Note*, The Learning Bar Inc., Fredericton, New Brunswick.

Warm, T. A. (1989), Weighted likelihood estimation of ability in item response theory, *Psychometrika*, Vol. 54, pp. 427-450.

Willms, J. D. and L. Tramonte (in press). "Measures of classroom and school material resources for educational studies in low- and middle-income countries", *Research Note,* Fredericton, NB: The Learning Bar Inc.

Willms, J. D., L. Tramonte, J. Duarte and S. Bos (2012), *Assessing Educational Equality and Equity with Large-Scale Assessment Data: Brazil as a Case Study*, Inter-American Development Bank, Washington.

Wu, M. and R. Adams (2007), "Applying the Rasch model to psycho-social measurement: A practical approach", *Educational Measurement Solutions*, Melbourne, Australia.

APPENDIX

Table 16.A1. Item parameters for Sense of Belonging

Item	beta	d_1	d_2	d_3	alpha
ST068Q01TA	-0.00458	0.56688	0.37422	-0.94110	1.21518
ST068Q02TA	0.00475	1.02240	0.57396	-1.59636	0.77746
ST068Q03TA	0.15553	1.14692	0.59957	-1.74650	0.61414
ST068Q04TA	-0.00104	0.74923	0.34099	-1.09022	1.12698
ST068Q05TA	0.04790	1.35674	0.85709	-2.21383	0.66787
ST068Q06TA	-0.07787	0.53076	0.30405	-0.83481	1.59837

Table 16.A2. Item parameters for Classroom Disciplinary Climate

Item	beta	d_1	d_2	d_3	alpha
ST074Q01TA	0.19029	1.25309	0.51737	-1.77046	0.94803
ST074Q02TA	0.19407	1.22680	0.34986	-1.57666	1.29726
ST074Q03TA	-0.00888	1.07093	0.31662	-1.38755	1.14809
ST074Q04TA	-0.3381	1.08205	0.48490	-1.56696	0.79547
ST074Q05TA	-0.18866	0.99587	0.37880	-1.37468	0.81114

Table 16.A3. Item parameters for Teacher Satisfaction with Current Job

Item	beta	d_1	d_2	d_3	alpha
TC033Q04TA	0.08511	1.62541	0.54615	-2.17156	1.25762
TC033Q06TA	0.36952	1.70996	0.43825	-2.14821	1.13821
TC033Q07TA	-0.33945	1.85826	0.92920	-2.78746	0.74091
TC033Q08TA	-0.31986	1.81916	0.78320	-2.60236	0.86326

Table 16.A4. Item parameters for Teacher Satisfaction with the Profession

Item	beta	d_1	d_2	d_3	alpha
TC033Q01TA	0.009637	1.50124	0.58113	-2.08236	0.57877
TC033Q02TA	0.082150	1.14089	0.25229	-1.39318	1.33432
TC033Q03TA	-0.337690	0.93611	0.39061	-1.32672	1.25278
TC033Q05TA	0.369083	1.63123	-0.20855	-1.42269	0.83412

Table 16.A5. Item parameters for Teachers' Views on School Leadership

Item	beta	d_1	d_2	d_3	alpha
TC032Q03TA	-0.10263	1.86481	0.49417	-2.35898	1.04441
TC032Q04TA	0.13401	2.15078	0.39237	-2.54316	1.05004
TC032Q01TA	0.25823	2.21824	0.34937	-2.56761	0.75192
TC032Q02TA	-0.62260	1.59634	0.73520	-2.33154	0.92705
TC032Q05TA	0.28493	1.95661	0.30381	-2.26042	1.22657

Table 16.A6. Item parameters for Household Possessions

Item	beta	d_1	d_2	d_3	alpha
ST062Q01TA	-1.36226				1.163981
ST062Q02TA	-1.13818				0.981662
ST062Q03TA	-1.65830				0.858233
ST062Q04TA	-1.03435				3.177346
ST062Q05TA	0.48202				1.231151
ST062Q06TA	-1.46040				2.567074
ST062Q10TA	-1.41380				0.602710
ST062Q12TA	-2.43987				1.250521
ST063Q01TA	-0.70843	1.80899	-0.64006	-1.16892	1.088631
ST063Q02TA	0.86955	0.93075	-0.03666	-0.89409	1.232472
ST063Q03TA	0.30502	1.87318	-0.58952	-1.28365	1.127248
ST063Q05NA	-0.74917	0.39305	-0.73634	0.34330	1.126953
ST063Q06NA	0.78970	1.65749	-0.31055	-1.34694	2.509960
ST063Q09NA	0.82099	0.02957	-0.27544	0.24587	0.729314
ST063Q07NA	1.08752	0.76733	-0.45244	-0.31489	1.155935
ST063Q08NA	2.01433	-0.13910	-0.22065	0.35975	0.885854
ST064Q01NA	-2.66710				1.680299
ST064Q03NA	-0.61033				1.667431
ST064Q04NA	-2.59320				2.450752
ST064Q06NA	-2.73295				1.782885
ST066Q01NA	-0.13249	1.06106	-0.43941	-0.62166	0.475449

Table 16.A7. Item parameters for Depression

Item	beta	d_1	d_2	d_3	alpha
ST017Q06NA	1.62467	-0.67298	0.19481	0.47817	0.83759
ST017Q07NA	1.58161	0.11499	-0.14868	0.03369	1.31759
ST017Q08NA	0.92158	-0.37023	-0.23997	0.61020	0.65987
ST017Q09NA	2.02280	0.84802	-0.14036	-0.70765	1.69834
ST017Q10NA	0.74208	-0.53743	0.04770	0.48973	0.48094
ST017Q11NA	0.55137	0.04502	-0.28125	0.23623	0.68198

Table 16.A8. Item parameters for Family Resources

Item	beta	d_1	d_2	d_3	alpha
ST062Q01	-1.10771				0.62127
ST062Q02	-1.14486				0.65306
ST062Q03	-1.62141				0.52495
ST062Q04	-0.80094				1.89321
ST062Q05	0.34902				0.90769
ST062Q06	-1.02203				1.31740
ST062Q10	-1.57718				0.46136
ST062Q12	-3.17830				1.23058
ST064Q01	-3.04817				1.33153
ST064Q03	-0.73868				1.22466
ST064Q04	-3.14500				2.00151
ST064Q06	-2.93396				1.28216
ST063Q01	-0.88866	1.950902	-0.70163	-1.24927	0.91666
ST063Q02	0.73938	0.53967	-0.20545	-0.33422	0.72477
ST063Q03	0.01226	1.769387	-0.68908	-1.08031	0.90159
ST063Q04	-0.72873	0.311918	-0.77645	0.464534	0.71095
ST063Q05	0.76155	1.202482	-0.34374	-0.85874	1.35597
ST063Q06	0.79568	-0.02775	-0.26271	0.290454	0.52419
ST066Q01	-0.15848	1.047106	-0.44006	-0.60705	0.31572
ST049Q01	-1.12832				0.31081
ST048Q01	-2.56594				1.51018
ST057Q01	-1.48420				1.03872
ST051Q01	-1.36991	0.96702	-0.96702		0.70874
ST059Q01	-1.56820				0.52368

Table 16.A9. Item parameters for Attitudes towards School

Item	beta	d_1	d_2	d_3	alpha
ST067Q03TA	-1.45628	0.31463	1.17265	-1.48729	0.75383
ST067Q04TA	-2.65817	0.35543	1.28210	-1.63752	1.36664
ST067Q05TA	-3.49172	0.88161	1.19050	-2.07211	1.92350
ST067Q06TA	-3.52080	0.82398	1.22416	-2.04813	1.91353
ST067Q07TA	-3.24366	0.14840	1.71298	-1.86137	1.56427
ST067Q08TA	-3.74013	0.45394	1.73658	-2.19052	2.00834

Table 16.A10. Item parameters for Teacher-Student Relationships

Item	beta	d_1	d_2	d_3	alpha
ST072Q01NA	-2.11445	1.01331	1.14387	-2.15718	1.46310
ST072Q02NA	-2.00095	1.69847	0.89792	-2.59639	1.72577
ST072Q03NA	-2.12551	1.96972	0.76832	-2.73805	1.93457
ST072Q04NA	-1.84861	1.48313	0.77883	-2.26196	1.61381
ST072Q05NA	-1.64217	1.10888	0.87593	-1.98481	1.42256
ST072Q06NA	-2.82600	1.32896	1.08011	-2.40907	2.11237
ST072Q07NA	-2.78109	1.28222	1.10382	-2.38604	2.03271

Table 16.A11. Item parameters for Expectations for Success

Item	beta	d_1	d_2	d_3	alpha
ST072Q08NA	-3.42842	0.96793	1.51200	-2.47993	2.13167
ST072Q09NA	-3.86014	1.24239	1.40692	-2.64931	2.32977
ST072Q10NA	-3.84293	1.38277	1.48054	-2.86331	2.47474
ST072Q11NA	-2.19374	1.02606	1.28794	-2.31399	1.45088

Table 16.A12. Item parameters for Structured Lessons in Mathematics

Item	beta	d_1	d_2	d_3	alpha
ST075Q01NA	-1.44126	0.48312	-0.46698	-0.01614	0.95356
ST075Q02NA	-1.26551	0.85263	-0.31944	-0.53319	1.04171
ST075Q03NA	-2.24844	0.83008	-0.23380	-0.59628	1.47636
ST075Q04NA	-1.99614	0.84541	-0.14197	-0.70344	1.50857
ST075Q05NA	-2.14094	1.03756	-0.18112	-0.85644	1.62419
ST075Q06NA	-1.63777	0.78131	-0.04365	-0.73766	1.29763
ST075Q07NA	-2.39732	1.28703	-0.26608	-1.02095	1.71740
ST075Q08NA	-1.61122	0.82553	-0.26502	-0.56051	1.04975
ST075Q09NA	-1.17211	0.59528	-0.17373	-0.42155	0.99431
ST075Q11NA	-0.80405	0.59597	-0.30817	-0.28780	0.85614
ST075Q12NA	-1.37442	1.01744	-0.44597	-0.57147	0.84630

Table 16.A13. Item parameters for Instructional Resources

Item	beta	d_1	d_2	d_3	alpha
TC017Q01NA	-0.79019	-0.15737	1.005965	-0.8486	0.70337
TC017Q02NA	-0.68180	-0.19495	1.021168	-0.82621	0.65683
TC017Q03NA	-1.91261	1.15621	0.137509	-1.29372	0.81469
TC017Q04NA	-0.52833	-1.77578	0.452706	1.323069	0.21196
TC017Q05NA	-0.11225	-1.24966	0.890691	0.358972	0.30164
TC017Q06NA	0.19006	-1.58342	1.138494	0.444926	0.28834
TC017Q07NA	-0.17112	-2.0189	1.069184	0.949715	0.44957
TC017Q08NA	-0.18781	-2.49102	1.262943	1.228076	0.42538
TC017Q09NA	-0.28524	-1.34722	0.551853	0.795365	0.40804
TC017Q10NA	-0.97807	-0.56891	0.306417	0.262498	0.67493
TC017Q11NA	-0.46825	-1.17179	0.729814	0.441974	0.59438
TC017Q12NA	-0.43350	-1.69174	0.691081	1.000654	0.41947
TC017Q13NA	-0.30964	-0.93653	0.516923	0.41961	0.52928
TC017Q14NA	0.37319	-2.30045	1.366559	0.933888	0.32230
TC017Q15NA	0.90041	-2.51129	1.421376	1.089917	0.54690
TC017Q16NA	0.88057	-2.46732	1.408456	1.058867	0.48730
TC017Q17NA	-0.59428	0.195462	0.267435	-0.4629	0.70799
TC017Q18NA	0.10158	-1.84752	1.194809	0.652714	0.59762
TC017Q19NA	0.34765	-2.15041	1.177081	0.973333	0.58436
TC035Q18NA	0.54064	-1.4335	0.876647	0.556848	0.52979
TC035Q01NA	-0.02638	-0.60496	0.929605	-0.32464	0.76047
TC035Q02NA	0.52833	-1.50421	0.735121	0.76909	0.61548
TC035Q03NA	0.39493	-1.77024	1.177698	0.592545	0.44294
TC035Q04NA	0.35986	-1.42342	0.635073	0.788344	0.49608
TC035Q05NA	-0.85007	-0.46437	0.713161	-0.24879	0.77325
TC035Q06NA	-0.30348	-0.85001	0.886703	-0.03669	0.90289
TC035Q07NA	0.31902	-1.12492	0.839534	0.285383	0.58559
TC035Q08NA	-0.01605	-1.24485	0.770126	0.474722	0.40630
TC035Q09NA	-0.02427	-1.67428	0.857344	0.816931	0.45347

Item	beta	d_1	d_2	d_3	alpha
TC035Q10NA	0.27467	-1.72027	1.04922	0.671048	0.56408
TC035Q11NA	0.45016	-2.23696	1.075703	1.161256	0.33180
TC035Q12NA	0.25550	-1.6875	0.772367	0.915135	0.38981
TC035Q13NA	0.06500	-2.56333	0.764978	1.798352	0.35273
TC035Q14NA	-0.16096	-0.96477	0.634378	0.330389	0.61351
TC035Q15NA	-1.28560	-0.26661	0.599288	-0.33267	0.96894
TC035Q16NA	-0.38956	-0.02369	0.237875	-0.21418	0.92228
TC035Q17NA	0.39687	-2.03295	1.132611	0.90034	0.61781
TC017Q20NA	0.56931	-1.44636	0.786577	0.659784	0.54657
TC019Q01NA	-0.83607	0.18503	-0.16496	-0.02007	0.50829
TC019Q01NA	-0.72384	-0.06425	-0.03229	0.096539	0.44071

Table 16.A14. Item parameters for School Infrastructure

Item	beta	d_1	d_2	d_3	alpha
SC011Q01NA	-2.89193	1.82206	0.16570	-1.98777	1.41943
SC011Q02NA	-3.85073	1.40004	0.60236	-2.00239	2.00286
SC011Q03NA	-2.09184	0.62034	0.84084	-1.46118	1.43115
SC011Q04NA	-2.13614	0.44246	0.68251	-1.12497	1.31161
SC011Q05NA	-3.03857	1.92103	0.51779	-2.43883	1.88582
SC011Q06NA	-3.54011	2.47167	0.42555	-2.89722	2.18127
SC011Q07NA	-2.19711	-0.20227	1.20696	-1.00469	1.44717
SC011Q08NA	-2.68592	0.93517	1.06797	-2.00314	1.72939
SC012Q01NA	-0.80137	-0.10920	1.01643	-0.90723	1.14462
SC012Q03NA	0.31277	-1.90362	1.36004	0.54358	0.44458
SC012Q04NA	-0.62681	-1.05192	0.91483	0.13708	0.57344
SC012Q05NA	-1.14608	-0.95227	0.95892	-0.00664	0.90737
SC012Q06NA	-1.86906	0.19409	0.43246	-0.62654	1.24680
SC012Q07NA	-0.65108	-0.52115	1.24759	-0.72644	1.08870
SC012Q08NA	0.57437	-2.74434	1.97040	0.77394	0.53331
SC012Q09NA	0.78307	-3.33798	1.90777	1.43021	0.48143
SC012Q10NA	0.13774	-1.81444	1.30713	0.50731	0.59025
SC012Q11NA	-0.81922	0.23157	0.38657	-0.61814	0.68701
SC012Q12NA	-0.76745	-0.12368	0.42259	-0.29891	0.78636
SC012Q13NA	0.14278	-2.06947	1.44970	0.61977	0.37604
SC012Q14NA	-0.03832	-1.11802	0.99316	0.12486	0.63837
SC012Q15NA	-1.37361	0.14142	0.74383	-0.88525	1.27919
SC013Q01NA	-2.21127	1.39752	0.36547	-1.76299	1.56064
SC013Q02NA	-0.75085	-0.60245	1.02699	-0.42454	0.67838
SC014Q01NA	-0.77665	0.65585	-0.72499	0.06913	0.40972
SC015Q01NA	-0.63509	0.41804	-0.62327	0.20523	0.29024

Table 16.A.15 Cross-country comparability — RMSD values

Construct	ECU	GTM	HND	KHM	PRY	SEN	ZMB	WRMSD
BELONG	0.096	0.132	0.140	0.114	0.135	0.097	0.101	0.120
DISCLI	0.088	0.081	0.074	0.064	0.082	0.100	0.068	0.084
DEPRESSION	0.022	0.021	0.030	0.051	0.023	0.034	0.041	0.033
FAMRES	0.061	0.077	0.063	0.105	0.088	0.096	0.106	0.094
ATSCH	0.024	0.022	0.024	0.060	0.023	0.042	0.045	0.038
STTCHREL	0.045	0.037	0.034	0.078	0.052	0.075	0.053	0.057
TCEXPSUC	0.037	0.033	0.037	0.076	0.050	0.061	0.040	0.050
STRLSMAT	0.054	0.041	0.038	0.062	0.044	0.045	0.047	0.049
SATJOB	0.129	0.099	0.104	0.134	0.108	0.087	0.118	0.120
SATTEACH	0.085	0.066	0.063	0.102	0.106	0.087	0.090	0.098
TCLEAD	0.097	0.090	0.073	0.118	0.113	0.069	0.096	0.103
INSTRES	0.110	0.104	0.082	0.085	0.095	0.087	0.117	0.105
SCHMATRES	0.131	0.112	0.116	0.153	0.036	0.118	0.128	0.133