

PISA 2022 Technical Report



14 **Scaling outcomes**

This chapter reports on the outcomes of implementing the item response theory (IRT) and population modelling methods described in Chapter 11 for the PISA 2022 main survey cognitive assessment data. It provides results of the assessments of the invariance of the IRT item parameters within and across countries/economies, estimates of the reliability and correlations across assessments domains, and estimates of the linking errors between the 2022 and prior PISA cycles. The location of the items across the full range of proficiencies based on their common international parameters are also reported. Finally, the correlations between scales and the percentage of students in each country at each proficiency level are presented for each cognitive domain.

IRT scaling outcomes

IRT scaling outcomes include the proportions of item were invariant across countries and PISA cycles, as well as the common and unique items parameters and the dropped items used for the population modelling of each country/economy data. The international (common) item parameters are provided in this technical report's Annex A and unique country/economy's item parameters are provided in Annex F. The next section provides an assessment of item parameter invariance across countries/economies supporting that the comparability of the PISA scales across cycles and countries was achieved in each domain by reaching a desirable proportion of invariant item parameters across countries/economies and cycles. The following section describes the international characteristics of each domain's item pool and shows the item maps that locate the items on the reporting scales.

Invariance of item parameters

The item parameters for all the items used in the assessment were obtained through IRT scaling. In PISA 2022, IRT scaling was implemented through a multi-group (i.e., country-by-language groups) IRT concurrent calibration using the 2022 main survey data, using the trend items as fixed linking items and setting the scale to the PISA scale established in 2015 and 2018. That is, item parameters for trend items were fixed to the ones used in PISA 2018 (either common international or unique to a specific country-by-language group or groups), unless there was evidence that the 2018 parameters did not fit the 2022 data (see Chapter 11 for details).

In most cases the international item parameters fitted data for all country-by-language groups. When they did not fit a particular country-by-language group, unique or group-specific parameters were estimated and used, unless it was found that the unique parameters could not be estimated, still did not fit the data well enough, or were extreme. That is, an item was dropped if in the end, its RMSD fit could not be reduced to 0.15 or below, its slope parameter was below 0.1 or its difficulty parameter was larger than 5.0 in absolute value. These criteria were not applied to reading fluency items because they typically are very easy items. In rare cases, items were also dropped when, despite being checked in the field trial, content and/or translation issues were nonetheless found in the main survey—given feedback from countries/economies, content and psychometric reviews. In even rarer cases, items were dropped entirely (in all countries/economies) if analyses indicated that it did not fit the data collected in the majority of the

countries/economies. In this PISA 2022 cycle: one mathematics, one reading, one financial literacy and six creative thinking items were dropped from all groups ¹.

To assess the invariance of item parameters across country-by-language groups and cycles, items were categorized as:

- *invariant* when common international parameters could be used;
- *group-specific invariant* when the same unique parameters could be used across cycles (applies only to trend items);
- *variant* for all other cases where unique item parameters were estimated (new items) or when unique parameters were estimated that are different from the 2018 parameters (trend items); and
- *dropped* when the item could not be fitted to the data and was dropped for one or more country-by-language groups.

For countries with multiple language groups, the number of invariant, variant, or dropped items were averaged across the different language groups within the country to calculate the proportion of unique item parameters used. Sample weights were used for this calculation.

Table 14.1 shows the proportions of items categorized as invariant, variant, and dropped, averaged across countries participating in the 2022 computer-based assessment (CBA). The proportion of invariant items was large for all domains, ranging from 76.4% for the reading MSAT items to 93.7% for the reading fluency items. A large proportion of invariant items is critical for ensuring the comparability of scores across countries and cycles. Group-specific invariant items also contribute to the comparability of scores across cycles. The proportion of invariant total (invariant and group-specific invariant) was above 98.5% for all domains but creative thinking at 77.4%. Regarding the dropped category, the proportions were small for all domains (less than 2%).

Table 14.2 shows the proportion of items categorized as invariant, variant, and dropped, averaged across countries participating in the new 2022 paper-based assessment (new PBA). The results across the three new PBA participating countries showed somewhat lower proportions of invariance than with CBA. Nevertheless, proportions of total invariant items were above 80% for all domains and few items were dropped for any country.

An overview of the frequencies of invariant, variant, and dropped items for each domain is presented in Figure 14.1, Figure 14.2, Figure 14.3, Figure 14.4 and Figure 14.5 for CBA, new PBA and PBA participating countries/economies. Each country is represented by stacked vertical bars: above the horizontal line at zero, dark green represents the number of items classified as invariant and light green represents the number of group-specific invariant items (only trend items); below the 0 horizontal line, yellow represents the number of variant items² and red represents the number of items dropped from scaling. The frequencies of variant and dropped items are shown using negative values to highlight differences between the number of items that contribute to ensuring the comparability of the PISA scales (invariant) and the number of items that do not (variant). The countries are sorted from left to right by increasing number of invariant items, first CBA, new PBA, and PBA countries.

These plots show that while there is some variability across countries, the numbers of invariant item parameters and group-specific invariant item parameters are large enough to ensure the comparability of the proficiency estimates across countries/economies and across cycles.

Figure 14.1. Frequency of invariant, variant, and dropped items for mathematics, by country/economy

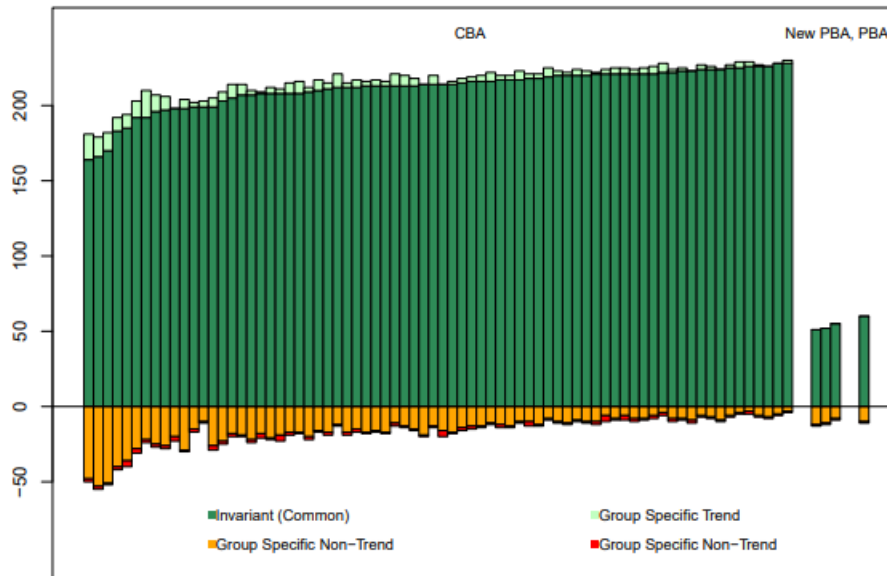
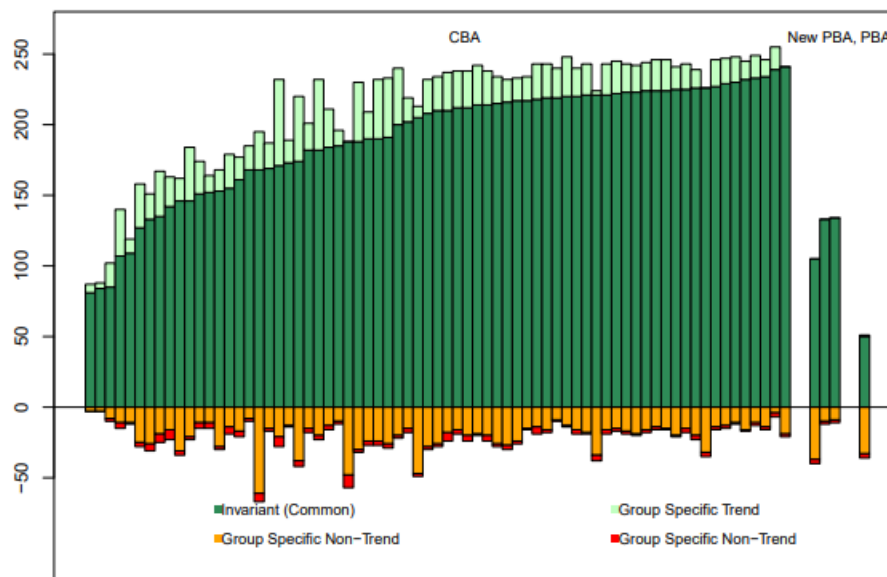


Figure 14.2. Frequency of invariant, variant, and dropped items for reading, by country/economy



Note: Because reading is a minor domain in 2022, in some countries, sample size was not enough to assess fit with the 2022 data. These cases are not included in this plot, resulting in fewer than the number of items used being displayed in these cases. However all items were evaluated for fit in 2018 when reading was the major domain--see PISA 2018 Technical report, Chapter 12 for these results).

Figure 14.3. Frequency of invariant, variant, and dropped items for science, by country/economy

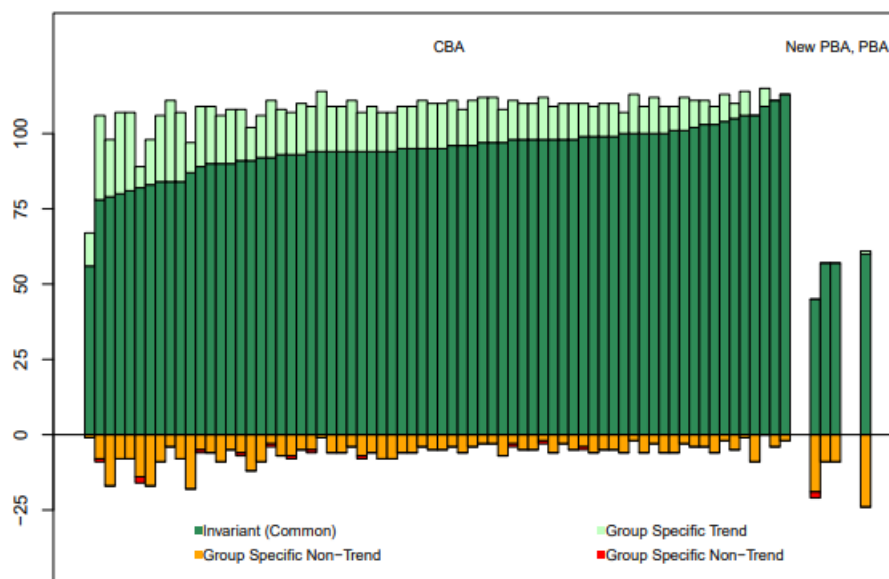


Figure 14.4. Frequency of invariant, variant, and dropped items for financial literacy, by country/economy

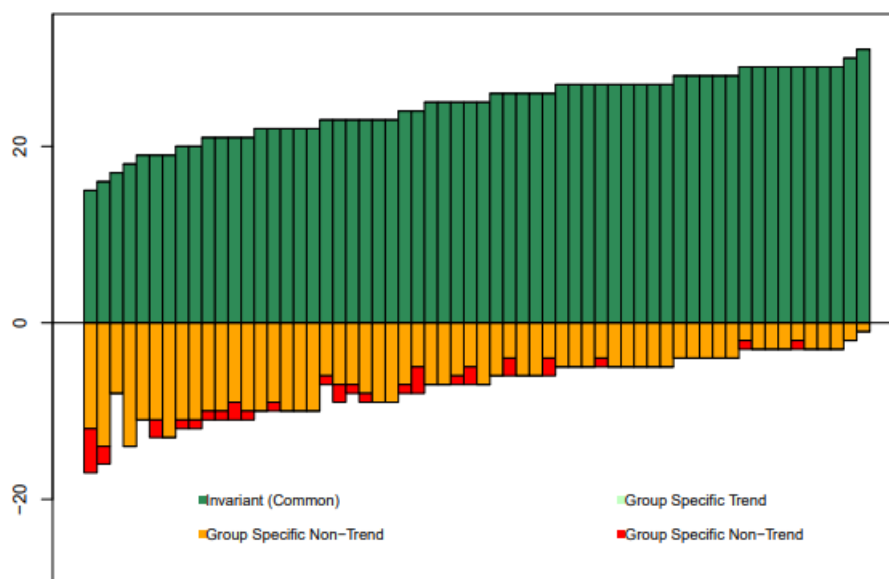
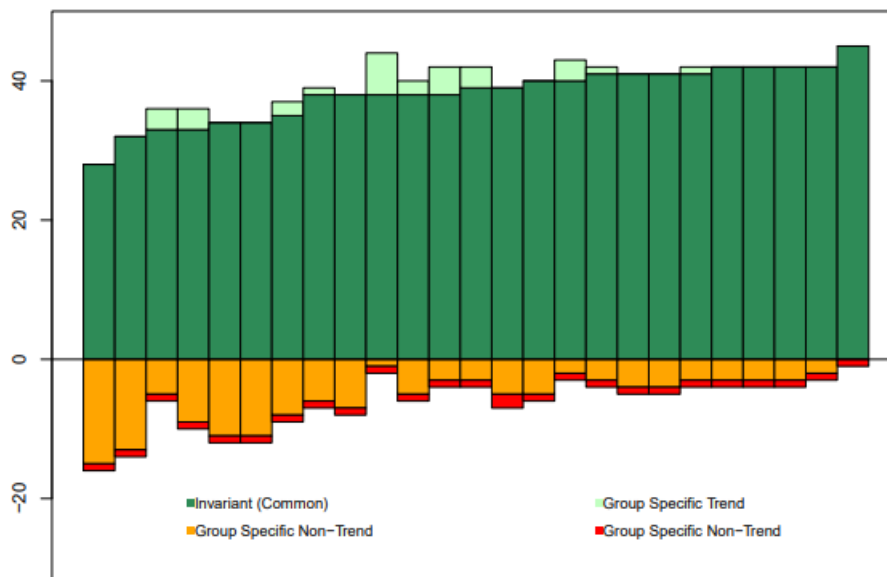


Figure 14.5. Frequency of invariant, variant, and dropped items for global competence, by country/economy



International characteristics of the item pools

This section provides an overview of the test targeting, the domain inter-correlations, and the correlations among the mathematics subscales.

Test targeting

Similar to assigning a specific score on a scale to students according to their performance on an assessment (OECD, 2022^[1]), each item in PISA 2022 was assigned a specific value on a scale based on response probability (RP) calculated using the item's IRT parameters (discrimination and difficulty). Chapter 17 describes how items can be placed along a scale based on their RP values and how these values can be used to classify items into proficiency levels.

Historically in PISA, a response probability of 0.62 (RP62) has been used to classify items into levels. Students with a proficiency located at or below this point have a probability of 0.62 or less of getting the item correct, while students with a proficiency above this point have a higher probability of getting the item correct higher than 0.62. The RP62 values for all items and their performance level classification are presented in Annex A, together with the final international/common item parameter estimates obtained from the IRT scaling. Note that for polytomous items, the RP62 value is provided for partial credit as well as full credit responses. The partial credit RP62 has been defined as the minimum proficiency level a student need to have an expected score that is 62% of the full credit.

Table 14.3, Table 14.4, Table 14.5, Table 14.6 and Results for Creative Thinking will be available in 2024

Table 14.7 show the proficiency levels defined for for each cognitive domain, along with the percentage of items and the percentages of students classified at each level of proficiency, using the first plausible value. Note that although polytomous items have two RP62 levels (partial credit and full credit), they were classified according to the full credit RP62 only for all domains but creative thinking. For creative thinking, most of the items are polytomous items (28 out of 32), therefore we describes both partial- and full-credit RP62 levels.

Since RP62 values and the plausible values are on the same PISA scale, the distribution of students' latent ability and the items' RP62 values can be compared and contrasted. In Figure 14.6, Figure 14.7, Figure 14.8, Figure 14.9 and Figure 14.10, the left side of each figure illustrates the distribution of the first plausible values (PV1) across countries. In each figure, the blue line indicates the empirical density of the first plausible values across all countries, and the red line indicates the theoretical normal distribution with the mean and the variance of plausible values across all countries. The figures show that the distribution of the plausible values for each domain are approximately normal. On the right side of each figure, the RP62 value for each of the items is plotted. As with the tables above, in all domains but creative thinking, only the RP62 values for full-credit are shown.

Figure 14.6. Distribution of the first plausible values and item RP62 values in mathematics

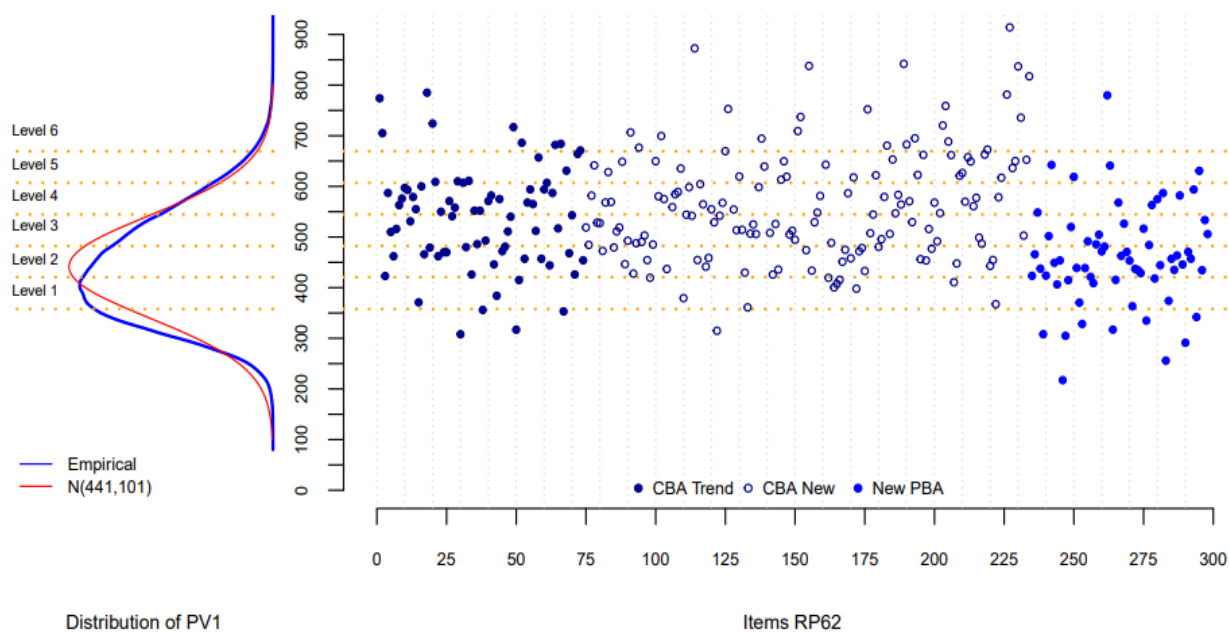


Figure 14.7. Distribution of the first plausible values and item RP62 values in reading

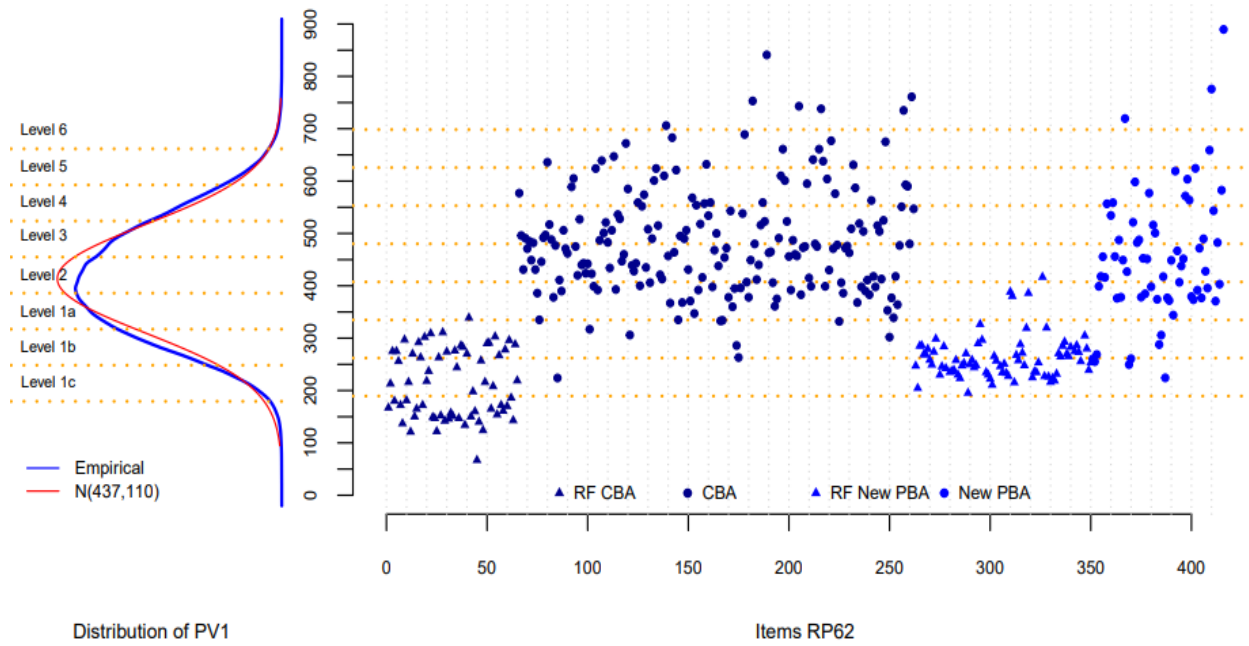


Figure 14.8. Distribution of the first plausible values and item RP62 values in science

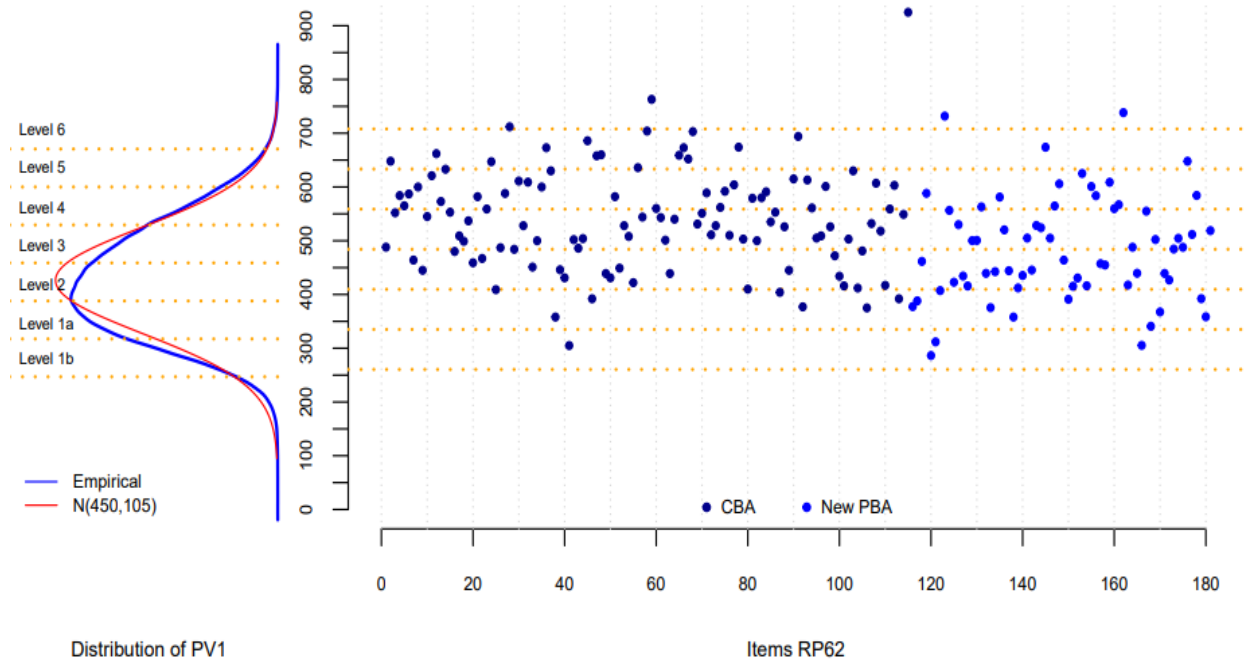


Figure 14.9. Distribution of the first plausible values and item RP62 values in creative thinking

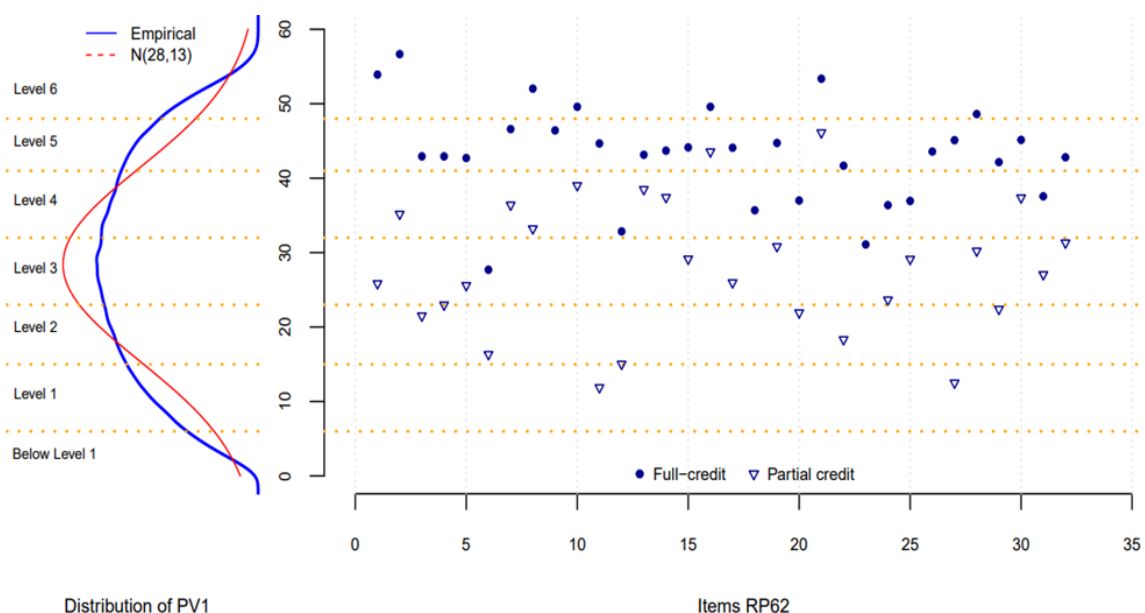
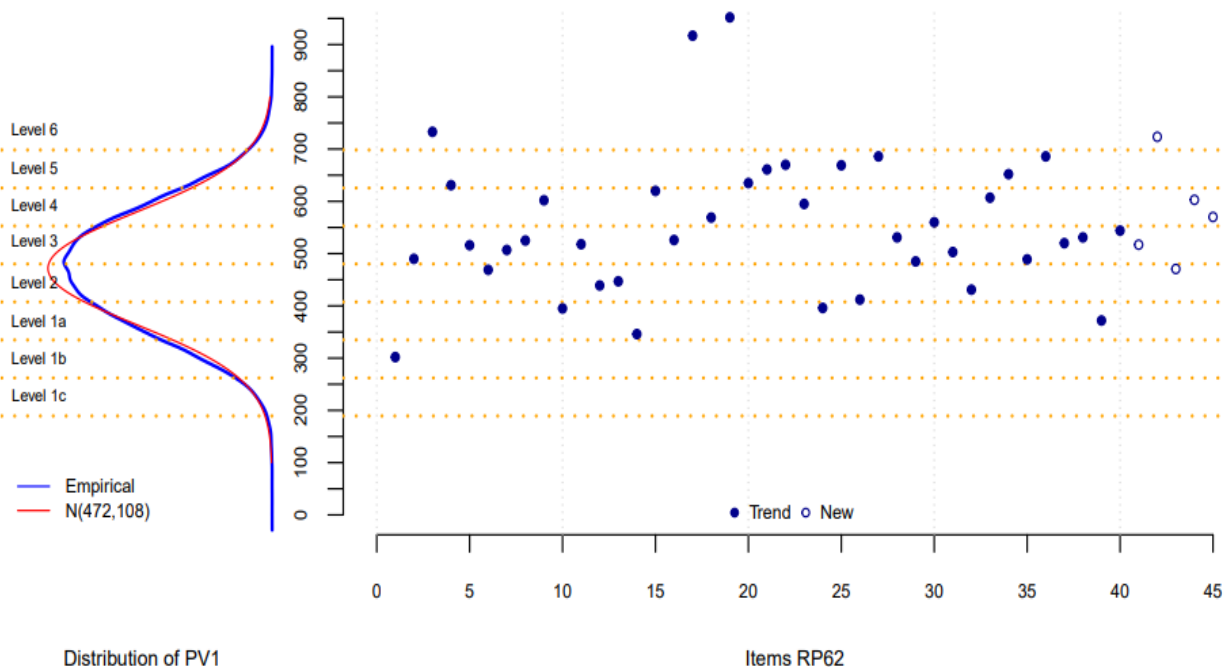


Figure 14.10. Distribution of the first plausible values and item RP62 values in financial literacy



Population modelling outcomes

The population modelling outcomes include the multivariate latent regressions models estimated for each country/economy and the plausible values (PVs) generated from them, which are included in the international and national databases. Because the latent part of the population model comes from the IRT scaling, the plausible value are generated on their underlying PISA IRT metric used when estimating IRT

item and group parameters and then transformed to the PISA scale. For example, mathematics IRT scaled PVs are produced and then transformed to the PISA metric of mean 500 and standard deviation 100 across all participating OECD countries during the first mathematics assessment. Based on these PVs, then the overall and country/economy-level PISA scale reliability, average performance and students percentile by proficiency levels, and finally the correlations between domain scales were estimated. In the next sections, the methods used to transform the PVs from the IRT scale to the PISA reporting scales are described and the outcomes are reported.

Mathematics, reading, science and financial literacy scaling transformations

The mathematics, reading, and science PISA reporting scales were set when the domain became a major domain for the first time—in 2006 for science, 2009 for reading and 2012 in mathematics. This was done using a linear transformations of the senate weighted OECD participating countries/economies IRT scaled plausible values available at the time, so that the overall mean was 500 and the standard deviation 100, resulting in nearly all reported plausible values being between 200 and 800. The same approach was used for each new innovative domain and for the optional financial literacy domain.

However, because the IRT models used for scaling were updated in 2015, a bridge study was completed as part of the 2015 scaling analyses to establish new IRT to reported PISA scale parameters. This did not change the scales or the scores reported prior to 2015, but the new transformations have been applied since. Detailed descriptions of the bridge study and its results are provided in the PISA 2015 technical report OECD (2017^[2]), Chapter 12)

Results for Financial Literacy will be available in 2024

Table 14.8 provides the PISA IRT theta to reported PISA proficiency scale linear transformation A and B coefficients for the core and financial literacy domains. Given any IRT scaled theta (θ) value (e.g., item difficulty, item step parameters, or student PV or proficiency), the transformed value on the PISA scale is $A * \theta + B$.

Creative Thinking scaling transformation

For the creative thinking innovative domain developed for the PISA 2022 main survey, it was found that the use of a non-linear transformation provided a more appropriate reporting scale. This was because the particular challenges in creating such an innovative measure resulted in a relatively small pool that did not provide much information towards the lower end of the scale. To best support scale interpretations, the creative thinking item pool IRT test characteristic curve transformation of the theta plausible values PV_θ was applied to obtain the reported plausible values $PV_{NC} = \sum_{i \in V_p} T_i(PV_\theta)$, where V_p indicates the set of 32 creative thinking items and $T_i(PV_\theta)$ is the expected score on item i as a function of PV_θ and item i 's IRT parameters. In this way the reported creative thinking plausible values can be interpreted as the expected number correct on a hypothetical form made up of all the items in the creative thinking item pool, given the proficiency level that the plausible value represents.

Reliability of the PISA scales

As was done in prior PISA cycle, test reliability was estimated using the well known theoretical formula: $1 - (\text{expected error variance}/\text{total variance})$. In practice, the expected error variance is the weighted average of the students' posterior variance, computed as the variance of the 10 plausible values, which is an expression of the posterior measurement error. The total variance was computed using a resampling approach (Efron, 1982^[3]), using each country/economy set of resampling weights.

Table 14.9 presents the test reliability descriptive statistics across countries/economies for the cognitive domains and the mathematics subscales. The reliabilities for each country/economy are presented in

Table 14.10. Overall, we observe that in average test reliability is high for the core and financial literacy domains (0.84 to 0.90) and a bit less for creative thinking (0.8), and that most countries/economies' reliability is close to the average. As expected, since the number of items is smaller than for the full mathematics instrument, the reliability of the mathematics subscales are much lower and more variable cross countries/economies.

Table 14.11 shows the average transformed plausible values as well as the resampling-based standard errors for each country and domain.

Domain inter-correlations

Estimated correlations between the domains, based on the 10 reported plausible values and averaged across all countries and assessment modes, are presented in Table 14.12a and Table 14.12b for the core domains and creative thinking, and in Table 14.13 for the financial literacy sample. The estimated correlations for each country are presented in Table 14.14.

Mathematics subscales correlations

There were two sets of subscales reported for mathematics. The first set, measuring content domains, was composed of the following four subscales: space and shape (MCSS), quantity (MCQN), change and relationships (MCCR), and uncertainty and data (MCUD). The second set, based on the cognitive processes, comprised the following four subscales: employing mathematical concepts, facts, and procedures (MPEM), interpreting, applying, and evaluating mathematical outcomes (MPIN), formulating situations mathematically (MPFS), and reasoning (MPRE).

The correlations between reading, science and the mathematics content domain subscales are presented in Table 14.15a. Table 14.15b shows the correlations between reading, science and the cognitive process domains.

Note that, as indicated in Chapter 11, because of the way in which these subscale plausible values were estimated, it is not appropriate to correlate the cognitive process subscales with the cognitive contents subscales, or any of the subscales with the overall mathematics proficiency.

Countries/economies average proficiency and percentages of students at each proficiency level

Figure 14.11, Refer to the PISA 2022 Technical Report website to view this figure on line ([link](#)).

Figure 14.12, Figure 14.13, Refer to the PISA 2022 Technical Report website to view this figure on line ([link](#)).

Figure 14.14 and Creative Thinking results will be available online by 2024.

Figure 14.15 show the average proficiency and percentages of students at each proficiency level across countries/economies for each domain.

Figure 14.11. Percentage of students in each country/economy at each proficiency level for mathematics

Refer to the PISA 2022 Technical Report website to view this figure on line ([link](#)).

Figure 14.12. Percentage of students in each country/economy at each proficiency level for reading

Refer to the PISA 2022 Technical Report website to view this figure on line ([link](#)).

Figure 14.13. Percentage of students in each country/economy at each proficiency level for science

Refer to the PISA 2022 Technical Report website to view this figure on line (link).

Figure 14.14. Percentage of students in each country/economy at each proficiency level for creative thinking

Creative Thinking results will be available online by 2024.

Figure 14.15. Percentage of students in each country/economy at each proficiency level for financial literacy

Financial Literacy results will be available online by 2024.

Linking error

The estimation of the linking error between two PISA cycles was accomplished by considering the differences between the reported country means from the previous PISA cycles and new estimates of these country means based on the new PISA cycle item parameters. To estimate the linking error for trend comparisons between PISA 2022 and a previous PISA cycle down to 2006, the subset of countries that had participated in both cycles being compared was used. In the cases of trends to 2000 or 2006 or financial literacy, since the number of participating countries was relatively small, all countries were used.

The 2022 linking errors are reported in Table 14.16. Using these values help evaluate the extent to which changes in a country/economy or subgroup's performance between PISA 2022 and a previous PISA cycle are significantly different.

Note that for each domain, the earliest cycle for which comparisons can be made between PISA 2022 and a previous PISA cycle is the cycle in which the domain first became a major domain. Thus, the comparison of mathematics scores between PISA 2022 and PISA 2000 is not possible, nor is the comparison of science scores between PISA 2022 and PISA 2000 or between PISA 2022 and PISA 2003. Detail on the methodology used to calculate the linking errors can be found in Chapter 11.

References

- Efron, B. (1982), "The jackknife, the bootstrap, and other resampling plans", *CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38*, <https://doi.org/10.1137/1.9781611970319>. [3]
- OECD (2022), *PISA 2018 Technical Report*. [1]
- OECD (2017), *PISA 2015 Technical Report*, OECD Publishing, Paris, <http://www.oecd.org/pisa/data/2015-technical-report/>. [2]

Notes

1. The dropped items are: CMA112Q02, CR547Q07S, DF082Q01C, and DT520Q01C, DT560Q01C, DT560Q02C, DT450Q01C, DT450Q02C and DT450Q03C.
2. For the trend items classified as variant in a specific group (yellow), the 2018 parameters did not appropriately fit the 2022 data; thus, new unique parameters were estimated. For new items classified as variant in a specific group (yellow), unique parameters were needed due to the misfit of the common international parameters to the 2022 data.

Chapter 14 tables

Tables	Title
Table 14.1	Proportion of invariant, variant, and dropped CBA items averaged across countries/economies, for each domain
Table 14.2	Proportion of invariant, variant, and dropped new PBA items averaged across countries/economies, for each domain
Table 14.3	Proficiency levels for mathematics and the classification of items and students
Table 14.4	Proficiency levels for reading and the classification of items and students
Table 14.5	Proficiency levels for science and the classification of items and students
Table 14.6	Proficiency levels for creative thinking and the classification of items and students
Table 14.7	Proficiency levels for financial literacy and the classification of items and students
Table 14.8	PISA IRT theta to reported PISA proficiency scale linear transformation coefficients
Table 14.9	Test reliability descriptive statistics across countries/economies for the cognitive domains and the mathematics subscales
Table 14.10	Countries/economies reliability values for the cognitive domains
Table 14.11	Average plausible values (PV) and resampling-based standard errors (SE) by country and domain. (CrT and FLit results will be available in 2024.)
Table 14.12a	Core domain inter-correlations for the main sample
Table 14.12b	Creative Thinking inter-correlations with core domains for the main sample
Table 14.13	Domain inter-correlations for the financial literacy sample
Table 14.14	Domain inter-correlations by country/economy
Table 14.15a	Mathematics content subscales inter-correlations
Table 14.15b	Mathematics cognitive process subscales inter-correlations
Table 14.16	Linking error for score comparisons between PISA 2022 and previous PISA cycles

Table 14.1. Proportion of invariant, variant, and dropped CBA items averaged across countries/economies, for each domain

	Mathematics		Reading		Science	Financial	Creative
	Trend	New	Fluency	MSAT	All	Literacy	Thinking
Total items	74	159	65	196	115	40	32
Total countries	68	68	68	68	68	19	55
Invariant	86.0%	92.6%	93.7%	76.4%	83.0%	84.1%	77.4%
Group-specific invariant	6.5%	-	3.5%	11.8%	11.8%	2.6%	-
Invariant total ¹	92.5%	92.6%	97.2%	88.2%	94.8%	86.7%	77.4%
Noninvariant	6.1%	7.0%	1.7%	10.6%	4.0%	11.0%	20.8%
Dropped	1.4%	0.4%	1.1%	1.2%	1.2%	2.3%	1.8%

Table 14.2. Proportion of invariant, variant, and dropped new PBA items averaged across countries/economies, for each domain

	Mathematics	Reading		Science
		Fluency	Reading	
Total items	64	79	66	66
Total countries	3	3	3	3
Invariant	82.3%	90.7%	79.3%	80.3%
Noninvariant	16.1%	9.3%	17.2%	18.7%
Dropped	1.6%	0.0%	3.5%	1.0%

Table 14.3. Proficiency levels for mathematics and the classification of items and students

Classification	Number of items		Percentage of items		Percentage of respondents	
	CBA	New PBA	CBA	New PBA	CBA	New PBA
Level 6	38	1	16%	2%	2%	
Level 5	34	4	15%	6%	5%	0%
Level 4	50	7	21%	11%	10%	0%
Level 3	48	10	21%	16%	16%	3%
Level 2	45	25	19%	39%	21%	10%
Level 1a	13	8	6%	13%	23%	26%
Level 1b	5	6	2%	9%	24%	33%
Level 1c		2		3%		20%
Below Level 1		1		2%		7%

Table 14.4. Proficiency levels for reading and the classification of items and students

Classification	Number of items				Percentage of items				Percentage of respondents	
	RF CBA	CBA	RF New PBA	New PBA	RF CBA	CBA	RF New PBA	New PBA	CBA	New PBA
Level 6		7		4		4%		6%	1%	
Level 5		14		1		7%		2%	4%	0%
Level 4		25		10		13%		15%	11%	1%
Level 3		42		10		21%		15%	19%	5%
Level 2		59	1	15		30%	1%	23%	24%	18%
Level 1a	1	40	3	18	2%	20%	4%	27%	23%	37%
Level 1b	22	8	30	3	34%	4%	38%	5%	15%	30%
Level 1c	11	1	46	5	17%	1%	58%	8%	5%	8%
Below Level 1	31				48%				1%	1%

Table 14.5. Proficiency levels for science and the classification of items and students

Classification	Number of items		Percentage of items		Percentage of respondents	
	CBA	New PBA	CBA	New PBA	CBA	New PBA
Level 6	3	3	3%	4%	1%	
Level 5	15	8	13%	9%	4%	0%
Level 4	31	23	27%	27%	11%	0%
Level 3	36	30	31%	35%	20%	4%
Level 2	22	17	19%	20%	25%	18%
Level 1a	7	3	6%	4%	24%	43%
Level 1b	1	1	1%	1%	13%	30%
Below 1b					2%	5%

Table 14.6. Proficiency levels for creative thinking and the classification of items and students

Results for Creative Thinking will be available in 2024

Table 14.7. Proficiency levels for financial literacy and the classification of items and students

Results for Financial Literacy will be available in 2024

Table 14.8. PISA IRT theta to reported PISA proficiency scale linear transformation coefficients

Domain	A	B
Mathematics	135.9030	514.1848
Reading	131.5532	437.9244
Science	168.3189	494.5360
Financial literacy	140.0807	490.7259
Creative thinking*	-	-

Note: * Not applicable because a non-linear test characteristic curve transformation was used.

Table 14.9. Test reliability descriptive statistics across countries/economies for the cognitive domains and the mathematics subscales

MODE	Domains	Median	S.D.	Max	Min	
CBA	Mathematics	0.90	0.03	0.93	0.81	
	Content	Change and Relationships	0.85	0.05	0.91	0.66
		Quantity	0.87	0.04	0.91	0.75
		Space and Shape	0.80	0.08	0.87	0.57
		Uncertainty and Data	0.84	0.05	0.90	0.71
	Cog. Process	Employing Mathematical Concepts, Facts, and Procedures	0.87	0.04	0.91	0.75
		Formulating Situations Mathematically	0.83	0.08	0.90	0.57
		Interpreting, Applying, and Evaluating Mathematical Outcomes	0.86	0.04	0.90	0.74
		Reasoning	0.85	0.08	0.91	0.59
	Reading	0.86	0.03	0.91	0.77	
	Science	0.87	0.03	0.92	0.79	
	Financial literacy	0.90	0.02	0.92	0.85	
	Creating Thinking	0.80	0.04	0.89	0.65	
PBA	Reading	0.87	0.03	0.90	0.84	
	Mathematics	0.87	0.01	0.89	0.85	
	Science	0.84	0.03	0.87	0.81	

Table 14.10. Countries/economies reliability values for the cognitive domains

Mode	Country/Economy	Mathematics	Reading	Science	Financial Literacy	Creative Thinking
CBA	Albania	0.85	0.77	0.80		0.80
CBA	United Arab Emirates	0.90	0.86	0.85	0.89	0.77
CBA	Argentina	0.87	0.85	0.85		
CBA	Australia	0.92	0.85	0.87		0.76
CBA	Austria	0.93	0.90	0.91	0.91	
CBA	Belgium*	0.92	0.86	0.89	0.90	0.79
CBA	Bulgaria	0.90	0.87	0.86	0.89	0.82
CBA	Brazil	0.86	0.84	0.85	0.87	0.77
CBA	Brunei Darussalam	0.92	0.91	0.91		0.87
CBA	Canada*	0.89	0.82	0.83	0.90	0.69

Mode	Country/Economy	Mathematics	Reading	Science	Financial Literacy	Creative Thinking
CBA	Switzerland	0.92	0.90	0.91		
CBA	Chile	0.87	0.84	0.86		0.77
CBA	Colombia	0.87	0.85	0.86		0.81
CBA	Costa Rica	0.86	0.84	0.83	0.87	0.82
CBA	Czech Republic	0.92	0.87	0.89	0.90	0.79
CBA	Germany	0.92	0.88	0.90		0.82
CBA	Denmark	0.90	0.85	0.89	0.90	0.78
CBA	Dominican Republic	0.82	0.85	0.81		0.79
CBA	Spain	0.89	0.82	0.83	0.86	0.65
CBA	Estonia	0.90	0.84	0.86		0.77
CBA	Finland	0.91	0.86	0.87		0.81
CBA	France	0.92	0.87	0.88		0.80
CBA	United Kingdom	0.92	0.87	0.89		
CBA	Georgia	0.88	0.84	0.83		
CBA	Greece	0.89	0.84	0.86		0.81
CBA	Hong Kong (China)	0.92	0.85	0.86		0.77
CBA	Croatia	0.91	0.84	0.87		0.77
CBA	Hungary	0.92	0.88	0.90	0.90	0.84
CBA	Indonesia	0.85	0.87	0.85		0.81
CBA	Ireland	0.91	0.88	0.89		
CBA	Iceland	0.89	0.84	0.86		0.77
CBA	Israel	0.92	0.86	0.88		0.85
CBA	Italy	0.91	0.86	0.88	0.90	0.80
CBA	Jamaica	0.88	0.89	0.88		0.89
CBA	Jordan	0.81	0.81	0.82		0.79
CBA	Japan	0.92	0.86	0.89		
CBA	Kazakhstan	0.85	0.83	0.81		0.74
CBA	Korea	0.92	0.85	0.88		0.80
CBA	Kosovo	0.85	0.85	0.84		
CBA	Lithuania	0.91	0.85	0.89		0.81
CBA	Latvia	0.90	0.85	0.88		0.74
CBA	Macao (China)	0.91	0.84	0.88		0.80
CBA	Morocco	0.84	0.82	0.81		0.83
CBA	Republic of Moldova	0.89	0.88	0.86		0.81
CBA	Mexico	0.86	0.86	0.86		0.79
CBA	North Macedonia	0.88	0.84	0.84		0.85
CBA	Malta	0.91	0.87	0.88		0.85
CBA	Montenegro	0.89	0.86	0.86		
CBA	Mongolia	0.89	0.84	0.86		0.79
CBA	Malaysia	0.90	0.88	0.88	0.92	0.85
CBA	Netherlands	0.93	0.89	0.91	0.91	0.84
CBA	Norway	0.91	0.86	0.87	0.85	
CBA	New Zealand	0.92	0.88	0.89		0.83
CBA	Panama	0.87	0.88	0.88		0.85
CBA	Peru	0.87	0.84	0.86	0.89	0.80
CBA	Philippines	0.88	0.90	0.87		0.89
CBA	Poland	0.91	0.87	0.87	0.90	0.78
CBA	Portugal	0.91	0.85	0.88	0.87	0.78
CBA	Palestinian Authority	0.83	0.82	0.81		0.81
CBA	Qatar	0.91	0.87	0.88		0.84
CBA	Baku (Azerbaijan)	0.87	0.81	0.81		0.73
CBA	Cyprus	0.90	0.84	0.84		0.81
CBA	Ukrainian regions (18 of 27)	0.89	0.86	0.87		0.82
CBA	Romania	0.92	0.90	0.90		0.85
CBA	Saudi Arabia	0.83	0.81	0.79	0.86	0.76

Mode	Country/Economy	Mathematics	Reading	Science	Financial Literacy	Creative Thinking
CBA	Singapore	0.92	0.86	0.88		0.78
CBA	El Salvador	0.82	0.84	0.84		0.79
CBA	Serbia	0.90	0.86	0.87		0.79
CBA	Slovak Republic	0.92	0.87	0.89		0.85
CBA	Slovenia	0.91	0.87	0.90		0.82
CBA	Sweden	0.92	0.88	0.90		
CBA	Chinese Taipei	0.93	0.89	0.90		0.81
CBA	Thailand	0.88	0.86	0.87		0.85
CBA	Türkiye	0.92	0.88	0.90		
CBA	Uruguay	0.89	0.85	0.87		0.81
CBA	United States	0.92	0.90	0.92	0.91	
CBA	Uzbekistan	0.81	0.80	0.79		0.76
New PBA	Guatemala	0.88	0.90	0.87		
New PBA	Cambodia	0.85	0.84	0.81		
New PBA	Paraguay	0.89	0.89	0.87		
PBA	Viet Nam	0.87	0.84	0.81		

Note: Ukrainian regions (18 out of 27) administered the assessment.

*Denotes a country/economy for which the financial literacy domain was not fully sampled across the population; it is not a nationally-representative sample.

Table 14.11. Average plausible values (PV) and resampling-based standard errors (SE) by country and domain (CrT and Financial Literacy results will be made available in 2024).

Country	Reading		Mathematics		Science	
	Average PV	SE	Average PV	SE	Average PV	SE
International average	435.04	0.30	437.63	0.27	446.89	0.28
Albania	358.43	1.93	368.22	2.09	375.97	2.22
Argentina	400.74	2.57	377.53	2.25	406.19	2.49
Australia	498.05	2.01	487.08	1.78	507.00	1.93
Austria	480.41	2.67	487.27	2.34	491.27	2.65
Baku (Azerbaijan)	365.21	2.45	396.88	2.38	380.14	2.21
Belgium*	478.85	2.52	489.49	2.20	490.58	2.48
Brazil	410.36	2.09	378.69	1.58	403.00	1.93
Brunei Darussalam	429.23	1.16	442.09	0.93	445.86	1.32
Bulgaria	404.30	3.40	417.30	3.30	420.99	3.17
Cambodia	328.84	2.08	336.40	2.69	347.10	2.10
Canada*	507.13	1.97	496.95	1.56	515.02	1.93
Chile	447.98	2.63	411.70	2.08	443.54	2.47
Chinese Taipei	515.17	3.25	547.09	3.78	537.38	3.31
Colombia	408.67	3.75	382.70	3.03	411.12	3.28
Costa Rica	415.23	2.66	384.58	1.89	410.99	2.42
Croatia	475.50	2.44	463.11	2.38	482.67	2.40
Cyprus	381.08	1.16	418.31	1.18	410.90	1.46
Czech Republic	488.60	2.25	487.00	2.09	497.74	2.30
Denmark	488.80	2.58	489.27	1.95	493.82	2.50
Dominican Republic	351.31	2.44	339.11	1.62	360.43	2.04
El Salvador	364.90	2.80	343.47	2.00	373.14	2.62
Estonia	511.03	2.36	509.95	1.98	525.81	2.07
Finland	490.22	2.26	484.14	1.86	510.96	2.50
France	473.85	3.07	473.94	2.49	487.23	2.73
Georgia	373.86	2.29	390.02	2.37	384.07	2.31
Germany	479.79	3.61	474.83	3.06	492.43	3.48

Country	Reading		Mathematics		Science	
	Average PV	SE	Average PV	SE	Average PV	SE
Greece	438.44	2.83	430.15	2.34	440.79	2.77
Guatemala	374.12	2.44	344.20	2.21	372.96	2.23
Hong Kong (China)	499.70	2.85	540.35	2.99	520.42	2.79
Hungary	472.97	2.83	472.78	2.51	485.89	2.71
Iceland	435.90	2.06	458.90	1.58	446.93	1.76
Indonesia	358.57	2.91	365.53	2.35	382.86	2.56
Ireland	516.01	2.33	491.65	2.02	503.85	2.26
Israel	473.83	3.49	457.90	3.27	464.75	3.38
Italy	481.60	2.68	471.26	3.09	477.46	3.18
Jamaica	409.63	4.21	377.42	3.14	402.93	3.88
Japan	515.85	3.18	535.58	2.93	546.63	2.80
Jordan	342.17	2.40	361.23	2.03	374.53	2.35
Kazakhstan	386.28	1.66	425.44	1.69	423.17	1.72
Korea	515.42	3.63	527.30	3.86	527.82	3.58
Kosovo	342.19	1.06	354.96	1.02	357.02	1.26
Latvia	474.57	2.46	483.16	2.03	493.84	2.30
Lithuania	471.83	2.21	475.15	1.84	484.46	2.33
Macao (China)	510.41	1.35	551.92	1.10	543.10	1.11
Malaysia	388.09	2.75	408.69	2.40	416.31	2.35
Malta	445.30	1.90	466.02	1.58	465.59	1.70
Mexico	415.36	2.92	395.03	2.27	409.89	2.42
Mongolia	378.42	2.25	424.59	2.57	412.38	2.36
Montenegro	405.02	1.35	405.60	1.12	403.13	1.21
Morocco	339.36	3.97	364.77	3.35	365.40	3.38
Netherlands	459.24	4.28	492.68	3.77	488.32	4.07
New Zealand	500.85	2.12	479.07	1.99	504.13	2.24
North Macedonia	358.52	0.81	388.58	0.87	379.88	0.93
Norway	476.52	2.54	468.45	2.06	478.23	2.37
Palestinian Authority	349.16	2.03	365.75	1.84	368.82	2.10
Panama	391.95	3.41	356.57	2.84	387.77	3.54
Paraguay	373.16	2.44	337.54	2.16	368.33	2.06
Peru	408.25	2.73	391.24	2.34	407.78	2.64
Philippines	346.55	3.40	354.72	2.58	356.17	3.11
Poland	488.71	2.74	488.96	2.27	499.16	2.55
Portugal	476.59	2.66	471.91	2.35	484.37	2.56
Qatar	419.30	1.45	414.11	1.14	432.40	1.48
Republic of Moldova	410.94	2.51	414.20	2.31	416.86	2.39
Romania	428.50	3.98	427.76	4.00	427.51	3.87
Saudi Arabia	382.55	1.99	388.78	1.76	390.39	1.96
Serbia	440.35	2.79	439.88	2.97	447.46	2.89
Singapore	542.55	1.87	574.66	1.23	561.43	1.33
Slovak Republic	446.86	3.10	463.99	2.89	462.27	3.03
Slovenia	468.54	1.64	484.53	1.24	499.96	1.45
Spain	474.31	1.65	473.14	1.50	484.53	1.60
Sweden	486.98	2.49	481.77	2.06	493.55	2.35
Switzerland	483.33	2.26	507.99	2.14	502.52	2.19
Thailand	378.66	2.82	393.95	2.68	409.26	2.78
Türkiye	456.08	1.85	453.15	1.59	475.94	1.93
Ukrainian regions (18 of 27)	427.53	3.93	440.85	4.06	450.19	3.78
United Arab Emirates	417.35	1.34	431.11	0.95	431.98	1.31
United Kingdom	494.40	2.37	488.98	2.22	499.67	2.38
United States	503.94	4.33	464.89	4.01	499.41	4.32

Country	Reading		Mathematics		Science	
	Average PV	SE	Average PV	SE	Average PV	SE
Uruguay	430.36	2.41	408.71	2.02	435.38	2.48
Uzbekistan	335.50	2.00	363.94	2.02	354.86	2.01
Viet Nam	461.89	3.94	469.40	3.93	472.38	3.59

Table 14.12a. Core domain inter-correlations for the main sample

DOMAIN		Reading	Science
Mathematics	Average	0.80	0.85
	Average (CBA)	0.80	0.86
	Average (PBA)	0.81	0.83
	Range	0.65 ~ 0.89	0.75 ~ 0.92
Reading	Average		0.79
	Average (CBA)		0.79
	Average (PBA)		0.81
	Range		0.67~ 0.88

Table 14.12b. Creative Thinking inter-correlations with core domains for the main sample

DOMAIN		Mathematics	Reading	Science
Creative Thinking	Average	0.68	0.68	0.67
	Range	0.53 ~ 0.80	0.55 ~ 0.83	0.54 ~ 0.80

Table 14.13. Domain inter-correlations for the financial literacy sample

DOMAIN		Mathematics	Reading
Financial Literacy	Average	0.86	0.84
	Range	0.80 ~ 0.90	0.79 ~ 0.88

Table 14.14. Domain inter-correlations by country/economy

Country	Mathematics & Reading	Mathematics & Science	Mathematics & Financial literacy	Mathematics & Creative Thinking	Reading & Science	Reading & Financial literacy	Reading & Creative Thinking	Science & Creative Thinking
Albania	0.69	0.76		0.66	0.67		0.58	0.60
Argentina	0.75	0.81			0.75			
Australia	0.80	0.86		0.65	0.78		0.63	0.64
Austria	0.84	0.90	0.88		0.85	0.86		
Baku (Azerbaijan)	0.75	0.82		0.64	0.72		0.63	0.63
Belgium*	0.82	0.90	0.89	0.69	0.82	0.85	0.69	0.69
Brazil	0.80	0.84	0.84	0.69	0.78	0.82	0.70	0.68
Brunei Darussalam	0.88	0.92		0.80	0.87		0.81	0.80
Bulgaria	0.83	0.86	0.87	0.76	0.80	0.85	0.74	0.74
Cambodia	0.79	0.78			0.75			
Canada*	0.75	0.80	0.85	0.56	0.72	0.81	0.55	0.54
Chile	0.79	0.86		0.61	0.78		0.58	0.57
Chinese Taipei	0.84	0.90		0.68	0.82		0.67	0.67
Colombia	0.80	0.86		0.69	0.77		0.68	0.68
Costa Rica	0.79	0.83	0.87	0.71	0.78	0.83	0.69	0.66
Croatia	0.79	0.86		0.67	0.77		0.67	0.68
Cyprus	0.75	0.82		0.72	0.74		0.70	0.68
Czech Republic	0.81	0.88	0.87	0.68	0.80	0.83	0.67	0.68
Denmark	0.79	0.87	0.87	0.62	0.77	0.84	0.61	0.61
Dominican Republic	0.80	0.80		0.64	0.77		0.67	0.62
El Salvador	0.80	0.81		0.67	0.76		0.66	0.66
Estonia	0.77	0.86		0.62	0.74		0.58	0.62
Finland	0.79	0.87		0.68	0.77		0.71	0.70
France	0.84	0.88		0.71	0.82		0.72	0.70
Georgia	0.75	0.81			0.74			
Germany	0.85	0.90		0.76	0.86		0.76	0.76
Greece	0.78	0.83		0.69	0.77		0.65	0.68
Guatemala	0.84	0.87			0.88			
Hong Kong (China)	0.79	0.84		0.63	0.76		0.61	0.60
Hungary	0.84	0.91	0.89	0.76	0.84	0.85	0.74	0.74
Iceland	0.77	0.85		0.67	0.77		0.68	0.68
Indonesia	0.78	0.77		0.57	0.72		0.55	0.54
Ireland	0.81	0.88			0.84			
Israel	0.81	0.88		0.76	0.80		0.74	0.73
Italy	0.76	0.84	0.82	0.64	0.75	0.79	0.62	0.61
Jamaica	0.84	0.86		0.67	0.82		0.71	0.69
Japan	0.81	0.88			0.84			
Jordan	0.72	0.80		0.66	0.74		0.68	0.66
Kazakhstan	0.65	0.75		0.53	0.71		0.62	0.59
Korea	0.76	0.85		0.59	0.74		0.59	0.61
Kosovo	0.78	0.83			0.77			
Latvia	0.79	0.88		0.57	0.78		0.55	0.57
Lithuania	0.81	0.88		0.71	0.81		0.69	0.69
Macao (China)	0.75	0.87		0.66	0.78		0.64	0.66
Malaysia	0.79	0.87	0.90	0.75	0.82	0.88	0.79	0.78
Malta	0.78	0.87		0.73	0.80		0.73	0.72
Mexico	0.82	0.86		0.66	0.80		0.67	0.66
Mongolia	0.79	0.87		0.71	0.78		0.69	0.70
Montenegro	0.79	0.86			0.77			
Morocco	0.77	0.83		0.72	0.75		0.70	0.68

Country	Mathematics & Reading	Mathematics & Science	Mathematics & Financial literacy	Mathematics & Creative Thinking	Reading & Science	Reading & Financial literacy	Reading & Creative Thinking	Science & Creative Thinking
Netherlands	0.86	0.90	0.90	0.72	0.85	0.88	0.74	0.71
New Zealand	0.81	0.88		0.69	0.85		0.71	0.71
North Macedonia	0.80	0.84		0.75	0.76		0.72	0.74
Norway	0.78	0.86	0.80		0.80	0.82		
Palestinian Authority	0.76	0.81		0.71	0.72		0.67	0.67
Panama	0.82	0.86		0.64	0.79		0.66	0.65
Paraguay	0.84	0.87			0.86			
Peru	0.82	0.86	0.86	0.71	0.79	0.87	0.70	0.68
Philippines	0.89	0.87		0.80	0.85		0.83	0.77
Poland	0.81	0.87	0.87	0.70	0.80	0.84	0.68	0.68
Portugal	0.81	0.87	0.85	0.70	0.80	0.83	0.70	0.69
Qatar	0.81	0.86		0.72	0.79		0.70	0.70
Republic of Moldova	0.83	0.87		0.70	0.81		0.74	0.71
Romania	0.86	0.90		0.78	0.85		0.77	0.77
Saudi Arabia	0.75	0.78	0.80	0.66	0.73	0.82	0.67	0.65
Serbia	0.81	0.87		0.70	0.79		0.68	0.70
Singapore	0.82	0.89		0.67	0.81		0.66	0.66
Slovak Republic	0.83	0.89		0.74	0.81		0.72	0.73
Slovenia	0.77	0.89		0.60	0.77		0.59	0.58
Spain	0.76	0.82	0.86	0.59	0.75	0.79	0.59	0.58
Sweden	0.81	0.88			0.84			
Switzerland	0.83	0.89			0.86			
Thailand	0.79	0.83		0.68	0.77		0.67	0.68
Türkiye	0.82	0.90			0.83			
Ukrainian regions (18 of 27)	0.79	0.86		0.72	0.79		0.67	0.72
United Arab Emirates	0.81	0.85	0.87	0.71	0.79	0.84	0.71	0.69
United Kingdom	0.81	0.86			0.79			
United States	0.83	0.89	0.89		0.87	0.86		
Uruguay	0.80	0.87		0.72	0.79		0.69	0.71
Uzbekistan	0.72	0.78		0.67	0.70		0.63	0.63
Viet Nam	0.77	0.82			0.75			

Note: Ukrainian regions (18 out of 27) administered the assessment.

*Denotes a country/economy for which the financial literacy domain was not fully sampled across the population; it is not a nationally-representative sample.

Table 14.15a. Mathematics content subscales inter-correlations

	MCCR1	MCQN2	MCSS3	MCUD4
Reading	0.71	0.72	0.63	0.71
Science	0.76	0.77	0.68	0.75
MCCR1		0.86	0.77	0.82
MCQN2			0.79	0.85
MCSS3				0.76

Table 14.15b. Mathematics cognitive process subscales inter-correlations

	MPEM1	MPFS2	MPIN3	MPRE4
Reading	0.72	0.66	0.73	0.69
Science	0.77	0.71	0.77	0.74
MPEM1		0.83	0.87	0.84
MPFS2			0.81	0.79
MPIN3				0.82

Table 14.16. Linking error for score comparisons between PISA 2022 and previous PISA cycles

Comparison	Mathematics	Reading	Science	Financial literacy
PISA 2000 to 2022		6.67		
PISA 2003 to 2022	5.55	5.25		
PISA 2006 to 2022	4.09	8.56	3.68	
PISA 2009 to 2022	4.28	4.66	5.92	
PISA 2012 to 2022	3.58	6.01	5.20	4.05
PISA 2015 to 2022	2.74	3.63	1.38	3.47
PISA 2018 to 2022	2.24	1.47	1.61	2.20

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

Note by the Republic of Türkiye

The information in this document with reference to “Cyprus” relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Türkiye recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Türkiye shall preserve its position concerning the “Cyprus issue”.

Note by all the European Union Member States of the OECD and the European Union

The Republic of Cyprus is recognised by all members of the United Nations with the exception of Türkiye. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

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