CHAPTER 13

Estimating Standard Errors in the PIRLS 2021 Results

Pierre Foy
Ahmed Almaskut

To obtain estimates of students' proficiency in reading that are both accurate and cost-effective, PIRLS 2021 made extensive use of probability sampling techniques to sample students from the national fourth grade student populations and applied matrix-sampling assessment designs to target individual students with a subset of the complete pool of assessment items. This approach made efficient use of resources, particularly keeping student response burden to a minimum, but at the cost of some variance or uncertainty in the reported statistics, such as the means and percentages computed to estimate population parameters.

To quantify this uncertainty, each statistic in *PIRLS 2021 International Results in Reading* is accompanied by an estimate of its standard error. Statistics based on differences between two estimated results also have standard errors, which serve to calculate confidence intervals or to perform statistical tests of significance. For statistics reporting student achievement, which are based on plausible values, standard errors are calculated based on two components. The first reflects the uncertainty due to generalizing from a student sample to the entire student population from which it was drawn, referred to as *sampling variance*. The second is known as *imputation variance* and reflects uncertainty due to inferring students' achievement estimates from their observed performance on a subset of achievement items and other achievement-related information. This imputation variance reflects the posterior variance of the achievement variables given all available information used in the achievement imputation model described in Chapter 10. For reported statistics that are not based on plausible values, the estimates of standard errors are based entirely on sampling variance.

Estimating Sampling Variance

PIRLS makes extensive use of probability sampling to derive achievement results from national samples of students. Because many such samples are possible but only one sample is drawn, some uncertainty about how well the sample represents the population is expected. The uncertainty caused by sampling students from a target population, known as *sampling variance*, can be estimated from the data of the one sample drawn.

https://doi.org/10.6017/lse.tpisc.tr2103.kb1576

Whereas estimating the sampling variance from simple random samples is a relatively simple task, estimating the sampling variance from the complex sample design of PIRLS is a more challenging endeavor. A common way to estimate the sampling variance in multistage cluster sampling designs is through resampling schemes (Efron, 1982) such as the balanced repeated replication and Jackknife techniques (Johnson & Rust, 1992; Quenouille, 1949; Tukey, 1958; Wolter, 1985). PIRLS uses a variation of the Jackknife, Jackknife Repeated Replication (JRR), to estimate sampling variances. JRR was chosen because it is computationally straightforward and provides approximately unbiased estimates of the sampling variance of means, totals, and percentages.

At the core of the JRR technique is the repeated resampling from the observed sample, under identical sample design conditions. In the context of PIRLS, this entails the grouping of primary sampling units into sampling zones based on the PIRLS sample design and repeated draws of subsamples from these zones. The main features of the PIRLS sample design that JRR incorporates in its repeated replication are the stratification of schools and the clustering of students within schools. This was done by defining Jackknife sampling zones as pairs of successive schools[1] to model the stratification and clustering from the national samples (see Chapter 3 for information on the PIRLS sample design). The repeated subsampling required by JRR is applied within each sampling zone.

Jackknife sampling zones of pairs of schools were constructed within explicit strata for each country or benchmarking participant. When an explicit stratum had an odd number of schools, either by design or because of non-responding schools, the students in the lone school of the last sampling zone were divided randomly to make up two "quasi" schools for the purposes of calculating jackknife standard errors.[2] This resulted in each sampling zone consisting of either two schools or two "quasi" schools. While national samples required a minimum of 150 schools, many exceeded that minimum, so a total of 125 zones were created, allowing for as many as 250 sampled schools. If more than 250 schools were sampled, the additional zones were collapsed into the first 125 zones.[3]

Exhibit 13.1 shows the effective school sample size and number of constructed Jackknife sampling zones, prior to collapsing, for the participating countries and benchmarking participants in PIRLS 2021.[4] It includes the full PIRLS samples—either paper or digital—as well as the bridge samples of digital countries.

1 When schools were sampled, they were ordered within explicit strata by implicit stratification variables and their measure of size. Based on this sorting, successively sampled schools had similar stratification attributes. More information can be found in Appendix 3A of Chapter 3.

2 If a remaining school consisted of two sampled classrooms, each classroom became a "quasi" school.

3 The randomization used in the resampling within sampling zones preserves the sampling variance measured in the original sampling zones after collapsing.

4 Note that Jackknife sampling zones may be constructed in a different manner under specific national conditions. Country-specific information on the construction of Jackknife sampling zones is available in Appendix 8A of Chapter 8.

Exhibit 13.1: Number of Schools and Jackknife Sampling Zones in the PIRLS 2021 Samples

Country	PIRLS Sample		Bridge Sample	
	Schools	Zones	Schools	Zones
Albania	177	90	—	—
Australia	281	142	—	—
Austria	160	80	—	—
Azerbaijan	184	92	—	—
Bahrain	186	120	—	—
Belgium (Flemish)	141	72	48	26
Belgium (French)	158	79	—	—
Brazil	187	98	—	—
Bulgaria	151	76	—	—
Chinese Taipei	184	92	68	34
Croatia	154	78	48	25
Cyprus	160	81	—	—
Czech Republic	196	101	58	29
Denmark	197	99	60	31
Egypt	192	96	—	—
England	162	82	—	—
Finland	219	111	62	31
France	184	92	—	—
Georgia	190	97	—	—
Germany	252	127	74	38
Hong Kong SAR	144	73	—	—
Hungary	157	80	52	27
Iran, Islamic Rep. of	218	109	—	—
Ireland	148	75	—	—
Israel	194	98	77	40
Italy	164	83	58	29
Jordan	216	119	—	—
Kazakhstan	267	134	122	62
Kosovo	150	76	—	—
Latvia	156	87	—	—
Lithuania	190	97	68	34

Exhibit 13.1: Number of Schools and Jackknife Sampling Zones in the PIRLS 2021 Samples (Continued)

Country	PIRLS Sample		Bridge Sample	
	Schools	Zones	Schools	Zones
Macao SAR	63	190	—	—
Malta	78	40	22	11
Montenegro	140	254	—	—
Morocco	266	151	—	—
Netherlands	131	69	—	—
New Zealand	184	96	65	33
North Macedonia	148	75	—	—
Northern Ireland	143	75	—	—
Norway (5)	158	80	55	28
Oman	215	110	—	—
Poland	150	75	—	—
Portugal	196	99	88	45
Qatar	259	150	66	34
Russian Federation	204	63	92	33
Saudi Arabia	142	72	51	27
Serbia	169	85	—	—
Singapore	183	187	60	30
Slovak Republic	169	87	73	38
Slovenia	160	81	51	26
South Africa	321	161	—	—
Spain	452	230	74	37
Sweden	146	75	49	26
Turkiye	192	96	—	—
United Arab Emirates	663	736	92	47
United States	—	—	78	41
Uzbekistan	180	91	—	—
Benchmarking Participants				
Alberta, Canada	116	58	—	—
British Columbia, Canada	179	90	—	—
Newfoundland & Labrador, Canada	133	67	—	—
Quebec, Canada	112	58	—	—

Exhibit 13.1: Number of Schools and Jackknife Sampling Zones in the PIRLS 2021 Samples (Continued)

Country	PIRLS Sample		Bridge Sample	
	Schools	Zones	Schools	Zones
Moscow City, Russian Federation	174	88	66	33
South Africa (6)	253	129	—	—
Abu Dhabi, UAE	262	287	—	—
Dubai, UAE	190	220	—	—

For estimating the sampling variance, the JRR procedure draws two subsamples from each sampling zone: one where the first school in the pair is included and the second school is removed, and the other where the second school is included and the first school is removed. When a school is removed from a sampling zone, the sampling weights of the students in the remaining school are doubled to make up for the omitted school. In both subsamples, all students in the other sampling zones are included with their sampling weights unchanged. With this process applied in each of the 125 sampling zones, the JRR procedure yields a total of 250 replicate subsamples, each with its own set of replicate sampling weights to account for the successive removal of each school from the pair of schools in any given sampling zone.

The process of creating replicate sampling weights for the replicate subsamples defines replicate factors k_{hi} as follows:

$$k_{hi} = \begin{cases} 2 \text{ for students in school } i \text{ of sampling zone } h \\ 0 \text{ for students in the other school of sampling zone } h \\ 1 \text{ for students in any other sampling zone} \end{cases}$$

These replicate factors are used to compute the 250 sets of replicate sampling weights as follows:

$$W_{hij} = k_{hi} \cdot W_{0j}$$

where W_{0j} is the overall sampling weight of student j and W_{hij} is the resulting replicate sampling weight of student j when school i from sampling zone h is included and the other school in the pair is removed.

Exhibit 13.2 illustrates the calculation of the replicate factors necessary to produce the replicate sampling weights. Within each sampling zone, each school is randomly assigned an indicator u_{hi}, coded either 0 or 1, such that one school has a value of 0 and the other a value of 1. This indicator determines how schools within each zone will be successively included and removed. When a school is removed from a zone, the replicate factor is set to 0 and the sampling weights of all students in that school are set to 0. When a school is included, the replicate factor is

set to 2 and the sampling weights of all students in that school are doubled. The sampling weights of students in all the other sampling zones remain unchanged.

Exhibit 13.2: Construction of Replicate Factors Across Sampling Zones

Sampling Zone	School Replicate Indicator (u_{hi})	Replicate Factors for Computing JRR Replicate Sampling Weights (k_{hi})											
		Zone 1		Zone 2		Zone 3		...	Zone *h*		...	Zone 125	
		(1)	(2)	(3)	(4)	(5)	(6)		(2*h*-1)	(2*h*)		(249)	(250)
1	0	2	0	1	1	1	1	⋯	1	1	⋯	1	1
	1	0	2										
2	0	1	1	2	0	1	1	⋯	1	1	⋯	1	1
	1			0	2								
3	0	1	1	1	1	2	0	⋯	1	1	⋯	1	1
	1					0	2						
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋮
h	0	1	1	1	1	1	1	⋯	2	0	⋯	1	1
	1								0	2			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋮
125	0	1	1	1	1	1	1	⋯	1	1	⋯	2	0
	1											0	2

For example, sampling Zone 1 yields two sets of replicate sampling weights, hence the two columns for Zone 1. The first set has doubled sampling weights (k_{11} = 2) for the students in the first school (u_{11} = 0) of Zone 1, zeroed sampling weights (k_{12} = 0) for the students in the second school (u_{12} = 1) of Zone 1, and unchanged sampling weights (k_{hi} = 1) for all students in the other sampling zones, e.g., Zones 2 through 125. This is shown in the first Zone 1 column. The second set of replicate sampling weights (shown in the second Zone 1 column) has zeroed sampling weights (k_{11} = 0) for the students in the first school (u_{11} = 0) of Zone 1, doubled sampling weights (k_{12} = 2) for the students in the second school (u_{12} = 1) of Zone 1, and unchanged sampling weights (k_{hi} = 1) for all students in the other sampling zones.

The process is repeated across all 125 possible sampling zones, generating 250 sets of replicate sampling weights. The replicate sampling weights are then used to estimate any statistic of interest 250 times. The variation across these 250 jackknife estimates determines the sampling variance.

Given a statistic t to be computed from a national sample, the formula used to estimate the sampling variance of that statistic, based on the PIRLS JRR algorithm, is given by the following equation:

$$Var_{jrr}(t_0) = \frac{1}{2}\sum_{h=1}^{125}\sum_{i=1}^{2}(t_{hi} - t_0)^2 \quad (13.1)$$

where the term t_0 denotes the statistic of interest estimated with the overall student sampling weights W_{0j} and the term t_{hi} denotes the same statistic computed using the set of replicate sampling weights W_{hij} obtained from sampling zone h (h = 1, …, 125), where the i^{th} school (1st or 2nd) in the zone is included and the other removed. Efron (1982) provides a proof of why the variance can be calculated based on these squared deviations of the t_{hi} from the total sample statistics in jackknife-based resampling schemes.

The sampling variance estimated with the PIRLS JRR method properly accounts for the variation arising from having sampled students using the PIRLS 2021 multistage stratified cluster sample design. Its square root is an estimate of the standard error for any statistic derived from variables other than those based on plausible values. Examples of such statistics include the mean age of students and the percentage of students with at least one parent with a university degree.

Estimating Imputation Variance

For variables other than those based on plausible values, standard errors are estimated solely using sampling variation, and were computed using the JRR technique. However, the situation with achievement estimates is more complex. Achievement estimates are based on observations of how students perform on a subset of the PIRLS 2021 items. Any estimate of achievement based on a finite set of observed variables is affected by measurement error. As described in PIRLS 2021 Assessment Design (Martin et al., 2019), the PIRLS 2021 item pool was far too extensive to be administered in its entirety to any one student, and a matrix-sampling assessment design was adopted whereby each student was given a single test booklet containing only a part of the entire assessment. The results from all students and booklets were then analyzed using item response theory to provide estimates of achievement on the PIRLS 2021 reporting scale. To generalize to the full assessment, an imputation (Rubin, 1987) model that incorporates performance on PIRLS 2021 of each student and information about similarities between students was applied. This imputation model is a latent regression model described in Chapter 10 and was used to derive estimates of student performance (plausible values). Student proficiency estimates incorporate uncertainty that can be quantified through measurement error. Proficiency estimates are associated with variability due to measurement error, even after accounting for context variables to reduce this uncertainty using a latent regression. PIRLS 2021 followed the customary procedure of generating

five imputations, or plausible values, for each student and using the variability among them to measure that uncertainty, known as *imputation variance*.

The general procedure for estimating the imputation variance when analyzing student achievement data follows the basic principle of performing any statistical analysis five times—once for each set of plausible values—and aggregating the five sets of results (Mislevy et al., 1992). Thus, for any given achievement-based statistic t, estimating that statistic from each plausible value yields five estimates t_m, m = 1, …, 5, all computed using the overall student sampling weights W_{0j}. The final estimate of that statistic, t_0, is the average of these five estimates:

$$t_0 = \frac{1}{5}\sum_{m=1}^{5} t_m.$$

The imputation variance of the statistic t_0 is simply the variance of the five results from the plausible values, computed as follows:

$$Var_{imp}(t_0) = \frac{6}{5}\sum_{m=1}^{5}\frac{(t_m - t_0)^2}{4}$$

where the factor $\frac{6}{5}$ is a correction factor required by the multiple imputation methodology (Rubin, 1987). This imputation variance is then added to the sampling variance to estimate the total variance of the statistic t_0, as follows:

$$Var_{tot}(t_0) = Var_{jrr}(t_0) + Var_{imp}(t_0). \tag{13.2}$$

The sampling variance $Var_{jrr}(t_0)$ in this context is the average of the sampling variances from the five plausible values $Var_{jrr}(t_m)$, m = 1, …, 5, as follows:

$$Var_{jrr}(t_0) = \frac{1}{5}\sum_{m=1}^{5} Var_{jrr}(t_m)$$

where

$$Var_{jrr}(t_m) = \frac{1}{2}\sum_{h=1}^{125}\sum_{i=1}^{2}(t_{mhi} - t_m)^2$$

and t_{mhi} is the appropriate JRR estimate for plausible value m and computed using the set of replicate sampling weights of sampling Zone h where school i is included. The square root of the total variance is then the standard error estimate for any statistic based on plausible values, such as the average PIRLS reading achievement for girls, or the percentage of students at or above the PIRLS Advanced International Benchmark of reading achievement.

Appendix 13A provides details on the jackknife sampling variance, the imputation variance, the total variance, and the overall standard error for each country's mean proficiency estimates

in reading at the fourth grade, including the purposes and processes domains. Appendix 13B provides the same details for the bridge samples.[5]

Estimating Standard Errors for International Averages

Some exhibits in *PIRLS 2021 International Results in Reading* include international averages and their standard errors. For example, Exhibit 1.5 reports the international average for the percentages of girls and boys and their fourth grade reading achievement. International averages are computed using the data from participating countries included in the main table of the exhibit. Results from the benchmarking participants are not included in the estimation of international averages.

For any given statistic t_0, its international average is given by

$$t_{intl} = \frac{1}{N}\sum_{k=1}^{N} t_{0k}$$

where N is the number of countries contributing to the international average and t_{0k} is the estimate of our statistic of interest for country k.

The total variance of the international average t_{intl} is given by

$$Var_{tot}(t_{intl}) = \frac{1}{N^2}\sum_{k=1}^{N} Var_{tot}(t_{0k}) \qquad (13.3)$$

where $Var_{tot}(t_{0k})$ is the total variance of our statistic of interest for country *k*. For statistics based on plausible values, the total variance includes the sampling variance and the imputation variance, as given in equation (13.2) above. For statistics not based on plausible values, such as percentages, the total variance is based entirely on the sampling variance, as shown in equation (13.1) above. The standard error of the international average is the square root of the total variance.

Estimating Standard Errors for Comparing Independent Results

Standard errors, along with providing a measure of uncertainty for PIRLS results, can also serve to perform a null hypothesis significance test when comparing two or more results. A basic objective of PIRLS is to provide fair and accurate comparisons of student achievement across PIRLS assessment cycles. Exhibit 2.1.2 in *PIRLS 2021 International Results in Reading* is one such example, showing fourth grade reading trend comparisons for the PIRLS 2021 countries across the PIRLS assessment cycles. The report also includes comparisons of results across the participating countries in PIRLS 2021. Exhibit 1.2 shows all pairwise country comparisons for fourth

5 Information on the bridge samples is available in Chapter 3.

grade reading achievement. All of these null hypothesis significance tests require the computation of a standard error for comparing the difference between two estimates, which has an expected value of zero (indicating no difference).

PIRLS 2021 results were reported by way of a statistic such as a mean or percentage, and each statistic is accompanied by its standard error, computed using either equation (13.1) or equation (13.2), as appropriate. When comparing results, either across assessment cycles or between countries, it is necessary to compute the standard error of the difference between two results. Because national samples are drawn independently of each other within an assessment cycle, as well as between assessment cycles, computing the standard error of a difference is straightforward.

When computing the difference between two PIRLS results t_A and t_B on the same PIRLS scale, such as comparing the reading achievement of countries A and B, or comparing the reading achievement of a country between assessment cycles A and B, the standard error of that difference is given by

$$SE(t_A - t_B) = \sqrt{Var_{tot}(t_A) + Var_{tot}(t_B)}$$

or, more simply

$$SE(t_A - t_B) = \sqrt{SE(t_A)^2 + SE(t_B)^2}$$

which can be stated as follows: the standard error of the difference between two independent results is the square root of the sum of their respective squared standard errors.

Estimating Standard Errors for Comparing Dependent Results

In the context of PIRLS, dependent results are statistics derived from the same national or benchmarking sample. The achievement difference between girls and boys, as shown in Exhibit 1.5 in *PIRLS 2021 International Results in Reading*, is an example of two dependent results and their difference. This dependence occurs because girls and boys are selected from the same sample of classrooms and schools. Attributes from girls and boys from the same school tend to be more similar compared to subgroups selected from different schools, thus inducing a correlation that needs to be accounted for in the computation of the standard error of their difference.

The difference between two statistics is itself a statistic. With this in mind, the standard error of any difference between two dependent results is computed in the same way as any other statistic, as was described earlier. The 250 sets of replicate weights produce 250 replicate estimates of the difference of interest and equations (13.1) and (13.2) apply.

Estimating Standard Errors for Comparing Against International Average

When comparing a country's result to the international average, PIRLS accounts for the fact that the country contributed to the international average and standard error. To correct for this contribution, the standard error of the difference needs to be adjusted. The total variance of the difference $t_k - t_{intl}$, comparing country k to the international average for a statistic t, is given by

$$Var_{tot}(t_k - t_{intl}) = Var_{tot}(t_{intl}) + \frac{(N-1)^2 - 1}{N^2} Var_{tot}(t_k) \quad (13.4)$$

where N is the number of countries contributing to the international average, $Var_{tot}(t_{intl})$ is the total variance of the international average as computed by equation (13.3), and $Var_{tot}(t_k)$ is the total variance for country k as computed by equation (13.2) for results based on plausible values, or equation (13.1) for results not based on plausible values.

Equation (13.4) can be simplified and expressed in terms of standard errors as follows:

$$SE(t_k - t_{intl}) = \sqrt{SE(t_{intl})^2 + \frac{N-2}{N} SE(t_k)^2}$$

where $SE(t_{intl})$ is the standard error of the international average and $SE(t_k)$ is the standard error for country k.

References

Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

Johnson, E. G., & Rust, K. F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics, 17*(2), 175-190.

Martin, M. O., von Davier, M., Foy, P., & Mullis, I. V. S. (2019). PIRLS 2021 assessment design. In I. V. S. Mullis & M. O. Martin's (Eds.), *PIRLS 2021 Assessment Frameworks*. Boston College, TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/pirls2021/frameworks/

Mislevy, R. J., Beaton, A., Kaplan, B. A., & Sheehan, K. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, *29*(2), 133-161.

Quenouille, M. H. (1949). Problems in plane sampling. *The Annals of Mathematical Statistics, 20*(3), 355–375. https://doi.org/10.1214/aoms/1177729989

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.

Tukey, J. W. (1958). Bias and confidence in not quite large samples. *The Annals of Mathematical Statistics, 29*(2), 614. https://doi.org/10.1214/aoms/1177706647

Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer-Verlag.

Appendix 13A: Summary Statistics and Standard Errors for Proficiency in Reading

Summary Statistics and Standard Errors for Proficiency in Overall Reading

Country	Sample Size	Overall Reading				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Albania	4,213	512.743	7.364	2.228	9.592	3.097
Australia	5,487	540.134	4.551	0.333	4.884	2.210
Austria	4,806	529.763	3.653	1.062	4.715	2.171
Azerbaijan	5,209	440.275	11.189	1.716	12.905	3.592
Bahrain	5,208	458.395	7.116	1.121	8.237	2.870
Belgium (Flemish)	5,114	510.723	4.344	0.965	5.309	2.304
Belgium (French)	4,279	494.266	5.021	2.042	7.063	2.658
Brazil	4,941	418.877	26.589	1.828	28.417	5.331
Bulgaria	4,043	539.894	8.522	0.554	9.076	3.013
Chinese Taipei	5,555	543.759	4.044	0.596	4.641	2.154
Croatia	3,937	556.554	6.015	0.331	6.347	2.519
Cyprus	4,589	510.867	7.309	0.930	8.238	2.870
Czech Republic	6,621	539.652	5.035	0.204	5.240	2.289
Denmark	4,821	539.002	3.964	1.016	4.981	2.232
Egypt	7,979	378.233	25.709	3.542	29.251	5.408
England	4,150	557.553	5.483	0.631	6.114	2.473
Finland	7,018	549.300	5.189	0.585	5.774	2.403
France	5,339	513.737	5.620	0.853	6.473	2.544
Georgia	5,241	494.026	5.888	0.875	6.763	2.601
Germany	4,611	523.974	3.655	0.862	4.517	2.125
Hong Kong SAR	3,830	572.821	6.578	0.922	7.501	2.739
Hungary	5,312	539.422	10.140	1.599	11.740	3.426
Iran, Islamic Rep. of	5,962	412.775	22.075	1.726	23.800	4.879
Ireland	4,663	577.328	5.678	0.437	6.114	2.473
Israel	4,890	510.054	4.867	0.051	4.918	2.218
Italy	5,440	537.178	4.012	0.808	4.820	2.195
Jordan	6,150	380.527	26.812	2.359	29.171	5.401

Summary Statistics and Standard Errors for Proficiency in Overall Reading (Continued)

Country	Sample Size	Overall Reading				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Kazakhstan	7,023	503.589	6.425	1.036	7.461	2.732
Kosovo	4,557	420.548	7.724	2.059	9.782	3.128
Latvia	4,369	527.683	6.527	0.271	6.798	2.607
Lithuania	4,623	552.220	4.971	0.163	5.134	2.266
Macao SAR	5,093	535.565	0.704	0.946	1.650	1.284
Malta	3,030	514.641	6.599	0.842	7.441	2.728
Montenegro	4,489	487.162	2.278	0.431	2.709	1.646
Morocco	7,017	372.390	17.073	3.409	20.482	4.526
Netherlands	4,313	527.238	5.912	0.420	6.332	2.516
New Zealand	5,557	521.467	4.715	0.604	5.318	2.306
North Macedonia	2,929	442.147	26.544	1.773	28.317	5.321
Northern Ireland	4,050	565.934	4.563	1.738	6.301	2.510
Norway (5)	5,382	538.790	3.587	0.371	3.958	1.989
Oman	5,321	429.480	12.300	1.399	13.699	3.701
Poland	4,179	549.121	4.007	0.748	4.755	2.180
Portugal	6,111	519.763	4.246	0.879	5.125	2.264
Qatar	5,258	484.726	13.256	0.791	14.047	3.748
Russian Federation	5,217	567.109	12.226	0.485	12.711	3.565
Saudi Arabia	4,778	448.549	10.876	2.060	12.935	3.597
Serbia	4,037	513.579	6.697	1.015	7.713	2.777
Singapore	6,719	587.138	9.192	0.664	9.857	3.140
Slovak Republic	4,841	529.072	5.220	2.236	7.456	2.731
Slovenia	5,110	519.658	2.659	0.792	3.451	1.858
South Africa	12,422	288.216	15.951	3.616	19.568	4.424
Spain	8,551	521.214	3.847	0.989	4.835	2.199
Sweden	5,175	543.519	4.138	0.426	4.565	2.136
Turkiye	6,032	496.447	10.199	1.292	11.491	3.390
United Arab Emirates	27,448	483.085	2.595	0.746	3.341	1.828
Uzbekistan	5,846	436.844	6.978	1.156	8.134	2.852

Summary Statistics and Standard Errors for Proficiency in Overall Reading (Continued)

Country	Sample Size	Overall Reading				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Benchmarking Participants						
Alberta, Canada	3,020	538.687	12.223	0.481	12.704	3.564
British Columbia, Canada	4,675	535.295	12.153	0.432	12.585	3.547
Newfoundland & Labrador, Can.	2,445	523.316	9.231	1.071	10.302	3.210
Quebec, Canada	3,739	551.002	6.129	1.326	7.455	2.730
Moscow City, Russian Fed.	5,745	598.235	3.286	1.113	4.399	2.097
South Africa (6)	9,317	384.269	18.734	1.567	20.301	4.506
Abu Dhabi, UAE	10,381	439.562	11.275	1.174	12.449	3.528
Dubai, UAE	7,711	551.860	1.828	0.424	2.252	1.501

Summary Statistics and Standard Errors for Proficiency in Literary Experience

Country	Sample Size	Literary Experience				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Albania	4,213	516.052	8.894	1.903	10.797	3.286
Australia	5,487	543.344	4.785	0.770	5.555	2.357
Austria	4,806	532.501	3.764	0.797	4.560	2.136
Azerbaijan	5,209	440.775	11.645	0.937	12.581	3.547
Bahrain	5,208	459.566	7.333	1.028	8.362	2.892
Belgium (Flemish)	5,114	511.352	5.008	2.429	7.436	2.727
Belgium (French)	4,279	498.939	5.492	1.052	6.545	2.558
Brazil	4,941	418.022	26.024	1.301	27.325	5.227
Bulgaria	4,043	543.893	8.829	1.958	10.787	3.284
Chinese Taipei	5,555	532.986	3.946	0.495	4.441	2.107
Croatia	3,937	567.383	7.666	0.245	7.911	2.813
Cyprus	4,589	517.034	7.323	0.795	8.117	2.849
Czech Republic	6,621	539.724	5.608	0.409	6.017	2.453
Denmark	4,821	545.530	4.947	1.809	6.756	2.599
Egypt	7,979	371.792	24.335	1.642	25.977	5.097
England	4,150	558.113	5.705	0.168	5.873	2.423
Finland	7,018	547.098	5.953	1.050	7.003	2.646
France	5,339	515.815	5.113	0.653	5.765	2.401
Georgia	5,241	501.060	6.410	1.391	7.801	2.793
Germany	4,611	529.469	3.862	1.748	5.609	2.368
Hong Kong SAR	3,830	564.353	6.632	0.604	7.237	2.690
Hungary	5,312	541.489	10.768	0.305	11.073	3.328
Iran, Islamic Rep. of	5,962	412.812	22.696	2.618	25.314	5.031
Ireland	4,663	583.761	5.845	0.426	6.271	2.504
Israel	4,890	515.279	6.345	1.345	7.690	2.773
Italy	5,440	536.424	4.261	2.163	6.424	2.534
Jordan	6,150	377.614	26.227	3.703	29.929	5.471
Kazakhstan	7,023	508.454	6.935	1.073	8.007	2.830
Kosovo	4,557	417.727	7.795	0.549	8.344	2.889
Latvia	4,369	527.320	7.501	1.202	8.703	2.950
Lithuania	4,623	552.423	5.078	2.079	7.157	2.675

Summary Statistics and Standard Errors for Proficiency in Literary Experience (Continued)

Country	Sample Size	Literary Experience				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Macao SAR	5,093	525.365	0.841	0.811	1.652	1.285
Malta	3,030	516.248	7.432	0.630	8.062	2.839
Montenegro	4,489	490.986	2.403	1.110	3.512	1.874
Morocco	7,017	371.969	16.008	2.907	18.915	4.349
Netherlands	4,313	528.059	5.825	2.194	8.019	2.832
New Zealand	5,557	523.358	4.762	1.221	5.982	2.446
North Macedonia	2,929	442.421	28.525	1.117	29.642	5.444
Northern Ireland	4,050	573.107	4.622	0.506	5.128	2.265
Norway (5)	5,382	538.063	3.744	0.187	3.931	1.983
Oman	5,321	425.036	12.612	1.649	14.261	3.776
Poland	4,179	551.912	4.447	0.831	5.278	2.297
Portugal	6,111	519.501	4.122	1.218	5.340	2.311
Qatar	5,258	480.728	14.902	0.177	15.079	3.883
Russian Federation	5,217	565.917	11.361	1.290	12.652	3.557
Saudi Arabia	4,778	443.895	11.468	1.475	12.944	3.598
Serbia	4,037	517.765	7.736	0.726	8.462	2.909
Singapore	6,719	591.457	9.654	0.568	10.222	3.197
Slovak Republic	4,841	530.231	5.857	1.037	6.894	2.626
Slovenia	5,110	521.659	2.626	1.810	4.436	2.106
South Africa	12,422	293.033	16.094	4.125	20.218	4.496
Spain	8,551	520.251	4.368	0.391	4.759	2.181
Sweden	5,175	545.489	4.479	1.833	6.313	2.513
Turkiye	6,032	494.691	11.093	2.191	13.284	3.645
United Arab Emirates	27,448	478.120	2.943	1.199	4.143	2.035
Uzbekistan	5,846	438.059	7.538	1.748	9.287	3.047

Summary Statistics and Standard Errors for Proficiency in Literary Experience (Continued)

Country	Sample Size	Literary Experience				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Benchmarking Participants						
Alberta, Canada	3,020	540.981	11.568	0.034	11.602	3.406
British Columbia, Canada	4,675	536.508	12.436	0.337	12.772	3.574
Newfoundland & Labrador, Can.	2,445	525.889	9.650	1.793	11.443	3.383
Quebec, Canada	3,739	560.587	7.305	1.326	8.631	2.938
Moscow City, Russian Fed.	5,745	597.499	3.212	0.522	3.734	1.932
South Africa (6)	9,317	382.442	19.390	1.983	21.373	4.623
Abu Dhabi, UAE	10,381	434.352	13.272	1.285	14.556	3.815
Dubai, UAE	7,711	550.059	2.325	0.217	2.542	1.594

Summary Statistics and Standard Errors for Proficiency in Acquire and Use Information

Country	Sample Size	Acquire and Use Information				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Albania	4,213	508.553	7.409	2.953	10.362	3.219
Australia	5,487	539.213	4.725	0.355	5.080	2.254
Austria	4,806	527.340	4.344	2.301	6.645	2.578
Azerbaijan	5,209	439.001	12.387	0.718	13.105	3.620
Bahrain	5,208	456.958	7.611	0.475	8.086	2.844
Belgium (Flemish)	5,114	510.098	4.377	1.116	5.493	2.344
Belgium (French)	4,279	489.877	5.352	0.564	5.916	2.432
Brazil	4,941	420.817	24.733	0.689	25.422	5.042
Bulgaria	4,043	537.557	9.229	0.451	9.680	3.111
Chinese Taipei	5,555	549.328	4.224	0.499	4.722	2.173
Croatia	3,937	552.529	6.126	0.428	6.554	2.560
Cyprus	4,589	504.771	7.355	1.298	8.653	2.942
Czech Republic	6,621	539.669	4.896	1.442	6.338	2.518
Denmark	4,821	536.246	3.893	0.473	4.367	2.090
Egypt	7,979	382.193	28.257	1.238	29.495	5.431
England	4,150	558.976	5.834	0.379	6.213	2.493
Finland	7,018	550.327	5.107	1.850	6.957	2.638
France	5,339	511.380	6.174	0.714	6.888	2.625
Georgia	5,241	485.601	6.313	1.510	7.823	2.797
Germany	4,611	521.637	3.683	0.878	4.561	2.136
Hong Kong SAR	3,830	582.489	6.616	0.650	7.266	2.696
Hungary	5,312	538.841	9.977	1.745	11.722	3.424
Iran, Islamic Rep. of	5,962	412.102	21.164	2.210	23.374	4.835
Ireland	4,663	573.741	5.793	0.072	5.865	2.422
Israel	4,890	507.701	4.665	0.507	5.173	2.274
Italy	5,440	537.796	3.730	0.489	4.219	2.054
Jordan	6,150	383.850	29.530	3.674	33.204	5.762
Kazakhstan	7,023	500.687	6.384	0.359	6.743	2.597
Kosovo	4,557	423.210	8.155	1.234	9.389	3.064
Latvia	4,369	529.221	6.621	0.883	7.503	2.739
Lithuania	4,623	552.682	4.856	1.223	6.079	2.466

Summary Statistics and Standard Errors for Proficiency in Acquire and Use Information (Continued)

Country	Sample Size	Acquire and Use Information				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Macao SAR	5,093	547.257	0.858	1.373	2.231	1.494
Malta	3,030	513.689	6.799	0.828	7.626	2.762
Montenegro	4,489	483.018	2.128	1.408	3.535	1.880
Morocco	7,017	372.994	17.964	2.243	20.207	4.495
Netherlands	4,313	528.084	6.464	2.124	8.588	2.931
New Zealand	5,557	520.657	4.919	1.369	6.288	2.508
North Macedonia	2,929	439.162	30.386	1.477	31.863	5.645
Northern Ireland	4,050	561.630	4.929	0.421	5.350	2.313
Norway (5)	5,382	540.130	3.661	0.657	4.318	2.078
Oman	5,321	432.138	13.407	1.115	14.522	3.811
Poland	4,179	547.938	4.362	0.558	4.920	2.218
Portugal	6,111	520.091	4.105	1.318	5.423	2.329
Qatar	5,258	485.698	13.118	0.644	13.762	3.710
Russian Federation	5,217	568.326	13.424	1.085	14.509	3.809
Saudi Arabia	4,778	450.619	12.110	1.462	13.572	3.684
Serbia	4,037	511.236	5.833	0.260	6.094	2.469
Singapore	6,719	586.401	9.386	0.330	9.717	3.117
Slovak Republic	4,841	529.667	5.719	1.114	6.834	2.614
Slovenia	5,110	518.755	2.952	1.416	4.368	2.090
South Africa	12,422	278.516	16.090	4.843	20.933	4.575
Spain	8,551	521.532	4.192	1.637	5.829	2.414
Sweden	5,175	543.541	4.266	0.354	4.620	2.149
Turkiye	6,032	498.080	10.547	0.880	11.427	3.380
United Arab Emirates	27,448	485.405	2.459	0.441	2.899	1.703
Uzbekistan	5,846	433.918	7.302	1.392	8.694	2.949

Summary Statistics and Standard Errors for Proficiency in Acquire and Use Information (Continued)

Country	Sample Size	Acquire and Use Information				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Benchmarking Participants						
Alberta, Canada	3,020	537.168	13.458	1.667	15.125	3.889
British Columbia, Canada	4,675	535.145	12.473	0.328	12.801	3.578
Newfoundland & Labrador, Can.	2,445	522.905	9.360	0.525	9.885	3.144
Quebec, Canada	3,739	548.071	5.862	1.057	6.919	2.630
Moscow City, Russian Fed.	5,745	599.867	3.367	0.248	3.615	1.901
South Africa (6)	9,317	384.269	18.046	3.670	21.717	4.660
Abu Dhabi, UAE	10,381	441.893	11.058	0.745	11.803	3.436
Dubai, UAE	7,711	552.949	1.789	0.414	2.203	1.484

Summary Statistics and Standard Errors for Proficiency in Retrieving and Straightforward Inferencing

Country	Sample Size	Retrieving and Straightforward Inferencing				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Albania	4,213	508.426	7.222	4.104	11.326	3.365
Australia	5,487	533.749	4.639	0.885	5.524	2.350
Austria	4,806	531.610	4.058	1.909	5.967	2.443
Azerbaijan	5,209	446.276	12.120	1.565	13.686	3.699
Bahrain	5,208	455.826	7.050	1.279	8.329	2.886
Belgium (Flemish)	5,114	511.041	4.578	0.236	4.814	2.194
Belgium (French)	4,279	496.573	5.439	0.209	5.648	2.377
Brazil	4,941	418.235	25.464	1.934	27.398	5.234
Bulgaria	4,043	541.071	8.659	1.216	9.875	3.142
Chinese Taipei	5,555	545.568	3.778	0.805	4.583	2.141
Croatia	3,937	552.416	5.837	1.044	6.882	2.623
Cyprus	4,589	508.534	6.245	0.125	6.369	2.524
Czech Republic	6,621	542.433	5.268	1.032	6.300	2.510
Denmark	4,821	539.022	4.064	0.451	4.515	2.125
Egypt	7,979	376.012	25.536	3.163	28.699	5.357
England	4,150	554.195	5.448	0.257	5.706	2.389
Finland	7,018	549.890	5.924	0.732	6.656	2.580
France	5,339	518.592	5.953	1.837	7.790	2.791
Georgia	5,241	488.505	5.512	0.072	5.584	2.363
Germany	4,611	524.691	3.518	0.754	4.273	2.067
Hong Kong SAR	3,830	577.289	7.772	0.635	8.408	2.900
Hungary	5,312	538.129	9.967	1.855	11.821	3.438
Iran, Islamic Rep. of	5,962	413.945	20.838	1.372	22.210	4.713
Ireland	4,663	571.248	5.145	0.253	5.398	2.323
Israel	4,890	507.892	4.740	0.483	5.223	2.285
Italy	5,440	536.736	3.898	1.849	5.747	2.397
Jordan	6,150	381.260	25.244	2.410	27.654	5.259
Kazakhstan	7,023	505.171	6.317	0.678	6.995	2.645
Kosovo	4,557	424.271	7.206	1.502	8.708	2.951
Latvia	4,369	524.650	6.871	0.155	7.026	2.651

Summary Statistics and Standard Errors for Proficiency in Retrieving and Straightforward Inferencing (Continued)

Country	Sample Size	Retrieving and Straightforward Inferencing				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Lithuania	4,623	554.377	4.845	1.554	6.399	2.530
Macao SAR	5,093	540.691	0.885	0.169	1.054	1.027
Malta	3,030	515.276	6.082	2.238	8.320	2.884
Montenegro	4,489	484.104	2.092	1.423	3.515	1.875
Morocco	7,017	373.840	15.658	1.326	16.984	4.121
Netherlands	4,313	526.712	6.086	1.622	7.708	2.776
New Zealand	5,557	520.884	4.491	0.843	5.333	2.309
North Macedonia	2,929	443.345	26.211	2.575	28.786	5.365
Northern Ireland	4,050	557.695	4.642	1.964	6.607	2.570
Norway (5)	5,382	540.184	3.803	0.193	3.996	1.999
Oman	5,321	425.532	11.590	1.359	12.949	3.599
Poland	4,179	545.295	4.068	0.664	4.732	2.175
Portugal	6,111	519.901	4.419	1.070	5.488	2.343
Qatar	5,258	485.915	12.826	0.841	13.668	3.697
Russian Federation	5,217	568.023	13.575	0.566	14.140	3.760
Saudi Arabia	4,778	450.071	11.414	0.349	11.763	3.430
Serbia	4,037	510.404	6.589	2.282	8.871	2.978
Singapore	6,719	583.768	8.735	0.200	8.935	2.989
Slovak Republic	4,841	530.478	5.680	0.823	6.503	2.550
Slovenia	5,110	519.576	2.977	0.809	3.785	1.946
South Africa	12,422	290.071	15.213	5.311	20.524	4.530
Spain	8,551	521.879	3.893	1.230	5.123	2.263
Sweden	5,175	545.628	3.917	1.233	5.150	2.269
Turkiye	6,032	499.345	10.714	2.133	12.848	3.584
United Arab Emirates	27,448	484.096	2.604	0.786	3.391	1.841
Uzbekistan	5,846	441.256	7.120	1.231	8.352	2.890

Summary Statistics and Standard Errors for Proficiency in Retrieving and Straightforward Inferencing (Continued)

Country	Sample Size	Retrieving and Straightforward Inferencing				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Benchmarking Participants						
Alberta, Canada	3,020	536.738	12.028	1.109	13.137	3.624
British Columbia, Canada	4,675	532.024	12.389	1.680	14.068	3.751
Newfoundland & Labrador, Can.	2,445	522.199	9.037	1.788	10.825	3.290
Quebec, Canada	3,739	551.329	6.370	0.634	7.004	2.647
Moscow City, Russian Fed.	5,745	601.981	3.647	0.238	3.885	1.971
South Africa (6)	9,317	385.763	18.273	1.618	19.891	4.460
Abu Dhabi, UAE	10,381	440.679	11.435	1.396	12.830	3.582
Dubai, UAE	7,711	549.827	1.920	0.755	2.675	1.636

Summary Statistics and Standard Errors for Proficiency in Interpreting, Integrating, and Evaluating

Country	Sample Size	Interpreting, Integrating, and Evaluating				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Albania	4,213	517.894	7.598	2.314	9.912	3.148
Australia	5,487	547.056	4.320	0.981	5.300	2.302
Austria	4,806	527.809	3.782	1.241	5.023	2.241
Azerbaijan	5,209	430.773	12.167	1.641	13.809	3.716
Bahrain	5,208	461.719	7.562	1.352	8.914	2.986
Belgium (Flemish)	5,114	510.254	4.478	0.949	5.428	2.330
Belgium (French)	4,279	492.261	4.919	0.924	5.843	2.417
Brazil	4,941	420.161	23.190	4.720	27.910	5.283
Bulgaria	4,043	540.926	8.741	0.901	9.642	3.105
Chinese Taipei	5,555	541.526	4.399	0.648	5.047	2.247
Croatia	3,937	561.157	6.369	0.713	7.082	2.661
Cyprus	4,589	512.376	7.396	3.199	10.594	3.255
Czech Republic	6,621	536.851	4.999	0.692	5.691	2.386
Denmark	4,821	539.779	3.983	0.714	4.697	2.167
Egypt	7,979	380.280	25.022	1.201	26.222	5.121
England	4,150	561.174	5.241	0.764	6.005	2.450
Finland	7,018	549.031	4.780	0.827	5.608	2.368
France	5,339	509.593	5.060	1.891	6.951	2.636
Georgia	5,241	499.859	6.595	0.939	7.534	2.745
Germany	4,611	522.061	3.809	0.325	4.134	2.033
Hong Kong SAR	3,830	572.345	6.059	0.949	7.008	2.647
Hungary	5,312	541.189	10.560	0.364	10.925	3.305
Iran, Islamic Rep. of	5,962	410.520	20.920	1.159	22.079	4.699
Ireland	4,663	581.850	5.745	1.596	7.341	2.709
Israel	4,890	512.010	5.286	1.823	7.109	2.666
Italy	5,440	537.913	3.928	1.025	4.952	2.225
Jordan	6,150	378.530	27.436	2.726	30.161	5.492
Kazakhstan	7,023	501.904	6.595	0.511	7.106	2.666
Kosovo	4,557	411.549	9.419	0.485	9.903	3.147
Latvia	4,369	532.065	5.828	1.545	7.373	2.715

Summary Statistics and Standard Errors for Proficiency in Interpreting, Integrating, and Evaluating (Continued)

Country	Sample Size	Interpreting, Integrating, and Evaluating				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Lithuania	4,623	551.037	5.368	1.726	7.094	2.664
Macao SAR	5,093	533.581	0.717	0.479	1.195	1.093
Malta	3,030	513.307	6.914	1.231	8.145	2.854
Montenegro	4,489	491.165	2.262	3.354	5.617	2.370
Morocco	7,017	366.405	18.714	3.772	22.487	4.742
Netherlands	4,313	528.967	5.936	0.589	6.525	2.554
New Zealand	5,557	522.313	4.976	0.829	5.806	2.410
North Macedonia	2,929	438.707	33.527	2.886	36.414	6.034
Northern Ireland	4,050	573.435	4.299	0.465	4.763	2.183
Norway (5)	5,382	537.598	3.572	1.979	5.551	2.356
Oman	5,321	433.423	13.234	2.132	15.366	3.920
Poland	4,179	552.378	3.658	0.177	3.835	1.958
Portugal	6,111	520.213	4.265	0.145	4.410	2.100
Qatar	5,258	481.910	14.100	0.087	14.187	3.767
Russian Federation	5,217	568.043	13.715	0.913	14.628	3.825
Saudi Arabia	4,778	443.334	11.796	2.365	14.161	3.763
Serbia	4,037	516.432	5.934	1.365	7.299	2.702
Singapore	6,719	590.960	9.517	0.594	10.110	3.180
Slovak Republic	4,841	529.186	5.948	0.689	6.637	2.576
Slovenia	5,110	519.339	2.443	0.900	3.343	1.828
South Africa	12,422	278.790	17.297	3.357	20.654	4.545
Spain	8,551	520.450	4.438	0.254	4.692	2.166
Sweden	5,175	542.132	4.401	0.231	4.632	2.152
Turkiye	6,032	494.082	10.901	0.394	11.296	3.361
United Arab Emirates	27,448	481.790	2.755	0.848	3.603	1.898
Uzbekistan	5,846	429.823	8.345	1.835	10.180	3.191

Summary Statistics and Standard Errors for Proficiency in Interpreting, Integrating, and Evaluating (Continued)

Country	Sample Size	Interpreting, Integrating, and Evaluating				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Benchmarking Participants						
Alberta, Canada	3,020	543.318	12.609	0.428	13.037	3.611
British Columbia, Canada	4,675	539.758	12.480	0.250	12.730	3.568
Newfoundland & Labrador, Can.	2,445	525.602	9.083	1.882	10.965	3.311
Quebec, Canada	3,739	552.371	5.214	0.845	6.059	2.461
Moscow City, Russian Fed.	5,745	597.215	3.208	0.389	3.597	1.896
South Africa (6)	9,317	380.933	19.582	0.398	19.980	4.470
Abu Dhabi, UAE	10,381	438.378	12.080	1.985	14.065	3.750
Dubai, UAE	7,711	553.864	1.945	1.125	3.070	1.752

Appendix 13B: Summary Statistics and Standard Errors for Proficiency in Reading for the Bridge Samples

Summary Statistics and Standard Errors for Proficiency in Overall Reading—Bridge Samples

Country	Sample Size	Overall Reading				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Belgium (Flemish)	1,623	502.624	12.956	0.231	13.187	3.631
Chinese Taipei	1,669	559.330	9.941	0.444	10.385	3.223
Croatia	1,226	553.078	18.076	5.229	23.306	4.828
Czech Republic	1,906	545.414	9.456	1.034	10.489	3.239
Denmark	1,403	539.466	11.366	1.243	12.609	3.551
Finland	2,069	544.051	15.913	3.189	19.102	4.371
Germany	1,343	532.215	13.020	0.201	13.221	3.636
Hungary	1,697	541.912	25.726	0.504	26.230	5.122
Israel	1,780	508.624	14.849	2.237	17.085	4.133
Italy	1,979	544.728	10.563	0.672	11.235	3.352
Kazakhstan	3,207	505.288	14.107	0.857	14.964	3.868
Lithuania	1,519	542.058	16.391	1.424	17.814	4.221
Malta	835	498.888	60.964	1.150	62.114	7.881
New Zealand	2,221	516.673	16.644	1.651	18.295	4.277
Norway (5)	1,673	523.283	11.347	1.469	12.816	3.580
Portugal	2,098	530.775	11.019	0.798	11.818	3.438
Qatar	1,343	481.560	58.944	1.781	60.725	7.793
Russian Federation	2,187	574.533	20.951	0.163	21.114	4.595
Saudi Arabia	1,872	417.320	68.150	5.886	74.036	8.604
Singapore	1,988	581.659	20.750	2.598	23.348	4.832
Slovak Republic	1,640	526.710	30.813	1.151	31.963	5.654
Slovenia	1,414	520.165	11.553	1.307	12.860	3.586
Spain	1,572	512.554	14.054	0.968	15.022	3.876
Sweden	1,863	535.915	15.141	3.297	18.438	4.294
United Arab Emirates	1,990	467.793	82.079	2.357	84.436	9.189
United States	1,657	547.544	45.369	0.284	45.653	6.757
Benchmarking Participants						
Moscow City, Russian Fed.	1,695	607.962	13.024	0.342	13.367	3.656

Summary Statistics and Standard Errors for Proficiency in Literary Experience—Bridge Samples

Country	Sample Size	Literary Experience				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Belgium (Flemish)	1,623	503.660	14.423	3.944	18.368	4.286
Chinese Taipei	1,669	548.951	9.943	0.840	10.783	3.284
Croatia	1,226	560.157	22.244	3.236	25.480	5.048
Czech Republic	1,906	550.062	10.451	0.327	10.779	3.283
Denmark	1,403	543.174	13.365	0.141	13.506	3.675
Finland	2,069	542.480	14.331	4.771	19.102	4.371
Germany	1,343	537.324	15.093	2.381	17.474	4.180
Hungary	1,697	544.946	27.270	3.084	30.354	5.509
Israel	1,780	515.905	16.872	2.224	19.096	4.370
Italy	1,979	547.676	12.271	1.254	13.525	3.678
Kazakhstan	3,207	504.259	16.526	0.468	16.994	4.122
Lithuania	1,519	541.392	15.381	1.365	16.746	4.092
Malta	835	502.288	43.549	4.109	47.658	6.904
New Zealand	2,221	521.159	18.802	1.397	20.199	4.494
Norway (5)	1,673	522.289	9.373	2.083	11.456	3.385
Portugal	2,098	531.925	13.221	2.253	15.474	3.934
Qatar	1,343	481.934	60.149	2.532	62.680	7.917
Russian Federation	2,187	572.602	16.438	2.272	18.709	4.325
Saudi Arabia	1,872	420.178	71.416	4.488	75.904	8.712
Singapore	1,988	590.045	24.909	5.030	29.939	5.472
Slovak Republic	1,640	532.341	31.637	0.661	32.298	5.683
Slovenia	1,414	520.452	12.658	1.087	13.745	3.707
Spain	1,572	513.549	14.155	1.097	15.252	3.905
Sweden	1,863	539.950	15.983	1.187	17.170	4.144
United Arab Emirates	1,990	466.915	84.808	0.769	85.578	9.251
United States	1,657	558.453	48.094	3.765	51.859	7.201
Benchmarking Participants						
Moscow City, Russian Fed.	1,695	609.052	12.750	3.636	16.387	4.048

Summary Statistics and Standard Errors for Proficiency in Acquire and Use Information—Bridge Samples

Country	Sample Size	Acquire and Use Information				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Belgium (Flemish)	1,623	501.419	12.430	2.260	14.690	3.833
Chinese Taipei	1,669	569.001	12.109	1.752	13.860	3.723
Croatia	1,226	547.025	21.262	1.508	22.771	4.772
Czech Republic	1,906	540.674	8.501	1.800	10.301	3.210
Denmark	1,403	533.703	13.702	2.418	16.120	4.015
Finland	2,069	547.744	18.222	2.639	20.861	4.567
Germany	1,343	525.730	14.200	1.952	16.152	4.019
Hungary	1,697	538.525	29.444	4.010	33.454	5.784
Israel	1,780	503.258	14.012	0.255	14.267	3.777
Italy	1,979	544.640	11.195	0.693	11.889	3.448
Kazakhstan	3,207	506.237	14.034	0.776	14.810	3.848
Lithuania	1,519	544.417	20.557	1.230	21.787	4.668
Malta	835	495.685	68.929	3.474	72.403	8.509
New Zealand	2,221	510.800	15.561	0.793	16.354	4.044
Norway (5)	1,673	524.736	14.640	4.096	18.736	4.329
Portugal	2,098	531.949	10.187	0.576	10.763	3.281
Qatar	1,343	483.755	60.707	3.234	63.940	7.996
Russian Federation	2,187	578.189	25.392	1.752	27.144	5.210
Saudi Arabia	1,872	413.220	65.940	1.009	66.948	8.182
Singapore	1,988	575.967	20.926	1.448	22.374	4.730
Slovak Republic	1,640	520.839	35.254	3.555	38.809	6.230
Slovenia	1,414	520.695	11.944	0.603	12.547	3.542
Spain	1,572	512.026	16.021	4.311	20.332	4.509
Sweden	1,863	533.532	15.638	1.780	17.418	4.173
United Arab Emirates	1,990	472.623	73.388	0.540	73.928	8.598
United States	1,657	539.671	43.464	2.292	45.756	6.764
Benchmarking Participants						
Moscow City, Russian Fed.	1,695	608.737	13.886	1.852	15.738	3.967

Summary Statistics and Standard Errors for Proficiency in Retrieving and Straightforward Inferencing—Bridge Samples

Country	Sample Size	Retrieving and Straightforward Inferencing				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Belgium (Flemish)	1,623	500.166	11.763	3.167	14.930	3.864
Chinese Taipei	1,669	559.627	9.039	2.087	11.127	3.336
Croatia	1,226	544.347	14.096	3.003	17.099	4.135
Czech Republic	1,906	548.435	9.901	0.667	10.569	3.251
Denmark	1,403	535.520	12.526	3.242	15.768	3.971
Finland	2,069	543.401	16.554	2.548	19.102	4.371
Germany	1,343	536.606	15.918	4.200	20.118	4.485
Hungary	1,697	537.086	25.813	1.081	26.894	5.186
Israel	1,780	500.766	13.006	1.280	14.286	3.780
Italy	1,979	539.661	10.337	1.417	11.754	3.428
Kazakhstan	3,207	503.609	13.882	2.729	16.611	4.076
Lithuania	1,519	544.550	17.712	0.594	18.306	4.279
Malta	835	497.997	60.126	2.323	62.448	7.902
New Zealand	2,221	515.598	15.313	0.931	16.245	4.030
Norway (5)	1,673	519.963	10.467	1.244	11.711	3.422
Portugal	2,098	531.518	11.161	1.206	12.366	3.517
Qatar	1,343	480.765	56.512	1.881	58.393	7.642
Russian Federation	2,187	577.249	23.798	1.810	25.608	5.060
Saudi Arabia	1,872	420.233	63.871	5.686	69.557	8.340
Singapore	1,988	575.281	17.802	1.122	18.924	4.350
Slovak Republic	1,640	526.324	31.050	3.697	34.747	5.895
Slovenia	1,414	523.385	15.610	2.147	17.758	4.214
Spain	1,572	507.430	15.468	2.149	17.617	4.197
Sweden	1,863	535.798	14.544	2.354	16.898	4.111
United Arab Emirates	1,990	465.639	74.506	0.844	75.351	8.680
United States	1,657	539.100	35.705	1.376	37.081	6.089
Benchmarking Participants						
Moscow City, Russian Fed.	1,695	606.573	12.077	0.492	12.569	3.545

Summary Statistics and Standard Errors for Proficiency in Interpreting, Integrating, and Evaluating—Bridge Samples

Country	Sample Size	Interpreting, Integrating, and Evaluating				
		Mean Proficiency	Jackknife Sampling Variance	Imputation Variance	Total Variance	Overall Standard Error
Belgium (Flemish)	1,623	505.997	11.829	2.622	14.451	3.801
Chinese Taipei	1,669	559.575	8.877	1.824	10.701	3.271
Croatia	1,226	560.507	15.876	1.577	17.453	4.178
Czech Republic	1,906	543.607	8.303	2.909	11.212	3.348
Denmark	1,403	541.554	12.177	0.216	12.393	3.520
Finland	2,069	546.394	13.721	3.580	17.301	4.159
Germany	1,343	528.136	13.760	2.669	16.429	4.053
Hungary	1,697	545.338	27.008	1.818	28.826	5.369
Israel	1,780	513.413	15.143	0.907	16.051	4.006
Italy	1,979	550.837	11.642	1.735	13.377	3.657
Kazakhstan	3,207	507.742	13.232	4.211	17.443	4.176
Lithuania	1,519	540.672	16.216	1.381	17.597	4.195
Malta	835	498.664	58.535	4.848	63.384	7.961
New Zealand	2,221	516.471	18.168	0.678	18.846	4.341
Norway (5)	1,673	525.949	11.554	1.662	13.216	3.635
Portugal	2,098	528.651	11.642	1.029	12.670	3.560
Qatar	1,343	482.450	58.135	1.295	59.430	7.709
Russian Federation	2,187	574.216	19.140	3.061	22.202	4.712
Saudi Arabia	1,872	409.230	83.019	6.527	89.546	9.463
Singapore	1,988	584.686	22.288	1.304	23.593	4.857
Slovak Republic	1,640	527.838	32.547	0.571	33.118	5.755
Slovenia	1,414	518.196	8.855	1.231	10.086	3.176
Spain	1,572	516.585	17.040	1.742	18.782	4.334
Sweden	1,863	536.300	16.981	0.626	17.608	4.196
United Arab Emirates	1,990	471.075	82.894	4.073	86.967	9.326
United States	1,657	555.257	46.896	4.203	51.099	7.148
Benchmarking Participants						
Moscow City, Russian Fed.	1,695	609.298	13.373	0.588	13.961	3.736