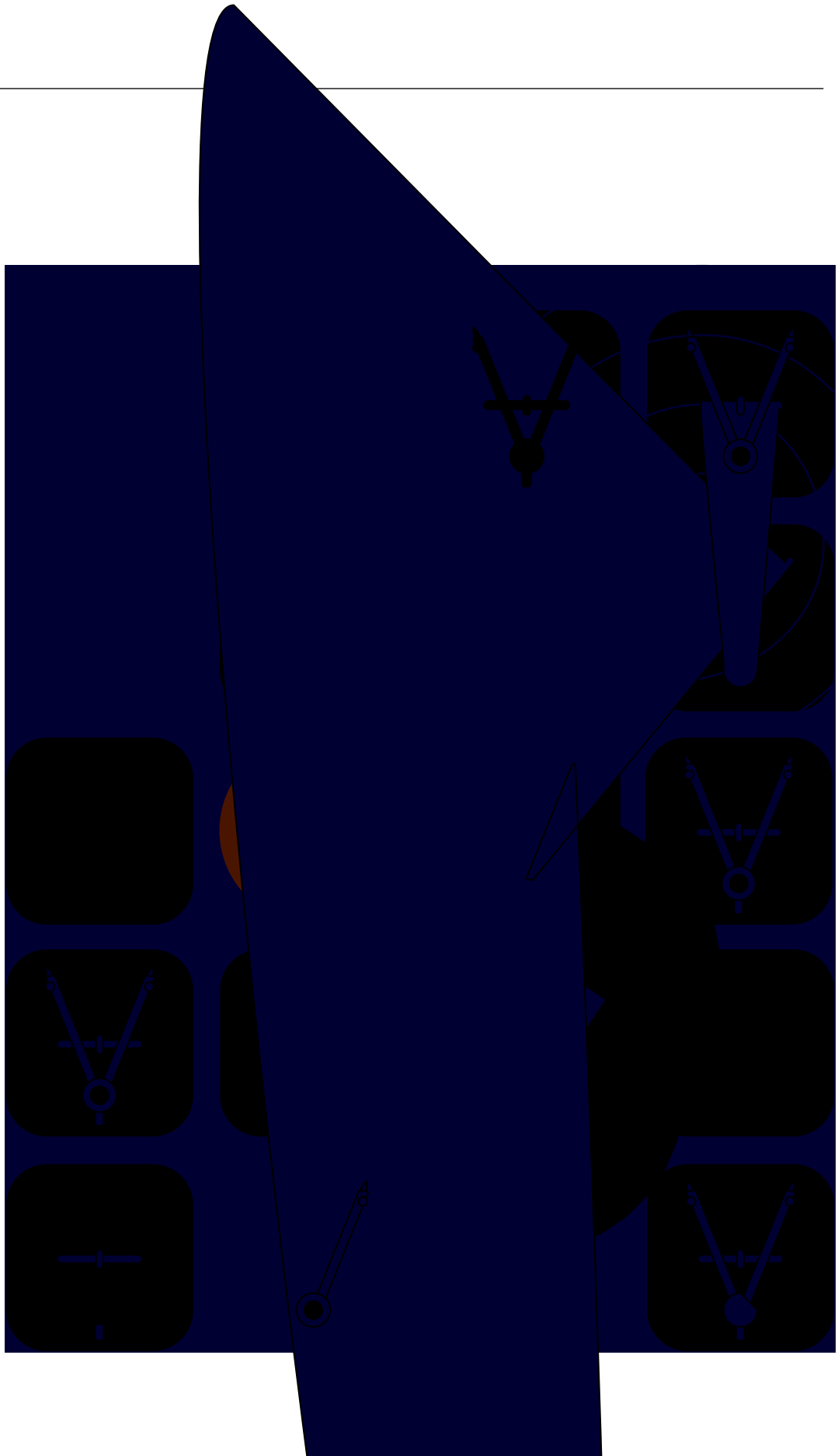

11 Estimation of Sampling and Imputation Variance for TIMSS 1999 Benchmarking[1]

Eugenio J. Gonzalez
Pierre Foy

11.1 Overview

To obtain estimates of student proficiency in mathematics and science that were both accurate and cost effective, TIMSS 1999

Exhibit 11.1 Range of Sampling Zones

Country	Zones
Australia	75

States	Zones
Connecticut	26

The computation of the JRR variance estimate for any statistic in TIMSS 1999 required the computation of the statistic up to 76 times for any given country or Benchmarking jurisdiction: once to obtain the statistic for the full sample, and up to 75 times to obtain the statistics for each of the jackknife replicates (J_h). The number of times a statistic needed to be computed for a given country depended on the number of implicit strata or sampling zones defined for that country.

Doubling and zeroing the weights of the selected units within the sampling zones was done by creating replicate weights that were then used in the calculations. This approach requires the user to create a new set of weights for each pseudo-replicate sample. Each replicate weight is equal to *k* times the overall sampling weight, where *k* can take values of 0, 1, or 2 depending on whether the case is to be removed from the computation, left as it is, or have its weight doubled. The value of *k* for an individual student record for a given replicate depends on the assignment of the record to the specific PSU and zone.

Within each zone the members of the pair of schools are assigned an indicator (u_i), coded randomly to 1 or 0 so that one of them has a value of 1 on the variable u_i and the other a value of 0. This indicator determines whether the weights for the elements in the school in this zone are to be doubled or zeroed. The replicate weight $(w_h^{g,i,j})$ for the elements in a school assigned to zone *h* is computed as the product of k_h times their overall sampling weight, where k_h can take values of 0, 1, or 2 depending on whether the school is to be omitted, included with its usual weight, or have its weight doubled for the computation of the statistic of interest. In TIMSS 1999, the replicate weights were not permanent variables, but were created temporarily by the sampling variance estimation program as a useful computing device.

To create replicate weights, each sampled student was first assigned a vector of 75 weights, $w_h^{g,i,j}$, where *h* takes values from 1 to 75. The value of $w_0^{g,i,j}$ is the overall sampling weight, which is simply the product of the final school weight, the appropriate final classroom weight, and the appropriate final student weight, as described in chapters 5 and 6.

11.4 Combining Sampling and Imputation Variance

When reporting standard errors for proficiency estimates using plausible values, it is necessary to combine the sampling and

Exhibit 11.2 (continued) Summary Statistics and Standard Errors for Overall Mathematics Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
States					
Connecticut	2023	512	85	9.0	9.1
Idaho	1847	495	82	7.1	7.4
Illinois	4781	509	82	6.6	6.7
Indiana	2046	515	76	7.1	7.2
Maryland	3317	495	88	6.2	6.2
Massachusetts	2353	513	82	5.8	5.9
Michigan	2623	517	81	7.4	7.5
Missouri	1979	490	77	5.0	5.3
North Carolina	3089	495	84	6.7	7.0
Oregon	1889				

Exhibit 11.3 Summary Statistics and Standard Errors for Geometry Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	497	91	3.5	5.7
Belgium (Flemish)	5259	535	101	3.1	4.1
Bulgaria	3272	524	107	4.8	5.9
Canada	8770	507	89	1.5	

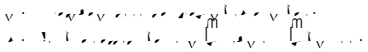

Exhibit 11.3 (continued) Summary Statistics and Standard Errors for Geometry Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
States					
Connecticut	2023	470	97	6.2	7.7
Idaho	1847	465	91	6.0	6.5
Illinois	4781	483	89	4.5	6.8
Indiana	2046	476	92	5.9	7.6
Maryland	3317	466	98	4.3	6.0
Massachusetts	2353	477	90	4.7	6.1
Michigan	2623	486	93	6.2	8.0
Missouri	1979	466	86	3.8	5.6
North Carolina	3089	475	89	5.1	5.6
Oregon	1889	486	92	5.0	6.8
Pennsylvania	3236	473	91	3.6	4.7
South Carolina	2011	476	97	6.5	7.8
Texas					

Exhibit 11.4 Summary Statistics and Standard Errors for Data Representation, Analysis and Probability Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]

Exhibit 11.4 (continued) Summary Statistics and Standard Errors for Data Representation, Analysis and Probability Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
States					
Connecticut	2023	516	105	8.2	9.9
Idaho	1847	501	102	6.1	7.2
Illinois	4781	510	98	6.3	7.1
Indiana	2046	518	95	5.5	6.3
Maryland	3317	504	99	5.5	6.4
Massachusetts	2353	521	102	6.0	6.3
Michigan	2623	517	101	6.7	6.8
Missouri	1979	500	96	4.4	5.0
North Carolina	3089	502	101	5.4	5.8

Exhibit 11.5 Summary Statistics and Standard Errors for Measurement Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	529	84	3.8	4.9
Belgium (Flemish)	5259	549	77	2.9	4.0
Bulgaria	3272	497	96	5.4	6.6
Canada	8770	521	80	2.0	2.4
Chile	5907	412	92	3.3	4.9
Chinese Taipei	5772	566	96	3.1	3.4
Cyprus	3116	471	93	2.2	4.0
Czech Republic	3453	535	83	3.3	5.0
England	2960	507	84	3.7	3.8
Finland	2920	521	74	2.6	4.7
Hong Kong, SAR	5179	567	79	4.0	5.8
Hungary	3183	538	84	2.6	3.5
Indonesia	5848	395	117	4.4	5.1
Iran, Islamic Rep.	5301	401	100	3.5	4.7
Israel	4195	457	97	3.9	5.1
Italy	3328	501	89	3.4	5.0
Japan	4745	558	75	1.7	2.4
Jordan	5052	438	106	3.2	4.4
Korea, Rep. of	6114	571	79	1.9	2.8
Latvia (LSS)	2873	505	89	3.1	3.5
Lithuania	2361	467	81	3.1	4.0
Macedonia, Rep. of	4023	451	101	3.4	5.2
Malaysia	5577	514	86	4.1	4.6
Moldova	3711	479	97	3.5	4.9
Morocco	5402	348	115	2.2	3.5
Netherlands	2962	538	73	5.4	5.8
New Zealand	3613	496	86	4.4	5.3
Philippines	6601	355	104	4.2	6.2
Romania	3425	491	99	4.4	4.9

Exhibit 11.6 Summary Statistics and Standard Errors for Algebra Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	520	81	4.1	5.1
Belgium (Flemish)	5259	540	86	3.2	4.6
Bulgaria	3272	512	88	4.8	5.1
Canada	8770	525	73	1.7	2.4
Chile	5907	399	96	3.9	4.3
Chinese Taipei	5772	586	114	4.3	4.4
Cyprus	3116	479	80	1.5	1.6
Czech Republic	3453	514	87	3.8	4.0
England	2960	498	77	3.3	4.9
Finland	2920	498	73	2.3	3.1
Hong Kong, SAR	5179	569	78	3.6	4.5
Hungary	3183	536	94	3.4	4.1
Indonesia	5848	424	104	3.9	5.7
Iran, Islamic Rep.	5301	434	88	2.8	4.9
Israel	4195	479	97	4.1	4.5
Italy	3328	481	84	3.3	3.6
Japan	4745	569	82	1.5	3.3
Jordan	5052	439	108	3.6	5.3
Korea, Rep. of	6114	585	90	1.9	2.7
Latvia (LSS)	2873	499	83	3.0	4.3
Lithuania	2361	487	74	3.4	3.7
Macedonia, Rep. of	4023	465	100	3.8	4.0
Malaysia	5577	505	81	3.8	4.8
Moldova	3711	477	91	3.2	3.7
Morocco	5402	353	111	2.2	4.7
Netherlands	2962	522	77	6.9	7.7
New Zealand	3613	497	81	4.3	4.7
Philippines	6601	345	119	5.2	5.8
Romania	3425	481	99	5.0	5.2
Russian Federation	4332	529	95	4.8	4.9
Singapore	4966	576	81	5.9	6.2
Slovak Republic	3497	525	76	3.6	4.6
Slovenia	3109	525	85	2.7	2.9
South Africa	8146	293	125	6.1	7.7
Thailand	5732	456	91	4.2	4.9
Tunisia	5051	455	74	1.9	2.7
Turkey	7841	432	98	4.3	4.6
United States	9072	506	90	3.4	4.1

Exhibit 11.6 (continued) Summary Statistics and Standard Errors for Algebra Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
States					
Connecticut	2023	513	84	8.0	8.2
Idaho	1847	500	83	6.5	7.3
Illinois	4781	513	84	5.5	

Exhibit 11.7 Summary Statistics and Standard Errors for Fractions and Number Sense Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	519	78	4.1	

Exhibit 11.7 (continued) Summary Statistics and Standard Errors for Fractions and Number Sense Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
States					

Exhibit 11.8 Summary Statistics and Standard Errors for Science Proficiency

Country	Sample Size	Mean of 5 Plausible Values	S.D.[a]	Error Due to Sampling	S.E.[b]
Australia	4032	540	87	4.3	4.4
Belgium (Flemish)	5259	535	69	2.6	3.1
Bulgaria	3272	518	93	5.3	5.4
Canada	8770	533	78	1.8	2.1
Chile	5907	420	88	3.7	3.7
Chinese Taipei	5772	569			

Exhibit 11.8 (continued) Summary Statistics and Standard Errors for Science Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
States					
Connecticut	2023	529	91	10.4	10.4
Idaho	1847	526	85	6.5	6.6
Illinois	4781	521	89	6.4	6.5
Indiana	2046	534	86	6.7	7.0
Maryland	3317	506	95	7.2	7.7
Massachusetts	2353	533	89	7.1	7.4
Michigan	2623	544	94	8.4	8.6
Missouri	1979	523	89	6.1	6.5
North Carolina	3089	508	90	6.2	6.5
Oregon	1889	536	91	5.7	6.1
Pennsylvania	3236	529	87	6.3	6.5
South Carolina	2011	511	95	6.7	6.7
Texas	1996	509	104	10.4	10.4
Districts and Consortia					
Academy School Dist. #20, CO	1233	559	77	1.7	2.1
Chicago Public Schools, IL	1132	449	90	9.4	9.5
Delaware Science Coalition, DE	1268	500	94	8.3	8.4
First in the World Consort., IL	750	565	78	4.0	5.3
Fremont/Lincoln/WestSide PS, NE	1093	511	91	4.8	5.8
Guilford County, NC	1018	534	93	7.0	7.1
Jersey City Public Schools, NJ	1004	440	96	9.6	9.8
Miami-Dade County PS, FL	1229	426	106	10.9	10.9
Michigan Invitational Group, MI	903	563	82	5.7	6.2
Montgomery County, MD	1155	531	92	3.5	4.3
Naperville Sch. Dist. #203, IL	1212	584	76	3.6	4.1
Project SMART Consortium, OH	1096	539	86	8.3	8.4
Rochester City Sch. Dist., NY	966	452	89	7.2	7.4
SW Math/Sci. Collaborative, PA	1538	543	85	7.3	7.4

Exhibit 11.9 Summary Statistics and Standard Errors for Life Science Proficiency

Country	Sample Size

Exhibit 11.9 (continued) Summary Statistics and Standard Errors for Life Science Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error

Exhibit 11.10 Summary Statistics and Standard Errors for Earth Science Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	519	96	3.9	6.1
Belgium (Flemish)	5259	533	92	2.8	3.5
Bulgaria	3272	520	115	5.4	5.7
Canada	8770	519	92	1.7	3.7
Chile	5907	435	93	3.0	7.0
Chinese Taipei	5772	538	89	2.0	3.0
Cyprus	3116	459	87	1.8	5.4
Czech Republic	3453	533	113	4.7	6.9
England	2960	525	88	3.6	3.9
Finland	2920	520	101	3.0	5.5
Hong Kong, SAR	5179	506	82	2.5	4.3
Hungary	3183	560	119	3.8	3.9
Indonesia	5848	431	99	3.7	6.4
Iran, Islamic Rep.	5301	459	96	2.8	5.2
Israel	4195	472	108	4.4	5.2
Italy	3328	502	103	3.6	5.9
Japan	4745	533	91	2.2	6.2
Jordan	5052	446	92	2.4	3.5
Korea, Rep. of	6114	532	98	2.1	2.7
Latvia (LSS)	2873	495	114	3.8	5.4
Lithuania	2361	476	91	3.2	4.4
Macedonia, Rep. of	4023	464	116	3.9	4.2
Malaysia	5577	491	90	3.4	4.2
Moldova	3711	466	117	3.0	4.2
Morocco	5402	363	112	2.0	3.3
Netherlands	2962	534	94	6.0	7.2
New Zealand	3613	504	90	3.7	5.8
Philippines	6601	390	103	4.9	5.0
Romania	3425	475	128	4.5	5.5
Russian Federation	4332	529	124	4.5	5.1
Singapore	4966	521	91	5.4	7.3
Slovak Republic	3497	537	99	4.0	4.3
Slovenia	3109	541	111	3.6	4.3
South Africa	8146	348	102	3.6	4.8
Thailand	5732	470	95	3.4	3.9
Tunisia	5051	442	89	1.6	2.7
Turkey	7841	435	90	3.6	4.6
United States	9072	504	98	3.4	4.2

Exhibit 11.10 Summary Statistics and Standard Errors for Earth Science Proficiency

Participants	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
States					
Connecticut	2023	508	93	6.1	6.5
Idaho	1847	513	96	5.5	6.6
Illinois	4781	505	95	5.1	7.2
Indiana	2046	515	92	5.8	6.3
Maryland	3317	495	94	4.7	6.1
Massachusetts	2353	516	95	6.6	7.6
Michigan	2623	526	101	7.3	7.9
Missouri	1979	[redacted]	98	4.4	5.8
[redacted]					
North Carolina	3089	500	92	5.2	7.0
Oregon	1889	528	97	4.8	6.7
Pennsylvania	[redacted]	515			

Exhibit 11.11 Summary Statistics and Standard Errors for Physics Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	531	90	3.6	6.3
Belgium (Flemish)	5259	530	82	2.0	3.5
Bulgaria	3272	505	109	4.8	5.8
Canada	8770	521	85	2.3	3.8
Chile	5907	428	93	2.6	5.6
Chinese Taipei	5772	552	96	3.0	3.9
Cyprus	3116	459	95	2.0	2.9
Czech Republic	3453	526	99	3.6	4.2
England	2960	528	86	3.7	4.5
Finland	2920	520	103	2.6	4.4
Hong Kong, SAR	5179	523	88	3.4	4.9
Hungary	3183	543	102	3.0	4.3
Indonesia	5848	452	94	3.2	5.5
Iran, Islamic Rep.	5301	445	105	4.0	5.7
Israel	4195	484	102	3.9	5.3
Italy	3328	480	93	3.5	4.1
Japan	4745	544	83	1.7	2.9
Jordan	5052	459	108	3.1	3.6
Korea, Rep. of	6114	544	92	2.3	5.1
Latvia (LSS)	2873	495	95	3.1	3.9
Lithuania	2361	510	85	3.5	4.3
Macedonia, Rep. of	4023	463	107	3.8	6.0
Malaysia	5577	494	89	3.2	4.1
Moldova	3711	457	112	3.9	5.5
Morocco	5402	352	120	2.2	4.2

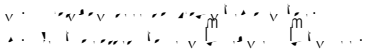

Exhibit 11.12 Summary Statistics and Standard Errors for Chemistry Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	520	101	4.2	5.0
Belgium (Flemish)	5259	508	92	2.4	3.3
Bulgaria	3272	527	115	4.5	5.7
Canada	8770	521	94	2.0	5.4
Chile	5907	435	97	3.2	5.2
Chinese Taipei	5772	563	105	3.0	4.3
Cyprus	3116	470	91	1.7	3.4
Czech Republic	3453	512	108	3.5	5.2
England	2960	524	95	3.8	5.5
Finland	2920	535	101	3.0	4.5
Hong Kong, SAR	5179	515	87	2.6	5.2
Hungary	3183	548	111	3.1	4.7
Indonesia	5848	425	88	3.5	3.9
Iran, Islamic Rep.	5301	487	92	2.4	4.1
Israel	4195	479	107	3.8	4.7
Italy	3328	493	94	3.2	4.8
Japan	4745	5304195			

Exhibit 11.13 Summary Statistics and Standard Errors for Scientific Inquiry and the Nature of Science Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	535	93	3.5	4.9
Belgium (Flemish)	5259	526	93	2.7	4.9
Bulgaria	3272	479	121	5.4	5.6
Canada	8770	532	86	1.2	5.1
Chile	5907	441	100	3.3	4.7
Chinese Taipei	5772	540	87	3.0	4.9
Cyprus	3116	467	104	2.1	4.6
Czech Republic	3453	522	108	4.8	5.7
England	2960	538	86	3.2	5.1
Finland	2920	528	101	2.6	4.0
Hong Kong, SAR	5179	531	82	2.3	2.8
Hungary	3183	526	103	2.9	5.9
Indonesia	5848	446	99	2.7	4.3
Iran, Islamic Rep.	5301	446	94	2.3	5.3
Israel	4195	476	112	3.8	8.3
Italy	3328	489	96	2.9	4.6
Japan	4745	543	77	1.8	2.8
Jordan	5052	440	109	2.6	5.5
Korea, Rep. of	6114	545	89	2.1	7.3
Latvia (LSS)	2873	495	104	3.2	4.7
Lithuania	2361	483	99	4.0	6.4
Macedonia, Rep. of	4023	464	117	3.2	3.6
Malaysia	5577	488	84	2.5	4.5
Moldova	3711	471	113	3.3	3.8
Morocco	5402	391	134	2.7	4.2
Netherlands	2962	534	98	5.1	6.5
New Zealand	3613	521	95	3.3	6.8
Philippines	6601	403	108	3.7	5.5
Romania	3425	456	118	3.4	5.5
Russian Federation	4332	491	109	3.3	4.9
Singapore	4966	550	85	4.2	5.9

Exhibit 11.14 Summary Statistics and Standard Errors for Environment and Resources Issues Proficiency

Country	Sample Size	Mean Proficiency[a]	Standard Deviation[a]	Jackknife Sampling Error	Overall Standard Error[b]
Australia	4032	530	104	3.9	6.3
Belgium (Flemish)	5259	513	98	2.3	3.5
Bulgaria	3272	483	126	5.5	6.4
Canada	8770	521	97	2.5	3.5
Chile	5907	449	97	2.6	4.8
Chinese Taipei	5772	567	101	2.4	4.0
Cyprus	3116	475	92	2.2	4.3
Czech Republic	3453	516	111	3.5	5.7
England	2960	518	108	4.1	5.8
Finland	2920	514	101	2.4	7.1
Hong Kong, SAR	5179	518	91	2.9	4.9
Hungary	3183	501	118	3.6	6.6
Indonesia	5848	489	84	2.2	4.8
Iran, Islamic Rep.	5301	470	86	2.6	5.5
Israel	4195	458	105	3.5	4.0
Italy	3328	491	93	2.5	5.4
Japan	4745	506	89	2.2	5.5
Jordan	5052	476	106	2.7	6.0
Korea, Rep. of	6114	523	96	1.5	4.5