

Appendix B

TEST-CURRICULUM MATCHING ANALYSIS

When comparing student achievement across countries, it is important that the comparisons be as “fair” as possible. TIMSS has worked towards this goal in a number of ways, including providing detailed procedures for standardizing the population definitions, sampling, test translations, test administration, scoring, and database formation. Developing the TIMSS tests involved the interaction of experts in the sciences with representatives of the participating countries and testing specialists.¹ The National Research Coordinators (NRCs) from each country formally approved the TIMSS test, thus accepting it as being sufficiently fair to compare their students’ science achievement with that of students from other countries.

Although the TIMSS test was developed to represent a set of agreed-upon science content areas, there are differences among the curricula of participating countries that result in various science topics being taught at different grades. To restrict test items not only to those topics in the curricula of all countries but also to those covered in the same sequence in all participating countries would severely limit test coverage and restrict the research questions about international differences that TIMSS is designed to address. The TIMSS tests, therefore, inevitably contain some items measuring topics unfamiliar to some students in some countries.

The Test-Curriculum Matching Analysis (TCMA) was developed and conducted to investigate the appropriateness of the TIMSS science test for third- and fourth-grade students in the participating countries, and to show how student performance for individual countries varied when based only on the test questions that were judged to be relevant to their own curriculum.²

To gather data about the extent to which the TIMSS tests were relevant to the curriculum of the participating countries, TIMSS asked the NRC of each country to report whether or not each item was in their country’s intended curriculum at each of the two grades being tested. The NRC was asked to choose a person or persons who were very familiar with the curricula at the grades being tested to make the determination. Since an item might be in the curriculum for some but not all students in a country, an item was determined appropriate if it was in the intended curriculum for more than 50% of the students. The NRCs had considerable flexibility in selecting items and may have considered items inappropriate for other reasons. All participating countries except Austria and Thailand returned the information for analysis.

¹ See Appendix A for more information on the test development.

² Because there also may be curriculum areas covered in some countries that are not covered by the TIMSS tests, the TCMA does not provide complete information about how well the TIMSS tests cover the curricula of the countries.

Tables B.1 and B.2 present the TCMA results for the fourth and third grades, respectively. The first row of each table indicates that at both grades the countries varied substantially in the number of items considered appropriate. At the fourth grade, half of the countries indicated that items representing two-thirds or more of the score points (70 out of a possible 105) were appropriate,³ with the percentage ranging from 100% in the United States to approximately 25% in Korea (25 score points) and Japan (29 score points). Fewer items were selected at the third grade, where about one-third of the countries selected at least half of the score points. All items were selected at the third grade as well as the fourth grade in the United States. At the third grade there were also several countries, including Ireland, Korea, and Japan, that retained less than 20% of the score points. That lower percentages of items were selected for the TCMA at the third grade is consistent with the instrument development process, which put more emphasis on the upper-grade curriculum. The low percentage of items considered appropriate for their curricula in several countries implies that science may not be emphasized at these grades by those countries.

Since most countries indicated that some items were not included in their intended curricula at the two grades tested, the question becomes whether the inclusion of these items had any effect on the international performance comparisons.⁴ The TCMA results offer a method for answering this question, providing evidence that the relative standings of countries generally do not vary much for the different sets of items selected from the TIMSS science test.

The first column in Tables B.1 and B.2 shows the overall average percent correct for each country (as discussed in Chapter 2 and reproduced here for convenience in making comparisons). The countries are presented in the order of their overall performance, from highest to lowest. To interpret these tables, reading across a row provides the average percent correct for the students in that country on the items selected by each country listed across the top of the table. For example, at the fourth grade Korea, where the average percent correct was 77% on its own set of items, had 79% for the items selected by Japan, 78% for those selected by the Netherlands, 74% for those selected by Australia, and so forth. The column for a country shows how each of the other countries performed on the subset of items selected for its own students. Using the set of items selected by Hong Kong as an example, on average, 81% of these were answered correctly by the Korean students, 76% by the Japanese students, 74% by the Dutch, and so forth. The shaded diagonal elements in each table show how each country performed on the subset of items that it selected based on its own curriculum. Thus, the Hong Kong students themselves averaged 72% correct responses on the items identified by Hong Kong for the analysis.

³ Of the 97 items in the test, some items were assigned more score points than others. In particular, some items had two parts, and some extended-response items were scored on a two-point scale. The total number of score points available for analysis was 105. The TCMA uses the score points in order to give the same weight to items that they received in the test scoring.

⁴ It should be noted that the performance levels presented in Tables B.1 and B.2 are based on the average percent correct as was done in Chapter 2, which is different from the average scale scores that were presented in Chapter 1. The cost and delay of scaling would have been prohibitive for the TCMA analyses.

Table B.1 Test-Curriculum Matching Analysis Results - Science - Upper Grade (Fourth Grade*)

Average Percent Correct Based on Subsets of Items Specially Identified by Each Country as Addressing Its Curriculum (See Table B.3 for corresponding standard errors)

Instructions: Read across the row to compare that country's performance based on the test items included by each of the countries across the top. Read down the column under a country name to compare the performance of the country down the items included by the country listed on the top. Read along the diagonal to compare performance for each different country based on its own decisions about the test items to include.

Country	Average Percent Correct on all Items																								
	(Number of Score Points Included)																								
	25	29	68	80	105	95	52	98	93	53	42	49	27	62	90	45	33	104	94	71	61	84	85	85	
Korea	77	79	78	74	74	75	78	74	75	76	81	78	83	77	76	81	76	74	73	76	78	75	75	75	75
Japan	68	79	74	71	70	70	73	70	71	71	76	72	80	73	72	76	71	70	70	72	72	72	72	72	70
Netherlands	65	70	72	68	67	67	69	68	67	68	74	69	77	72	68	75	69	67	66	70	71	70	70	68	68
Australia	66	67	70	68	66	67	70	67	66	68	73	67	78	70	68	75	67	66	66	69	70	68	67	67	67
United States	65	67	70	67	66	66	70	66	66	69	73	67	78	70	67	75	69	66	66	69	69	69	68	67	67
Czech Republic	65	68	70	67	65	66	68	66	65	66	72	68	79	70	68	71	67	65	66	68	69	68	67	66	66
Singapore	64	70	69	65	64	65	71	64	65	66	74	68	76	68	66	72	65	64	64	67	68	66	66	66	66
Slovenia	61	64	68	64	64	64	66	64	64	65	70	65	74	69	66	69	64	64	65	66	67	66	65	64	64
Canada	64	64	67	65	64	64	66	64	63	64	71	64	76	67	65	71	65	64	63	66	67	65	64	64	64
England	63	64	67	64	63	63	66	64	63	65	71	65	76	68	65	71	64	63	63	66	67	65	64	64	64
Hong Kong	59	68	66	64	62	63	66	63	62	63	72	65	76	68	64	67	61	62	62	65	65	64	64	64	63
Hungary	62	62	66	62	62	62	64	63	61	62	69	65	73	67	63	66	62	62	62	65	65	63	63	62	62
Ireland	60	61	64	62	61	61	64	61	61	62	69	62	74	65	62	69	64	61	61	63	66	63	62	62	62
Norway	58	61	65	61	60	61	63	61	60	61	68	62	71	67	62	68	59	61	61	63	64	63	61	61	61
New Zealand	59	60	64	61	60	61	64	61	60	61	67	61	72	64	62	68	61	60	60	63	65	62	61	61	61
Scotland	59	61	64	61	60	60	64	61	60	62	69	62	73	65	62	69	61	60	60	63	65	62	61	61	61
Israel	56	58	61	57	57	57	61	57	57	59	65	59	69	61	58	65	58	57	57	59	60	60	58	57	57
Latvia (LSS)	55	56	60	57	56	56	59	57	56	56	63	59	66	61	58	62	56	56	57	58	59	58	58	57	57
Iceland	55	55	59	56	55	56	59	56	55	56	61	58	65	61	57	64	55	55	56	57	57	57	57	56	56
Greece	54	54	58	54	54	54	58	54	54	54	62	56	65	55	55	62	55	54	54	57	58	56	55	55	55
Cyprus	52	52	55	51	51	50	55	51	50	49	59	53	61	56	52	58	53	50	51	53	54	52	52	52	52
Portugal	50	50	54	51	50	50	53	51	49	50	57	52	59	55	52	57	52	50	50	52	53	53	51	50	50
Iran, Islamic Rep.	40	39	44	40	40	40	45	40	39	38	48	42	48	44	41	44	37	40	40	42	41	41	41	41	41
Kuwait	39	42	43	39	39	40	43	39	38	37	47	41	48	43	41	45	40	39	39	42	41	40	40	40	41
International Average	60	(0.6)	61	64	60	60	63	60	60	60	67	62	71	64	61	67	61	60	60	62	63	61	61	60	60

*Fourth grade in most countries; see Table 2 for more information about the grades tested in each country.
 **Of the 97 items in the science test, some items had two parts and some extended-response items were scored on a two-point scale, resulting in 105 total score points.
 () Standard errors for the average percent of correct responses on all items appear in parentheses. Standard errors for scores based on subsets of items are provided in Table B.3.
 Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 Countries shown in italics did not satisfy one or more guidelines for sample participation rates, age/grade specifications, or classroom sampling procedures (see Figure A.3 for details).
 Because population coverage falls below 65% Latvia is annotated LSS for Latvian Speaking Schools only.
 SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

The international averages of each country's selected items presented across the last row of the tables show that the selection of items for the participating countries varied somewhat in average difficulty, ranging from 59% to 71% at the fourth grade and from 47% to 61% at third grade. Despite these differences, the overall picture provided by Tables B.1 and B.2 reveals that different item selections do not make a major difference in how well countries perform relative to each other. The items selected by some countries were more difficult than those selected by others. The relative performance of countries on the various item selections did vary somewhat, but generally not in a statistically significant manner.⁵

Comparing the diagonal element for a country with the overall average percentage correct shows the difference between performance on this subset of items and performance on the test as a whole. In general, there were small increases in each country's performance on its own subset of items. To illustrate, the average percent correct for fourth-grade students in Korea was 74%. The diagonal element shows that Korean students had about the same average percent correct (77%) on the smaller set of items selected as relevant to the curriculum in Korea as they did overall. In the fourth grade, most countries had a difference of less than 5 percentage points between the two performance measures, with the largest difference of 13% for Ireland (74% compared to 61%). Performance differences between the entire TIMSS test and the subset of items selected for the TCMA were, in general, somewhat larger for third-grade students; several countries had an average performance that was 10 percentage points or more higher on the subsets of items selected for their own students – Japan, the Netherlands, the Czech Republic, Hong Kong, Scotland, Ireland, Norway, and Greece.

It is clear that the selection of items does not have a major effect on the general relationship among countries. Countries that had substantially higher or lower performance on the overall test in comparison to each other also had higher or lower relative performance on the different sets of items selected for the TCMA. For example, at the fourth grade, Korea had the highest average percent correct on the test as a whole and on all of the item selections, with Japan, the Netherlands, and Australia among the four highest-performing countries in almost all cases. Although there are some changes in the ordering of countries based on the items selected for the TCMA, most of these differences are within the boundaries of sampling error.

As the most extreme example, consider the 27 score points selected by Ireland for the fourth grade. The Irish students did substantially better on these items than on the test as a whole, with 74% correct responses to these items, on average, compared to 61% average correct on the items on the test as a whole. However, all other countries also did better on these particular items, with an international average of 71% for the items selected by Ireland compared with 60% on the test as a whole. Insofar as countries

⁵ Small differences in performance in these tables are not statistically significant. The standard errors for the estimated average percent correct statistics can be found in Tables B.3 and B.4. We can say with 95% confidence that the value for the entire population will fall between the sample estimate plus or minus two standard errors.

rejected items that would be difficult for their own students, these items tended to be difficult for students in other countries as well. The analysis shows that omitting such items tends to improve the results for that country, but also tends to improve the results for all other countries, so that the relative standing of countries is largely unaffected.

Table B.3 Standard Errors for the Test-Curriculum Matching Analysis Results - Science - Upper Grade (Fourth Grade*)

See Table B.1 for the Test-Curriculum Matching Analysis Results

Instructions: Read across the row for the standard error for the score based on the test items included by each of the countries across the top. Read down the column under a country name for the standard error for the score of the country down the left on the items included by the country listed on the top. Read along the diagonal for the standard error for the score for each different country based on its own decisions about the test items to include.

Country	Average Percent Correct on all Items		(Number of Score Points Included)																								
	105	25	29	68	80	105	95	52	98	93	53	42	49	27	62	90	45	33	104	94	71	61	84	85	85		
Korea	74 (0.4)	0.4	0.5	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	
Japan	70 (0.3)	0.4	0.4	0.3	0.4	0.3	0.4	0.4	0.4	0.3	0.4	0.4	0.4	0.4	0.3	0.3	0.4	0.3	0.3	0.3	0.4	0.3	0.3	0.3	0.3	0.4	
Netherlands	67 (0.5)	0.7	0.7	0.5	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.5	
Australia	66 (0.5)	0.6	0.6	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.6	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	
United States	66 (0.5)	0.6	0.6	0.6	0.5	0.5	0.5	0.6	0.5	0.5	0.6	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	
Czech Republic	65 (0.5)	0.7	0.7	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.6	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	
Singapore	64 (0.8)	0.9	0.9	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.9	
Slovenia	64 (0.7)	0.8	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.7	0.7	0.7	0.7	0.6	0.7	0.7	0.6	0.7	0.7	0.7	0.7	
Canada	64 (0.6)	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	
England	63 (0.6)	0.8	0.8	0.6	0.6	0.6	0.6	0.7	0.6	0.6	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	
Hong Kong	62 (0.7)	0.8	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.6	0.7	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	
Hungary	62 (0.6)	0.8	0.8	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	
Ireland	61 (0.6)	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	
Norway	60 (0.6)	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	
New Zealand	60 (0.9)	1.1	1.0	0.9	0.9	0.9	0.9	1.0	0.9	0.9	1.0	0.9	1.0	1.0	0.9	0.9	0.9	1.1	1.0	0.9	0.9	0.9	1.0	0.9	0.9	0.9	
Scotland	60 (0.8)	0.9	1.0	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	
Israel	57 (0.8)	1.0	1.1	0.8	0.8	0.8	0.8	1.0	0.8	0.9	1.0	0.9	0.9	0.9	0.9	0.8	0.8	1.0	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	
Latvia (LSS)	56 (0.8)	0.9	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.8	
Iceland	55 (0.7)	0.8	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.9	0.8	0.8	0.8	0.7	0.7	0.7	0.8	0.8	0.7	0.7	0.8	0.8	0.7	0.7	0.7	
Greece	54 (0.8)	1.0	1.0	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.8	0.8	0.9	0.9	0.8	0.9	0.9	0.9	0.8	0.8	0.9	0.9	0.8	0.9	0.8	
Cyprus	51 (0.5)	0.6	0.6	0.6	0.5	0.5	0.5	0.6	0.5	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.5	0.5	0.6	0.6	0.6	0.5	0.5	
Portugal	50 (0.7)	0.9	1.0	0.8	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.7	0.8	0.8	0.9	0.8	0.8	0.7	0.8	0.8	0.8	0.8	0.7	
Iran, Islamic Rep.	40 (0.7)	0.7	0.9	0.8	0.7	0.7	0.7	0.7	0.8	0.7	0.8	0.8	0.8	0.7	0.9	0.8	0.8	0.8	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	
Kuwait	39 (0.5)	0.5	0.6	0.6	0.6	0.5	0.5	0.5	0.6	0.5	0.5	0.6	0.6	0.6	0.8	0.6	0.6	0.7	0.6	0.5	0.5	0.6	0.6	0.6	0.5	0.6	
International Average	60 (0.6)	0.8	0.8	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.6	0.7	0.7	0.6	0.6	0.6	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	

*Fourth grade in most countries; see Table 2 for more information about the grades tested in each country.
 **Of the 97 items in the science test, some items had two parts and some extended-response items were scored on a two-point scale, resulting in 105 total score points.
 () Standard errors for the average percent of correct responses on all items appear in parentheses. The matrix contains standard errors corresponding to the average percent of correct responses based on TCMA subsets of items, as displayed in Table B.1. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 Countries shown in italics did not satisfy one or more guidelines for sample participation rates, age/grade specifications, or classroom sampling procedures (see Figure A.3 for details).
 Because population coverage falls below 65% Latvia is annotated LSS for Latvian Speaking Schools only.
 SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Table B.4 Standard Errors for the Test-Curriculum Matching Analysis Results - Science - Lower Grade (Third Grade*)

See Table B.3 for the Test-Curriculum Matching Analysis Results

Instructions: Read across the row for the standard error for the score based on the test items included by each of the countries across the top.
 Read down the column under a country name for the standard error for the score of the country down the left on the items included by the country listed on the top.
 Read along the diagonal for the standard error for the score for each different country based on its own decisions about the test items to include.

Country	Average Percent Correct on all Items		(Number of Score Points Included)																					
	105	11	17	49	105	34	40	32	64	32	92	34	33	14	68	41	74	49	30	93	84	44	39	
Korea	67 (0.5)	0.8	0.6	0.5	0.5	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Japan	61 (0.3)	0.5	0.5	0.3	0.4	0.3	0.3	0.3	0.4	0.3	0.4	0.4	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4
Australia	57 (0.8)	0.9	1.0	0.8	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.9	0.8	0.8	0.7	0.7	0.7	0.8	0.8	0.7	0.7	0.7	0.7
United States	56 (0.6)	1.1	0.8	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.7	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6
Netherlands	56 (0.7)	1.0	0.8	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.7	0.7	0.7	0.7	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Czech Republic	55 (0.6)	0.9	0.7	0.6	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
England	55 (0.7)	0.9	1.0	0.7	0.7	0.7	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Canada	53 (0.5)	0.8	0.7	0.6	0.5	0.5	0.5	0.6	0.5	0.7	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6
Singapore	53 (0.9)	0.8	0.9	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	0.9	0.9	0.9	0.9	1.0	0.9	0.9	0.9	0.9	0.9	0.9
Slovenia	53 (0.5)	1.0	0.7	0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.5	0.7	0.7	0.5	0.6	0.5	0.5	0.5	0.7	0.5	0.5	0.5	0.7
Hong Kong	53 (0.6)	0.8	0.7	0.6	0.6	0.8	0.7	0.6	0.7	0.7	0.6	0.5	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.6
Scotland	51 (0.7)	1.1	0.9	0.8	0.7	0.9	0.8	0.8	0.8	0.9	0.8	0.7	0.8	0.8	0.9	0.8	0.8	0.8	0.8	0.8	0.7	0.7	0.7	0.8
Ireland	51 (0.7)	1.0	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.9	0.7	0.7	0.6	0.7	0.8	0.6	0.7	0.7	0.7	0.8
New Zealand	51 (0.9)	1.4	1.1	1.0	0.9	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.1	0.9	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0
Hungary	50 (0.8)	1.1	0.9	0.9	0.8	1.0	0.8	0.9	0.9	0.9	0.8	0.8	0.8	1.0	0.8	0.8	0.8	0.8	0.9	0.8	0.8	0.8	0.8	0.8
Latvia (LSS)	48 (0.9)	1.2	1.0	0.9	0.9	0.9	0.8	0.9	0.9	0.9	0.9	0.9	0.9	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
Norway	46 (0.7)	1.0	0.9	0.8	0.7	0.9	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.9	0.7	0.8	0.7	0.8	0.8	0.7	0.7	0.7	0.7	0.8
Greece	44 (0.7)	1.3	1.2	0.8	0.7	1.0	0.9	1.1	0.8	0.9	0.8	0.9	0.8	1.0	0.8	0.9	0.8	0.8	0.8	0.9	0.9	0.7	0.7	0.9
Iceland	42 (0.6)	1.1	1.1	0.7	0.6	0.8	0.8	0.8	0.6	0.8	0.6	0.7	0.9	0.8	0.7	0.8	0.7	0.7	0.8	0.6	0.7	0.8	0.7	0.8
Portugal	41 (0.8)	1.1	1.0	0.8	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.8	0.8	1.0	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Cyprus	39 (0.5)	1.0	0.8	0.5	0.5	0.7	0.6	0.6	0.6	0.6	0.5	0.6	0.7	0.8	0.6	0.7	0.5	0.6	0.6	0.5	0.5	0.6	0.6	0.6
Iran, Islamic Rep.	30 (0.7)	1.0	0.8	0.7	0.7	0.8	0.9	0.8	0.8	0.8	0.7	0.9	0.8	0.9	0.7	0.8	0.7	0.8	0.9	0.7	0.7	0.7	0.7	0.8
International Average	51 (0.7)	1.0	0.9	0.7	0.7	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7

*Third grade in most countries; see Table 2 for more information about the grades tested in each country.

**Of the 87 items in the science test, some items had two parts and some extended-response items were scored on a two-point scale, resulting in 105 total score points.

() Standard errors for the average percent of correct responses on all items appear in parentheses. The matrix contains standard errors corresponding to the average percent of correct responses based on TCMA subsets of items, as displayed in Table B.2. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Countries shown in italics did not satisfy one or more guidelines for sample participation rates, age/grade specifications, or classroom sampling procedures (see Figure A.3 for details). Because population coverage falls below 65% Latvia is annotated LSS for Latvian Speaking Schools only.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.