


# CHAPTER 1

## International Student Achievement in Science



1



Chapter 1 summarizes eighth-grade achievement on the TIMSS 1999 science assessment for each of the participating countries, and shows trends in student performance for those countries that also participated in TIMSS 1995 at the eighth grade. Comparisons of country performance against international benchmarks, as well as gender differences in performance, also are provided.

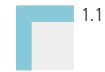




## How Do Countries Differ in Science Achievement?

Exhibit 1.1 presents the distribution of student achievement for the 38 countries that participated in TIMSS 1999.<sup>1</sup> Countries are shown in decreasing order of average (mean) scale score, together with an indication of whether the country average is significantly higher or lower than the international average. The international average of 488 was obtained by averaging across the mean scores for each of the 38 participating countries. The results reveal substantial differences in science achievement between the high- and low-performing countries, from an average of 569 for Chinese Taipei to 243 for South Africa. Nineteen countries had average science achievement that was significantly above the international average, including two countries that are participating in TIMSS for the first time – Chinese Taipei and Finland.<sup>2</sup> Thirteen countries had average achievement below the international average, including nine countries new to TIMSS—Moldova, the Republic of Macedonia, Jordan, Indonesia, Turkey, Tunisia, Chile, the Philippines, and Morocco.

The broad range of achievement both within and across countries is illustrated in Exhibit 1.1 by a graphical representation of the distribution of student performance within each country. Achievement for each country is shown for the 25th and 75th percentiles as well as for the 5th and 95th percentiles.<sup>3</sup> Each percentile point indicates the percentages of students performing below and above that point on the scale. For example, 25 percent of the eighth-grade students in each country performed below the 25th percentile for that country, and 75 percent performed above the 25th percentile. The range between the 25th and 75th percentiles represents performance by the middle half of the students. In most countries, the range of performance for the middle group was between 100 and 150 scale-score points. In contrast, performance at the 5th and 95th percentiles represents the extremes in both lower and higher achievement. The range of performance between these two score points, which include 90 percent of the population, is between 250 and 300 points in most countries. The dark boxes at the midpoints of the distributions show the 95 percent confidence intervals around the average achievement in each country.<sup>4</sup>



<sup>1</sup> TIMSS used item response theory (IRT) methods to summarize the achievement results on a scale with a mean of 500 and a standard deviation of 100. Given the matrix-sampling approach, scaling averages students' responses in a way that accounts for differences in the difficulty of different subsets of items. It allows students' performance to be summarized on a common metric even though individual students responded to different items in the science test. For more detailed information, see the "IRT Scaling and Data Analysis" section of Appendix A.


<sup>2</sup> The significance tests in Exhibits 1.1 and 1.2 are based on a Bonferroni procedure for multiple comparisons that holds to 5 percent the probability of erroneously stating the mean of one country to be different from that of another country.

<sup>3</sup> Tables of the percentile values and standard deviations for all countries are presented in Appendix D.

<sup>4</sup> See the "IRT Scaling and Data Analysis" section of Appendix A for more details about calculating standard errors and confidence intervals for the TIMSS statistics.

As well as showing the wide spread of student achievement within each country, the percentiles also provide a perspective on the size of the differences among countries. Even though performance generally differed very little between one country and the next higher- or lower-performing country, the range in performance across the 38 countries was very large. For example, average performance in Chinese Taipei exceeded performance at the 95th percentile in the lower-performing countries such as the Philippines, Morocco, and South Africa. This means that only the most proficient students in the lower-performing countries approached the level of achievement of students of average proficiency in Chinese Taipei.

To aid in interpretation, Exhibit 1.1 also includes the years of formal schooling and average age of the students in each country. Equivalence of chronological age does not necessarily mean that students have received the same number of years of formal schooling or studied the same curriculum. Most notably, students in Finland, Morocco, the Philippines, and parts of the Russian Federation had fewer years of formal schooling than their counterparts in other countries, while those in the Czech Republic, England, Moldova, and parts of Australia and New Zealand had more years of schooling. The average age of students ranged from 13.8 years in Cyprus and Finland to 15.5 years in South Africa.

1.2  Exhibit 1.2 compares overall mean achievement among individual countries. This figure shows whether or not the differences in average achievement between pairs of countries are statistically significant. Selecting a country of interest and reading across the table, a triangle pointing up indicates significantly higher performance than the comparison country listed across the top; a circle indicates no significant difference in performance; and a triangle pointing down indicates significantly lower performance.

The data in Exhibit 1.2 reinforce the point that, when ordered by average achievement, adjacent countries usually did not significantly differ from each other, although the differences in achievement between the high-performing and low-performing countries were very large. Because of this wide range in performance, the pattern for a number of countries was one of having lower mean achievement than some countries, about the same mean achievement as other countries, and higher mean achievement than a third group of countries.

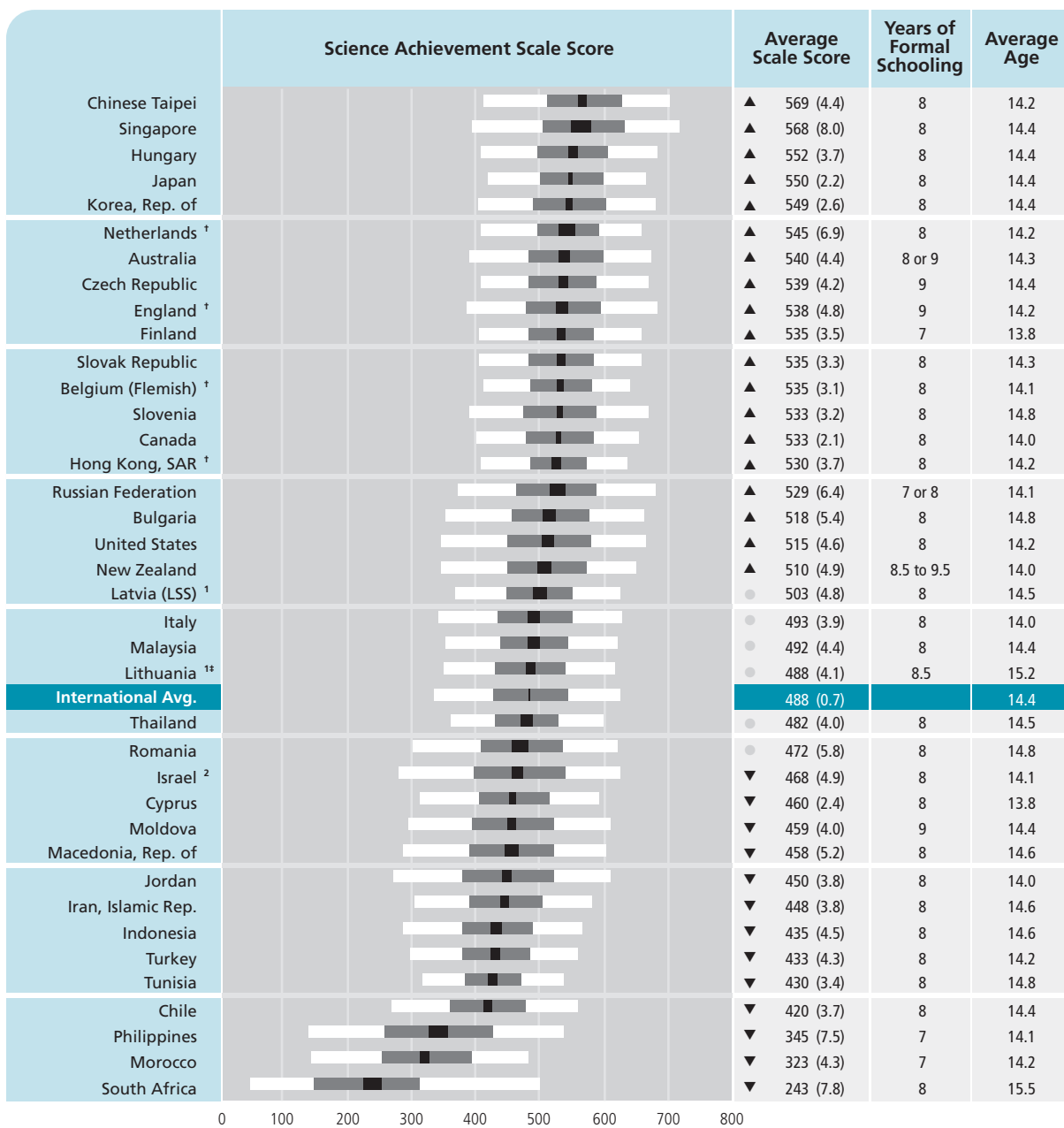


Chinese Taipei and Singapore had the highest average performance, closely followed by Hungary, Japan, and the Republic of Korea. Other countries that performed very well included the Netherlands,<sup>5</sup> Australia, the Czech Republic, and England. The latter group of countries had similar achievement levels. The difference in performance from one country to the next was often negligible. For example, Finland, the Slovak Republic, Belgium (Flemish), Slovenia, Canada, Hong Kong SAR, the Russian Federation, and Bulgaria outperformed about half of the participating countries. In turn, the United States, while performing less well than Chinese Taipei, Singapore, Hungary, Japan, Korea, the Netherlands, Australia, the Czech Republic, England, Finland, the Slovak Republic, Belgium (Flemish), Slovenia, and Canada, performed at about the same level as Hong Kong, the Russian Federation, Bulgaria, New Zealand, and Latvia (LSS),<sup>6</sup> and higher than all other countries. In contrast, the Philippines, Morocco, and South Africa performed less well than the other countries, with South Africa having significantly lower achievement than the other two.

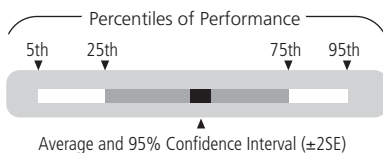
<sup>5</sup> Average achievement for the Netherlands was lower than that for Chinese Taipei, Singapore, Hungary, Japan, and Korea, but the difference was not statistically significant because the Netherlands had a larger than usual standard error.

<sup>6</sup> Because coverage of its eighth-grade population falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

# Exhibit 1.1 Distribution of Science Achievement



SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.



- ▲ Country average significantly higher than international average
- No statistically significant difference between country average and international average
- ▼ Country average significantly lower than international average

Significance tests adjusted for multiple comparisons

† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.8).

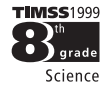
<sup>1</sup> National Desired Population does not cover all of International Desired Population (see Exhibit A.5). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

<sup>2</sup> National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.5).

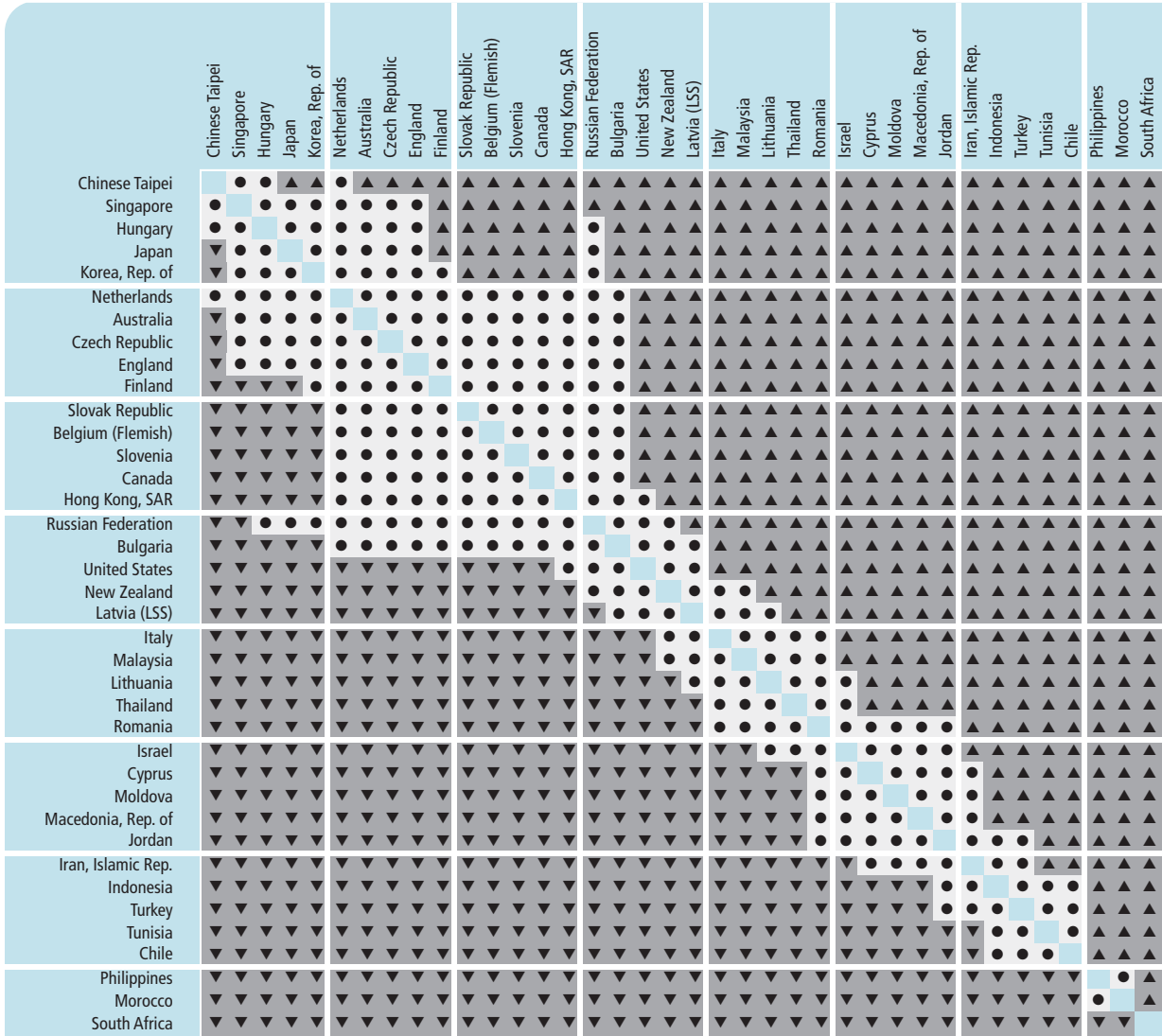
‡ Lithuania tested the same cohort of students as other countries, but later in 1999, at the beginning of the next school year.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

# Exhibit 1.2 Multiple Comparisons of Average Science Achievement



**Instructions:** Read across the row for a country to compare performance with the countries listed along the top of the chart. The symbols indicate whether the average achievement of the country in the row is significantly lower than that of the comparison country, significantly higher than that of the comparison country, or if there is no statistically significant difference between the average achievement of the two countries.



SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

- ▲ Average achievement significantly higher than comparison country
- No statistically significant difference from comparison country
- ▼ Average achievement significantly lower than comparison country

Significance tests adjusted for multiple comparisons

## How Has Science Achievement Changed Since 1995?

1.3

Twenty-six countries took part in the TIMSS eighth-grade assessments in both 1995 and 1999. For these countries, Exhibit 1.3 shows the results in 1995 and 1999 and the differences in average achievement between the two years.<sup>7</sup> Average science achievement across these 26 countries increased from a scale score of 518 in 1995 to 521 in 1999, although the gain was not statistically significant.

In some countries, average science achievement increased considerably between 1995 and 1999. The greatest increase was in Latvia (LSS), with an increase of 27 scale-score points. Lithuania showed a similar increase, although this should be interpreted with caution, since Lithuania conducted the assessment six months later than other participants, when the students were beginning ninth grade rather than finishing eighth grade. Other countries with significant increases in achievement were Canada and Hungary. Hong Kong and Australia also had large increases, although the somewhat larger estimates of measurement error for these countries meant that the differences were not statistically significant.

Several countries showed a small decrease in average achievement from 1995 to 1999, but only in the case of Bulgaria was it statistically significant. Israel, South Africa, and Thailand are shown in a separate panel in Exhibit 1.3 because they used unapproved sampling procedures at the classroom level in 1995. Israel and Thailand showed large decreases since 1995, which could indicate an upward bias in the 1995 results due to their sampling problems in the original TIMSS rather than actual decreases.

1.4

TIMSS in 1995 assessed both fourth- and eighth-grade students. This allowed participants to compare their performance relative to each other at the fourth and eighth grades, and gave a cross-sectional perspective on how relative performance changed between grades.<sup>8</sup> For example, as shown in Exhibit 1.4, the United States, Australia, and Canada in 1995 performed significantly above the international average at the fourth grade, but just similar to it at the eighth grade. These countries place considerable emphasis on science education in the early grades, so it could be that this apparent relative decline from fourth to eighth grade is partly because other countries begin to emphasize science after the fourth grade. That Singapore, Slovenia, and Hungary, the countries with just average fourth-grade performance but above average eighth-grade performance in 1995, each begin to emphasize science instruction prior to the eighth grade lends support to this interpretation.

<sup>7</sup> TIMSS used IRT methods to place the eighth-grade results from 1995 and 1999 on the same scale. See Appendix A for more detailed information.

<sup>8</sup> The science achievement scale for fourth grade is not comparable to that for eighth grade, and so results for fourth grade and eighth grade may be compared only in relative terms, for example with reference to the international average for countries that participated in 1995 at both the fourth and eighth grades.





It has also been argued, at least in the United States, that recent reforms in education had their greatest impact in the earlier grades, and that a second TIMSS assessment could show better results for eighth grade in 1999 than in 1995. Of the three countries with a relative decline from fourth to eighth grade in 1995, only the United States showed the same relative decline from fourth grade in 1995 to eighth grade in 1999. Hopes in that country that the benefits of educational reform would be evident in the 1999 eighth-grade results have not been realized. New Zealand also showed a relative decline at the eighth grade, from about the international average in 1995 to below it in 1999. In Canada and Australia, in contrast, the relative position has improved since 1995, with both countries above the international average at eighth grade in 1999.

## Exhibit 1.3 Trends in Science Achievement



SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

### Countries with Unapproved Sampling Procedures at the Classroom Level in 1995

Israel	509 (6.3)	484 (5.7)	-25 (8.3)
South Africa	263 (11.1)	243 (7.8)	-20 (13.7)
Thailand	510 (4.7)	482 (4.0)	-28 (6.2)

■ Difference statistically significant

□ Difference not statistically significant

Significance tests adjusted for multiple comparisons

<sup>§</sup> International average is for countries that participated and met sampling guidelines in both 1995 and 1999.

Trend notes: Because coverage fell below 65% in 1995 and 1999, Latvia is annotated LSS for Latvian-Speaking Schools only. Lithuania tested later in 1999 than in 1995, at the beginning of the next school year. In 1995, Italy and Israel were unable to cover their International Desired Population; 1999 data are based on their comparable populations.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

# Exhibit 1.4 Science Achievement for TIMSS 1999 Countries That Participated in 1995 at Both the Fourth and Eighth Grades in Relation to the Average Across These Countries



## 1995

Eighth Grade Difference From Average Across Countries <sup>§</sup>	
Singapore	60 (5.2)
Czech Republic	34 (4.4)
Japan	34 (1.9)
Korea, Rep. of	25 (2.2)
Netherlands	21 (5.8)
Slovenia	20 (2.8)
Hungary	16 (3.1)
England	13 (3.5)
Australia	6 (3.9)
Canada	-7 (2.7)
United States	-8 (5.3)
New Zealand	-10 (4.6)
Hong Kong	-11 (5.5)
Italy	-23 (3.4)
Latvia (LSS)	-44 (3.3)
Iran, Islamic Rep.	-58 (3.5)
Cyprus	-69 (2.2)
<b>Avg. Across Countries<sup>§</sup></b>	<b>521 (1.0)</b>

Fourth Grade Difference From Average Across Countries <sup>§</sup>	
Korea, Rep. of	62 (2.2)
Japan	39 (1.9)
United States	28 (3.2)
Australia	28 (3.5)
Czech Republic	18 (3.0)
Netherlands	17 (3.1)
England	14 (3.1)
Canada	12 (3.0)
Italy	10 (4.4)
Singapore	10 (4.6)
Slovenia	8 (3.9)
Hong Kong	-6 (3.3)
Hungary	-6 (3.3)
New Zealand	-9 (5.1)
Latvia (LSS)	-27 (4.7)
Cyprus	-64 (3.1)
Iran, Islamic Rep.	-134 (4.4)
<b>Avg. Across Countries<sup>§</sup></b>	<b>514 (0.9)</b>

## 1999

Eighth Grade Difference From Average Across Countries <sup>§</sup>	
Singapore	44 (7.6)
Hungary	28 (3.6)
Japan	25 (2.4)
Korea, Rep. of	24 (2.6)
Netherlands	21 (6.5)
Australia	16 (4.3)
Czech Republic	15 (4.1)
England	14 (4.5)
Canada	9 (2.1)
Slovenia	9 (3.3)
Hong Kong, SAR	5 (3.5)
United States	-9 (4.5)
New Zealand	-15 (4.8)
Latvia (LSS)	-21 (4.9)
Italy	-26 (4.5)
Cyprus	-64 (2.3)
Iran, Islamic Rep.	-76 (3.7)
<b>Avg. Across Countries<sup>§</sup></b>	<b>524 (1.1)</b>

Country average significantly higher than average across countries

Country average not significantly different from average across countries

Country average significantly lower than average across countries

Significance tests adjusted for multiple comparisons

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

<sup>§</sup> Average across the subset of TIMSS 1999 countries that participated and met sampling guidelines in 1995 at both the fourth and eighth grades.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

## How Do Countries Compare with International Benchmarks of Science Achievement?

1.5



The TIMSS science achievement scale summarizes student performance on test items designed to measure a wide range of student knowledge and proficiency. In order to provide meaningful descriptions of what performance on the scale could mean in terms of the science that students know and can do, TIMSS identified four points on the scale for use as international benchmarks, and conducted an ambitious scale-anchoring exercise to describe performance at these benchmarks. Exhibit 1.5 shows the four international benchmarks of science achievement and briefly describes what students scoring at these benchmarks typically know and can do. More detailed descriptions appear in Chapter 2, together with example test items illustrating performance at each benchmark.

The Top 10% Benchmark is defined at the 90th percentile on the TIMSS science scale, taking into account the performance of all students in all countries participating in 1999. This point on the scale, which corresponds to a scale score of 616, is the point above which the top 10 percent of the students in the TIMSS 1999 assessment scored. Students performing at this level demonstrated a grasp of some complex and abstract science concepts in earth science, life science, physics, and chemistry, and showed an understanding of the fundamentals of scientific investigation.

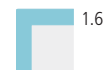
The Upper Quarter Benchmark is the 75th percentile on the science scale. This point, corresponding to a scale score of 558, is the point above which the top 25 percent of students scored. Students scoring at this benchmark typically demonstrated conceptual understanding of some science cycles, systems, and principles.

The Median Benchmark, with a score of 488, corresponds to the 50th percentile, or median. This is the point above which the top half of the students scored on the TIMSS 1999 assessment. Students performing at this level typically were able to recognize and communicate basic scientific information across a range of topics.

The Lower Quarter Benchmark is the 25th percentile and corresponds to a scale score of 410. This score point is reached by the top 75 percent of students, and may be used as a benchmark of performance for lower-achieving students. Students scoring at this level typically could recognize some basic facts from the earth, life, and physical sciences presented in non-technical language.

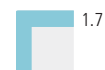


Exhibit 1.6 displays the percentage of students in each participating country that reached each international benchmark, in decreasing order by percentage reaching the Top 10% Benchmark. If student achievement in science were distributed in the same way in every country, then each country would be expected to have about 10 percent of its students reaching the Top 10% Benchmark, 25 percent the Upper Quarter Benchmark, 50 percent the Median Benchmark, and 75 percent the Lower Quarter Benchmark. Although countries such as Latvia (LSS), Italy, Israel, Malaysia, and Lithuania came fairly close, no country followed this pattern exactly. Instead, the high-performing countries generally had greater percentages of students reaching each benchmark, and the low-performing countries had lesser percentages. Among the high performers, for example, Singapore and Chinese Taipei had more than one-quarter of their students reaching the Top 10% Benchmark, more than half reaching the Upper Quarter Benchmark, four-fifths or more reaching the Median Benchmark, and almost all (94 to 95 percent) reaching the Lower Quarter Benchmark. In contrast, low-performing countries such as South Africa and Morocco had almost no students reaching the Top 10% Benchmark, only one or two percent reaching the Upper Quarter Benchmark, five or six percent reaching the Median Benchmark, and no more than 20 percent reaching the Lower Quarter Benchmark.



Although Exhibit 1.6 is organized to draw particular attention to the percentage of high-achieving students in each country, it conveys information about the distribution of middle and low performers also. For example, several countries, including Belgium (Flemish), Hong Kong, Malaysia, Lithuania, and Thailand, had greater percentages of students reaching the Median and Lower Quarter Benchmarks than might be expected from their percentages of high-performing students.

Exhibits 1.7 through 1.10 provide more information on the change in student performance from 1995 to 1999 by showing the percentages reaching each international benchmark in each of the years for the 26 countries that participated in both assessments.<sup>9</sup> Changes from 1995 to 1999 in the percentages of students reaching the Top 10% Benchmark are shown in Exhibit 1.7. Although on average across the 26 countries the percentage of students reaching this benchmark rose from 13 percent in 1995 to 14 percent in 1999, this increase was not statistically significant. Only in Hungary was there a significant increase, from 14 percent in 1995 to 22 percent in 1999. Bulgaria was the only country with a significant decrease, from 24 percent of students reaching the benchmark in 1995 to 14 percent in 1999.



<sup>9</sup> For Exhibits 1.7 through 1.10 the benchmarks were those computed from the 1999 data.



1.8

Countries generally had more success increasing the percentage of students reaching the Upper Quarter Benchmark (see Exhibit 1.8).

Although on average internationally there was little difference between the percentages reaching this benchmark in 1995 (34 percent) and in 1999 (35 percent), there was a significant increase in Canada, Hungary, Latvia (LSS), and Lithuania, and no country had a significant decrease.



1.9

Exhibit 1.9 shows the change from 1995 to 1999 in the percentage of students reaching the Median Benchmark. Like the two previous benchmarks, the average percentage of students reaching the benchmark increased slightly, in this instance from 65 percent in 1995 to 66 percent in 1999, but the increase was not statistically significant. At this benchmark also, Canada, Latvia (LSS), and Lithuania were the countries with the greatest increases. A somewhat similar situation was obtained for the Lower Quarter Benchmark (see Exhibit 1.10), as the international average percentage of students reaching it increased slightly, from 88 percent to 89 percent. Countries with significant increases at this benchmark were Canada, Hong Kong, and Latvia (LSS), and those with significant decreases were the Islamic Republic of Iran, Singapore, and Slovenia.



1.10

Taken together, the results from Exhibits 1.7 through 1.10 confirm that the modest increase in average student performance that was evident from Exhibit 1.3 was largely due to improved performance among a few countries. While in Hungary the increase was greatest among the more proficient students – those scoring above the Upper Quarter and Top 10% Benchmarks – in Canada, Latvia (LSS) and Lithuania the increase occurred more generally across the range of student proficiency.



**Exhibits 1.5–1.10 Overleaf**

### • Top 10% Benchmark

*Students demonstrate a grasp of some complex and abstract science concepts.* They can apply understanding of earth's formation and cycles and of the complexity of living organisms. They show understanding of the principles of energy efficiency, phase change, thermal expansion, light properties, gravitational force, basic structure of matter, and chemical versus physical changes. They demonstrate detailed knowledge of environmental and resource issues. They understand some fundamentals of scientific investigation and can apply basic physical principles to solve some quantitative problems. They can provide written explanations and use diagrams to communicate scientific knowledge.

90th Percentile: 616

### • Upper Quarter Benchmark

*Students demonstrate conceptual understanding of some science cycles, systems, and principles.* They have some understanding of the earth's processes, biological systems and populations, chemical reactions, and composition of matter. They solve physics problems related to light, speed, heat, and temperature and demonstrate basic knowledge of major environmental concerns. They demonstrate some scientific inquiry skills. They can combine information to draw conclusions; interpret information in diagrams, graphs and tables to solve problems; and provide short explanations conveying scientific knowledge in the life sciences.

75th Percentile: 558

### • Median Benchmark

*Students can recognize and communicate basic scientific knowledge across a range of topics.* They recognize some characteristics of the solar system, ecosystems, animals and plants, energy sources, force and motion, light reflection and radiation, sound, electrical circuits, and human impact on the environment. They can apply and briefly communicate practical knowledge, extract tabular information, extrapolate from data presented in a simple linear graph, and interpret representational diagrams.

50th Percentile: 488

### • Lower Quarter Benchmark

*Students recognize some basic facts from the earth, life, and physical sciences presented using non-technical language.* They can identify some of the earth's physical features, have some knowledge of the human body, and demonstrate familiarity with everyday physical phenomena. They can interpret and use information presented in simple diagrams.

25th Percentile: 410

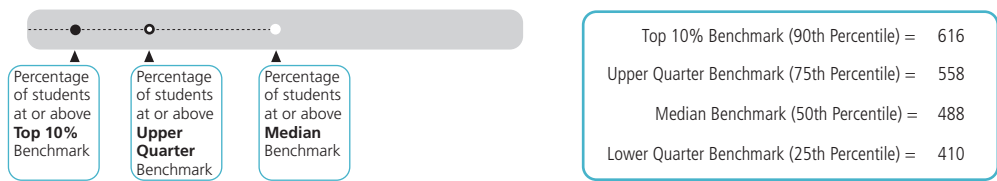
The international benchmarks are based on the combined data from the countries participating in 1999.



# Exhibit 1.6 Percentages of Students Reaching TIMSS 1999 International Benchmarks of Science Achievement



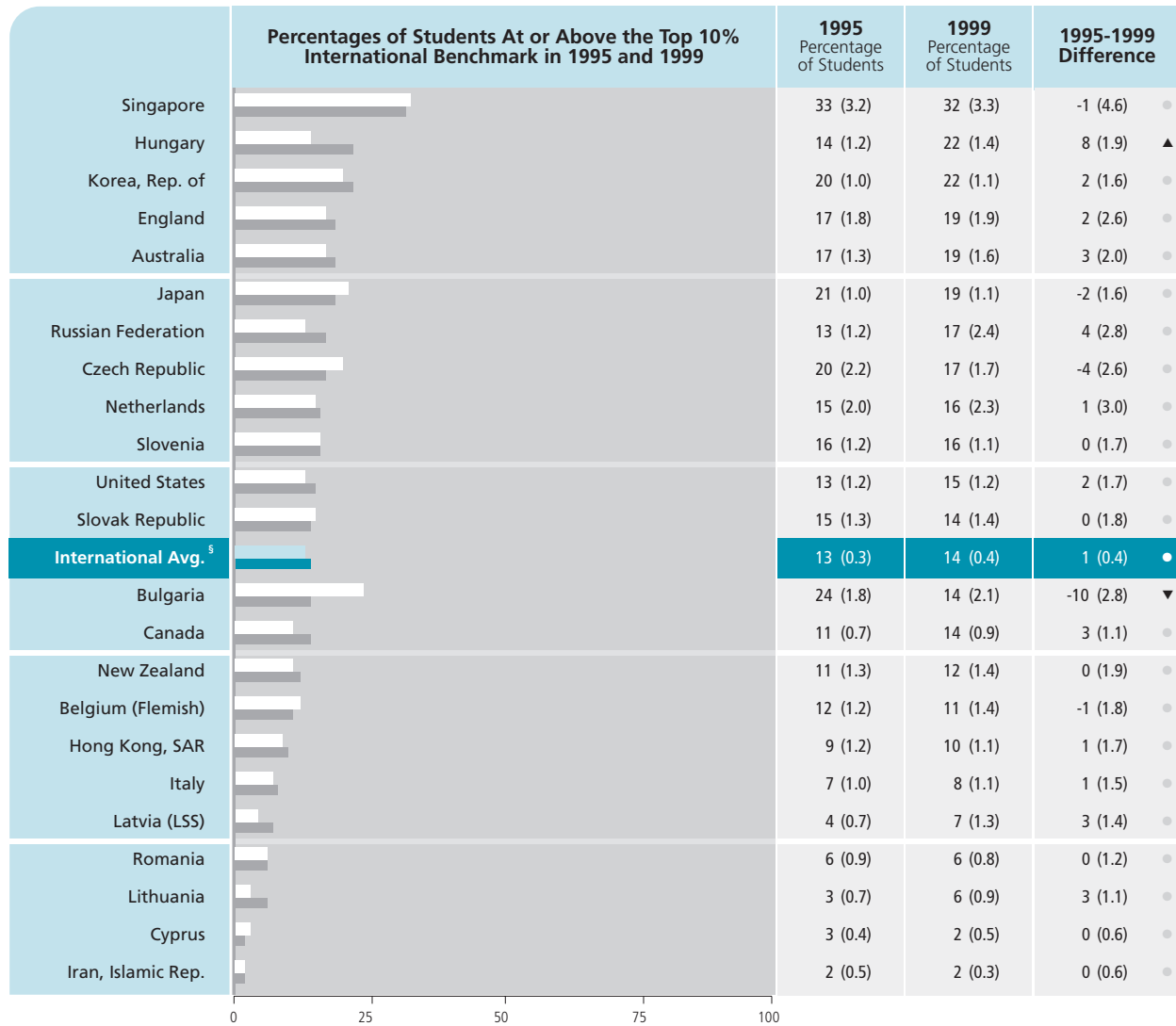
SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.



† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.8 for details).  
 1 National Desired Population does not cover all of International Desired Population (see Exhibit A.5). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.  
 2 National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.5).

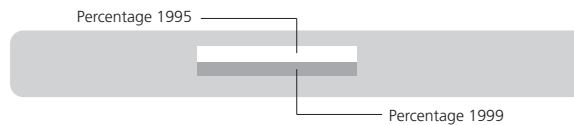
\* Lithuania tested the same cohort of students as other countries, but later in 1999, at the beginning of the next school year.  
 ( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

## Exhibit 1.7 Trends in Percentages of Students Reaching the TIMSS 1999 Top 10% International Benchmark of Science Achievement



### Countries with Unapproved Sampling Procedures at the Classroom Level in 1995

Israel		12 (1.8)	8 (0.8)	-4 (2.0)	●
South Africa		1 (0.5)	0 (0.2)	0 (0.6)	●
Thailand		6 (1.3)	3 (0.7)	-2 (1.5)	●



1999 significantly higher than 1995 ▲

No significant difference between 1995 and 1999 ●

1999 significantly lower than 1995 ▼

Significance tests adjusted for multiple comparisons

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

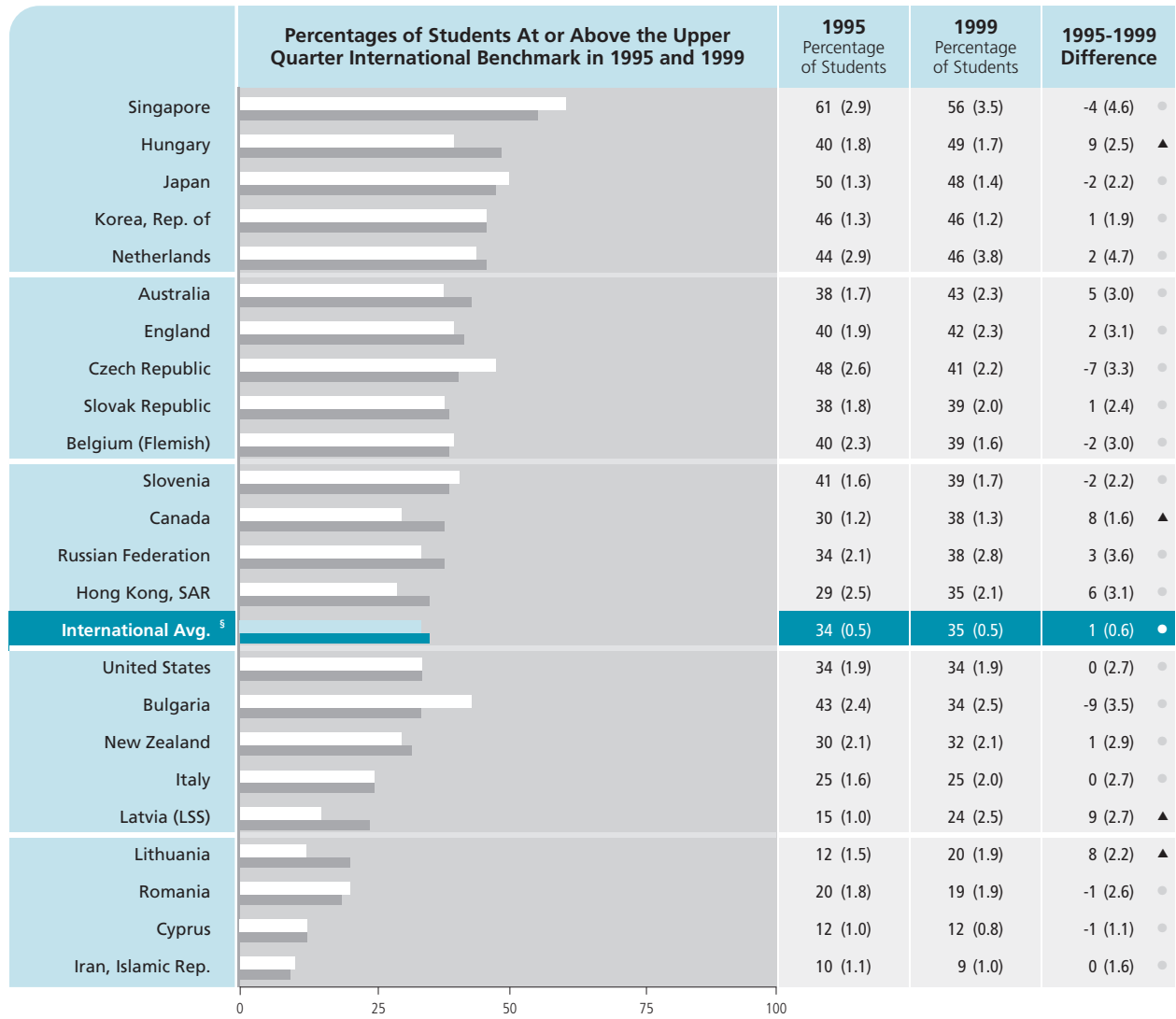
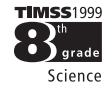
<sup>§</sup> International average is for countries that participated and met sampling guidelines in both 1995 and 1999.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Trend notes: Because coverage fell below 65% in 1995 and 1999, Latvia is annotated LSS for Latvian-Speaking Schools only. Lithuania tested later in 1999 than in 1995, at the beginning of the next school year. In 1995, Italy and Israel were unable to cover their International Desired Population; 1999 data are based on their comparable populations.

**Exhibit 1.8**

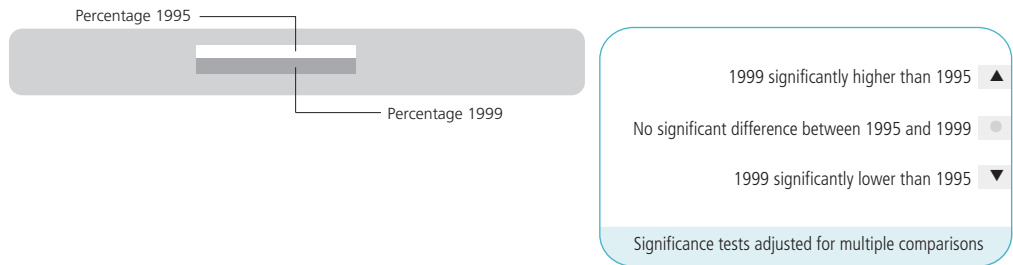
**Trends in Percentages of Students Reaching the TIMSS 1999 Upper Quarter International Benchmark of Science Achievement**



SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

**Countries with Unapproved Sampling Procedures at the Classroom Level in 1995**

Israel	31 (2.6)	24 (1.5)	-7 (3.0)	●
South Africa	4 (1.3)	2 (0.6)	-2 (1.4)	●
Thailand	23 (2.6)	15 (2.0)	-8 (3.2)	●

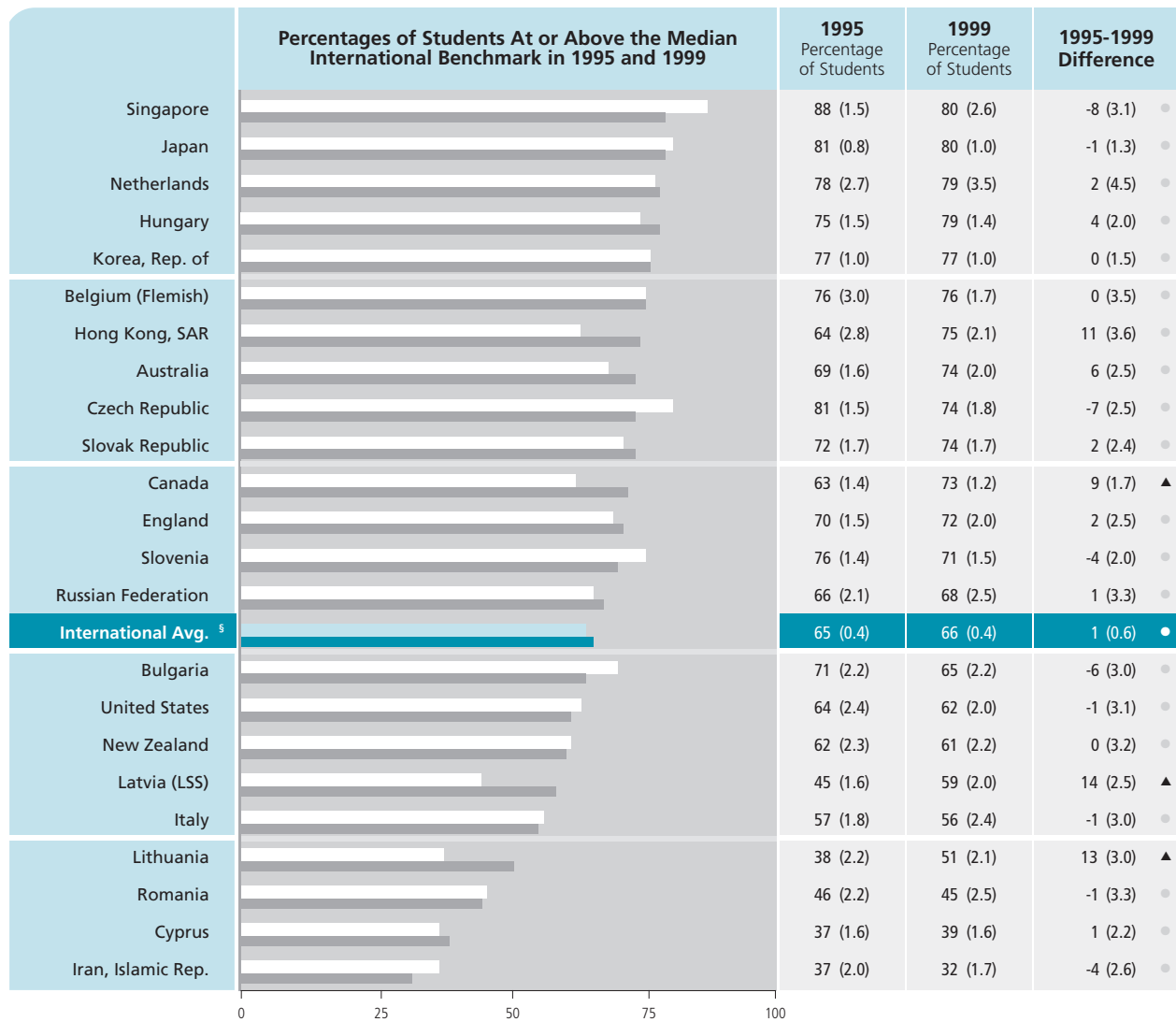


§ International average is for countries that participated and met sampling guidelines in both 1995 and 1999.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

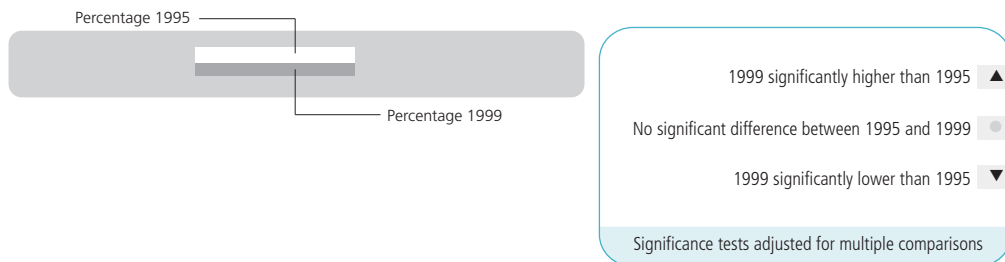
Trend notes: Because coverage fell below 65% in 1995 and 1999, Latvia is annotated LSS for Latvian-Speaking Schools only. Lithuania tested later in 1999 than in 1995, at the beginning of the next school year. In 1995, Italy and Israel were unable to cover their International Desired Population; 1999 data are based on their comparable populations.

## Exhibit 1.9 Trends in Percentages of Students Reaching the TIMSS 1999 Median International Benchmark of Science Achievement



### Countries with Unapproved Sampling Procedures at the Classroom Level in 1995

Israel		61 (2.8)	51 (2.3)	-10 (3.5) ●
South Africa		8 (2.2)	6 (1.4)	-2 (2.6) ●
Thailand		64 (2.5)	47 (2.5)	-16 (3.6) ▼



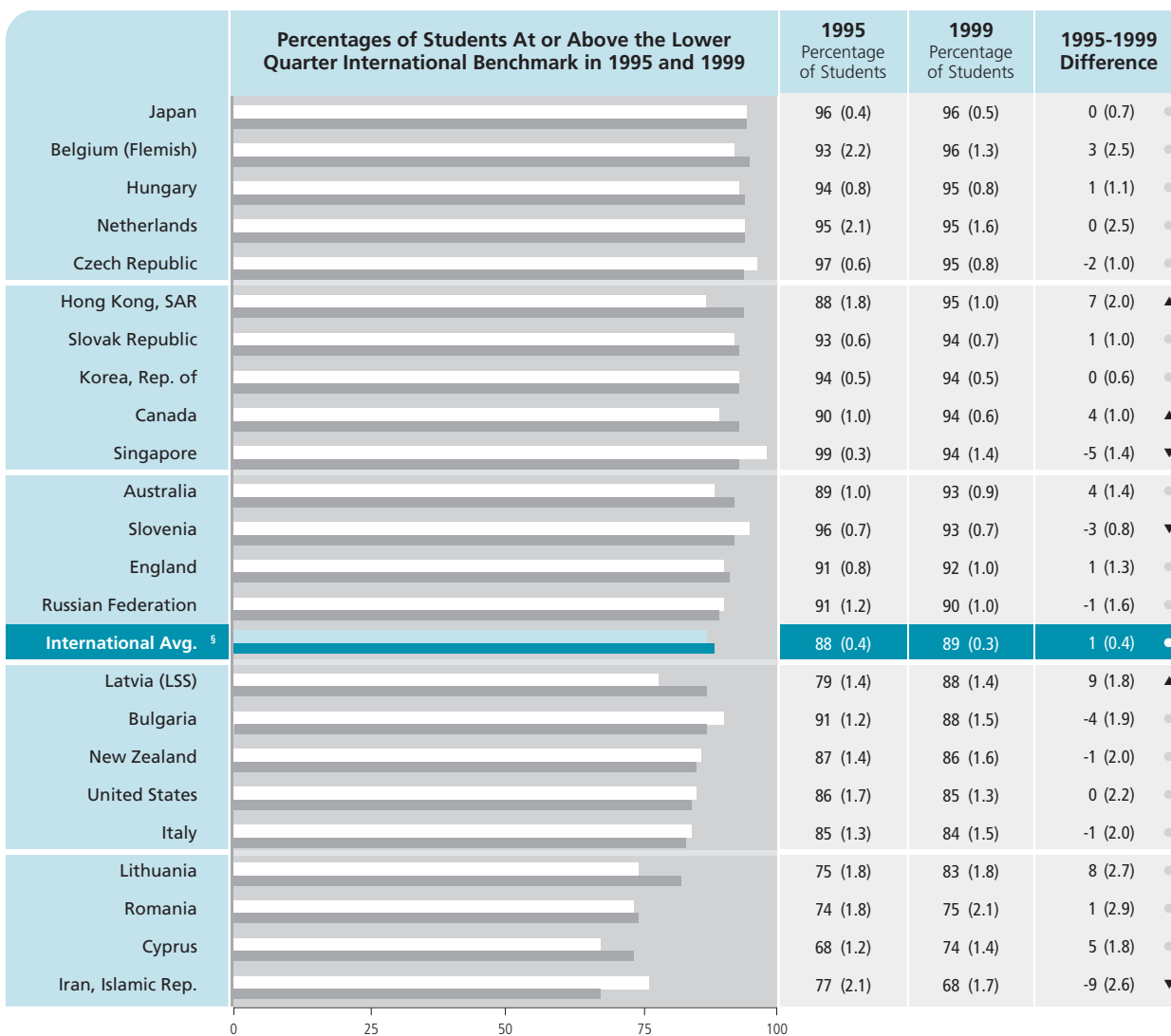
§ International average is for countries that participated and met sampling guidelines in both 1995 and 1999.

Trend notes: Because coverage fell below 65% in 1995 and 1999, Latvia is annotated LSS for Latvian-Speaking Schools only. Lithuania tested later in 1999 than in 1995, at the beginning of the next school year. In 1995, Italy and Israel were unable to cover their International Desired Population; 1999 data are based on their comparable populations.

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

**Exhibit 1.10**

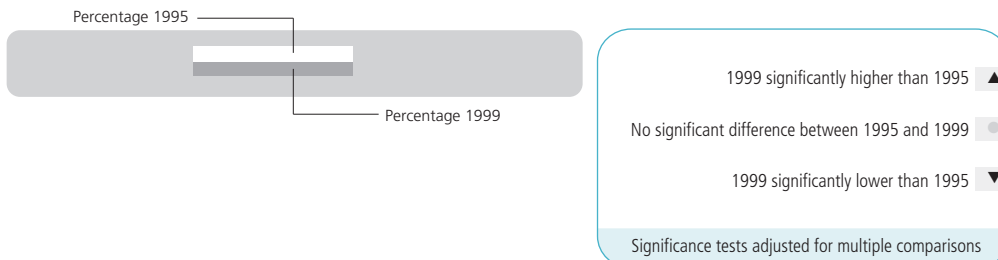
**Trends in Percentages of Students Reaching the TIMSS 1999 Lower Quarter International Benchmark of Science Achievement**



SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

**Countries with Unapproved Sampling Procedures at the Classroom Level in 1995**

Israel		86 (1.8)	78 (2.3)	-8 (2.9) ●
South Africa		15 (3.0)	13 (2.0)	-2 (3.6) ●
Thailand		93 (0.9)	84 (1.3)	-9 (1.5) ▼



<sup>§</sup> International average is for countries that participated and met sampling guidelines in both 1995 and 1999.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Trend notes: Because coverage fell below 65% in 1995 and 1999, Latvia is annotated LSS for Latvian-Speaking Schools only. Lithuania tested later in 1999 than in 1995, at the beginning of the next school year. In 1995, Italy and Israel were unable to cover their International Desired Population; 1999 data are based on their comparable populations.

## What Are the Gender Differences in Science Achievement?

1.11



Exhibits 1.11 through 1.14 show gender differences in eighth-grade science achievement in 1999, and also changes since 1995. Exhibit 1.11 presents average achievement separately for girls and boys for each of the TIMSS 1999 countries, as well as the difference between the means. Countries are shown in increasing order of this gender difference. The gender difference for each country is shown by a bar, indicating the amount of the difference, whether the direction of the difference favored girls or boys, and whether the difference is statistically significant (indicated by a darkened bar). On average across all countries, there was a significant difference of 15 scale-score points favoring boys, although the situation varied considerably from country to country. In many countries the gender difference was negligible. Among those with the smallest difference were Macedonia, Turkey, and Thailand. However, differences large enough to be statistically significant were found in 16 of the 38 countries. The countries with the largest differences were Iran, England, and the Czech Republic, where the mean for boys exceeded the mean for girls by more than 30 scale-score points.

1.12



Exhibit 1.12 provides information on gender differences in science achievement among students with high performance compared to those in the middle of the achievement distribution. For each country, score levels were computed for the highest-scoring 25 percent of students, called the upper quarter level, and for the top-scoring 50 percent of students, called the median level. The percentages of girls and boys in each country reaching each of the two levels were computed. For equitable performance, 25 percent each of girls and boys should have reached the upper quarter level, and 50 percent each the median level.

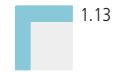
As may be seen from Exhibit 1.12, the gender difference in science at the country level is more apparent among high-performing students, although internationally it was about the same at both the upper quarter and median levels. On average across countries, 29 percent of boys reached the upper quarter level, compared with 21 percent of girls, a statistically significant difference of eight percentage points. Similarly, the international average percentage of boys reaching the median level was 54 percent and of girls 46 percent, also a significant difference of eight percentage points. Perhaps more important, however, Exhibit 1.12 shows that in 21 countries the percentage of boys reaching the upper



quarter level was significantly greater than the percentage of girls, whereas this was the case in 13 countries at the median level. In no country did the percentage of girls reaching either level significantly exceed the percentage of boys.

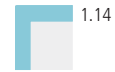
TIMSS in 1995 showed a pervasive difference in science achievement favoring boys, far more evident than in mathematics.<sup>10</sup> These findings were consistent with the results from the second IEA science study conducted in 1983-84, which for 14-year-olds found standard score differences favoring boys in all 23 of the participating countries.<sup>11</sup> In the light of this evidence of longstanding gender differences in science achievement, Exhibits 1.13 and 1.14 examine trends in gender differences from 1995 to 1999 for countries that participated in both assessments.

Achievement differences from 1995 to 1999 are presented separately for girls and for boys in Exhibit 1.13. Average science achievement across countries for girls increased significantly, from 506 to 512, over this period. Achievement for boys did not increase significantly, although the 1999 international average of 531 for boys remains well above the average for girls. Countries where science achievement for girls increased significantly from 1995 to 1999 were Latvia (LSS), Hong Kong, Lithuania, and Canada. Achievement for boys increased significantly in Lithuania, Canada, and Cyprus.



1.13

Taking the study of trends in gender differences one step further, Exhibit 1.14 presents the difference in average science achievement between girls and boys in 1995 and in 1999, and shows whether the difference has changed. On average across countries in 1995, achievement for boys significantly exceeded that for girls by 21 scale-score points. In 1999, the difference fell to 18 points, a statistically significant reduction in the gender gap. Average science achievement was greater for boys in 18 countries in 1995, but in just 13 countries in 1999. The countries that contributed to the overall decrease in gender difference were Hong Kong, Slovenia, and Israel, the only countries that had a significant reduction in the gender difference between 1995 and 1999.

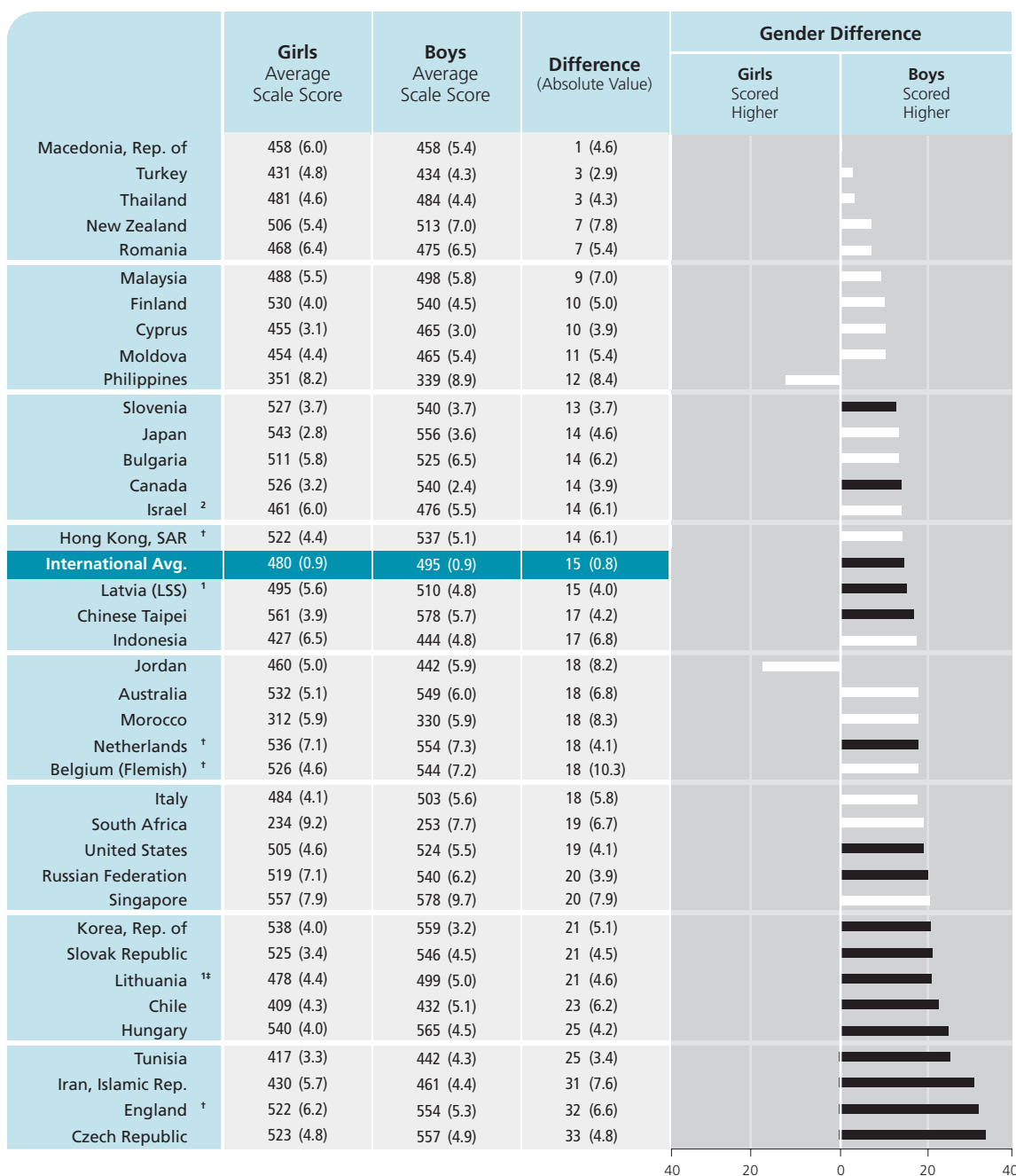


1.14

<sup>10</sup> Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L., and Smith, T.A. (1996), *Mathematics Achievement in the Middle School Years: The IEA's Third International Mathematics and Science Study (TIMSS)*, Chestnut Hill, MA: Boston College.

<sup>11</sup> Postlethwaite, T.N. and Wiley, D.E. (1992), *The IEA Study of Science II: Science Achievement in Twenty-Three Countries*, New York, NY: Pergamon Press.

# Exhibit 1.11 Average Science Achievement by Gender



SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

Gender difference statistically significant  
 Gender difference not statistically significant  
 Significance tests adjusted for multiple comparisons

<sup>†</sup> Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.8 for details).

<sup>1</sup> National Desired Population does not cover all of International Desired Population (see Exhibit A.5). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

<sup>2</sup> National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.5).

<sup>††</sup> Lithuania tested the same cohort of students as other countries, but later in 1999, at the beginning of the next school year.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.



## Percentages of Girls and Boys Reaching Each Country's Own Upper Quarter and Median Levels of Science Achievement

	Upper Quarter		Median	
	Percent of Girls	Percent of Boys	Percent of Girls	Percent of Boys
Australia	20 (1.8)	30 (2.4) ▲	46 (2.9)	55 (3.0)
Belgium (Flemish) †	20 (1.7)	30 (2.5)	44 (2.6)	56 (3.5)
Bulgaria	21 (2.6)	29 (2.9)	47 (2.8)	53 (3.2)
Canada	21 (1.5)	29 (1.3) ▲	46 (1.7)	54 (1.7)
Chile	19 (1.6)	31 (2.3) ▲	45 (2.2)	55 (2.3)
Chinese Taipei	20 (1.6)	30 (2.1) ▲	46 (2.0)	54 (2.4) ▲
Cyprus	21 (1.4)	29 (1.3) ▲	47 (1.4)	53 (1.4)
Czech Republic	18 (1.8)	32 (2.4) ▲	42 (2.5)	58 (2.5) ▲
England †	19 (2.5)	31 (2.4) ▲	43 (3.0)	56 (2.3) ▲
Finland	22 (2.0)	28 (2.1)	47 (2.3)	53 (2.3)
Hong Kong, SAR †	20 (2.5)	30 (2.4)	45 (2.8)	55 (2.6)
Hungary	19 (1.6)	31 (1.9) ▲	44 (2.0)	56 (2.1) ▲
Indonesia	22 (1.7)	28 (2.0)	46 (2.6)	55 (3.1)
Iran, Islamic Rep.	18 (2.4)	30 (2.1) ▲	40 (2.9)	57 (2.1) ▲
Israel <sup>2</sup>	21 (1.5)	29 (1.8) ▲	48 (2.4)	53 (2.3)
Italy	21 (1.8)	30 (2.0) ▲	45 (2.1)	55 (2.1) ▲
Japan	21 (1.3)	29 (1.4) ▲	46 (2.0)	54 (1.7)
Jordan	26 (1.8)	24 (1.6)	53 (1.9)	47 (2.3)
Korea, Rep. of	21 (1.4)	29 (1.4) ▲	44 (1.7)	55 (1.5) ▲
Latvia (LSS) <sup>1</sup>	21 (1.7)	29 (2.0) ▲	46 (2.3)	54 (2.2)
Lithuania <sup>‡</sup>	20 (2.0)	30 (2.4) ▲	46 (2.4)	54 (2.4) ▲
Macedonia, Rep. of	25 (1.9)	25 (1.8)	51 (2.6)	49 (2.2)
Malaysia	23 (2.2)	27 (3.0)	48 (2.6)	52 (3.0)
Moldova	23 (1.6)	28 (1.8)	47 (2.4)	53 (2.4)
Morocco	22 (1.8)	27 (1.3)	45 (2.3)	53 (1.9)
Netherlands †	21 (2.5)	30 (3.4) ▲	45 (4.1)	56 (4.0)
New Zealand	23 (2.1)	27 (2.9)	48 (2.7)	52 (3.3)
Philippines	26 (2.7)	24 (2.4)	52 (2.9)	47 (2.6)
Romania	24 (2.2)	26 (2.4)	49 (2.6)	51 (2.6)
Russian Federation	21 (2.7)	29 (2.8) ▲	45 (3.1)	55 (2.6) ▲
Singapore	20 (2.9)	30 (4.0)	45 (3.9)	55 (4.2)
Slovak Republic	19 (1.7)	31 (2.1) ▲	44 (2.0)	56 (2.2) ▲
Slovenia	21 (1.3)	29 (1.4) ▲	47 (1.7)	53 (2.0) ▲
South Africa	23 (2.7)	27 (2.5)	47 (2.5)	53 (2.1)
Thailand	24 (2.5)	26 (2.3)	49 (2.7)	51 (2.4)
Tunisia	19 (1.4)	31 (1.7) ▲	42 (1.6)	58 (1.6) ▲
Turkey	23 (1.9)	26 (1.6)	48 (2.1)	51 (2.0)
United States	20 (1.6)	30 (2.0) ▲	46 (2.1)	54 (2.2) ▲
<b>International Avg.</b>	21 (0.3)	29 (0.4) ▲	46 (0.4)	54 (0.4) ▲

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

▲ Significantly higher than other gender

Significance tests adjusted for multiple comparisons

† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.8 for details).

<sup>1</sup> National Desired Population does not cover all of International Desired Population (see Exhibit A.5). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

<sup>2</sup> National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.5).

<sup>‡</sup> Lithuania tested the same cohort of students as other countries, but later in 1999, at the beginning of the next school year.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

## Exhibit 1.13 Trends in Average Science Achievement by Gender

Girls				Boys			
	1995 Average Scale Score	1999 Average Scale Score	1995-1999 Difference		1995 Average Scale Score	1999 Average Scale Score	1995-1999 Difference
Latvia (LSS)	464 (3.8)	495 (5.6)	32 (6.5) ▲	Lithuania	477 (4.5)	499 (5.0)	22 (6.6) ▲
Hong Kong, SAR	492 (6.5)	522 (4.4)	30 (7.8) ▲	Latvia (LSS)	490 (4.3)	510 (4.8)	21 (7.0) ●
Lithuania	452 (4.3)	478 (4.4)	26 (6.1) ▲	Canada	521 (3.4)	540 (2.4)	19 (4.1) ▲
Canada	508 (3.2)	526 (3.2)	18 (4.4) ▲	Hungary	549 (3.5)	565 (4.5)	17 (5.6) ●
Hungary	525 (3.7)	540 (4.0)	15 (6.0) ●	Australia	533 (5.5)	549 (6.0)	16 (8.2) ●
Australia	520 (4.4)	532 (5.1)	12 (6.6) ●	Cyprus	451 (2.4)	465 (3.0)	14 (3.9) ▲
New Zealand	497 (5.6)	506 (5.4)	9 (7.9) ●	Hong Kong, SAR	525 (6.3)	537 (5.1)	12 (8.2) ●
Netherlands	528 (5.7)	536 (7.1)	8 (9.0) ●	England	543 (6.0)	554 (5.3)	11 (7.9) ●
Korea, Rep. of	530 (2.5)	538 (4.0)	8 (4.8) ●	Russian Federation	530 (5.1)	540 (6.2)	9 (8.2) ●
<b>International Avg. §</b>	<b>506 (1.1)</b>	<b>512 (1.0)</b>	<b>6 (1.5) ▲</b>	United States	520 (6.1)	524 (5.5)	5 (8.2) ●
Slovak Republic	520 (4.1)	525 (3.4)	5 (5.4) ●	<b>International Avg. §</b>	<b>527 (1.1)</b>	<b>531 (1.1)</b>	<b>3 (1.6) ●</b>
Romania	464 (5.4)	468 (6.4)	4 (8.4) ●	Italy	503 (3.8)	505 (6.4)	2 (7.1) ●
Russian Federation	516 (4.5)	519 (7.1)	4 (8.6) ●	Belgium (Flemish)	542 (9.0)	544 (7.2)	2 (11.7) ●
Belgium (Flemish)	524 (8.7)	526 (4.6)	2 (9.7) ●	Slovak Republic	545 (3.3)	546 (4.5)	1 (5.4) ●
Cyprus	454 (2.9)	455 (3.1)	1 (4.5) ●	Korea, Rep. of	559 (2.8)	559 (3.2)	0 (4.5) ●
Slovenia	526 (3.3)	527 (3.7)	0 (5.1) ●	Netherlands	554 (7.4)	554 (7.3)	0 (10.4) ●
United States	505 (5.4)	505 (4.6)	0 (7.1) ●	Romania	478 (5.6)	475 (6.5)	-3 (8.7) ●
England	522 (4.0)	522 (6.2)	-1 (7.5) ●	Japan	564 (2.2)	556 (3.6)	-7 (4.6) ●
Italy	492 (4.5)	491 (5.1)	-1 (6.9) ●	Singapore	587 (7.0)	578 (9.7)	-9 (12.0) ●
Japan	544 (1.9)	543 (2.8)	-2 (3.5) ●	New Zealand	524 (6.1)	513 (7.0)	-11 (9.4) ●
Czech Republic	538 (5.7)	523 (4.8)	-14 (7.6) ●	Iran, Islamic Rep.	475 (4.6)	461 (4.4)	-14 (6.2) ●
Singapore	574 (6.7)	557 (7.9)	-16 (10.4) ●	Czech Republic	572 (4.8)	557 (4.9)	-15 (6.8) ●
Iran, Islamic Rep.	448 (5.7)	430 (5.7)	-18 (8.2) ●	Slovenia	556 (3.3)	540 (3.7)	-16 (5.0) ▼
<b>Countries with Unapproved Sampling Procedures at the Classroom Level in 1995</b>				<b>Countries with Unapproved Sampling Procedures at the Classroom Level in 1995</b>			
Israel	494 (6.9)	476 (6.6)	-17 (9.2) ●	Israel	532 (6.8)	492 (6.2)	-39 (9.0) ▼
South Africa	243 (9.7)	234 (9.2)	-9 (13.4) ●	South Africa	283 (15.4)	253 (7.7)	-30 (17.3) ●
Thailand	511 (5.4)	481 (4.6)	-30 (7.1) ▼	Thailand	509 (4.9)	484 (4.4)	-25 (6.7) ▼

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

- ▲ 1999 significantly higher than 1995
  - No significant difference between 1995 and 1999
  - ▼ 1999 significantly lower than 1995
- Significance tests adjusted for multiple comparisons

§ International average is for countries that participated and met sampling guidelines in both 1995 and 1999.

Trend notes: Because coverage fell below 65% in 1995 and 1999, Latvia is annotated LSS for Latvian-Speaking Schools only. Lithuania tested later in 1999 than in 1995, at the beginning of the next school year. In 1995, Italy and Israel were unable to cover their International Desired Population; 1999 data are based on their comparable populations. Trends in gender data for Bulgaria are unavailable.

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

# Exhibit 1.14 Trends in Gender Differences in Average Science Achievement

	1995			1999			Change in Gender Difference*
	Girls Average Scale Score	Boys Average Scale Score	Difference (Absolute Value)	Girls Average Scale Score	Boys Average Scale Score	Difference (Absolute Value)	
Australia	520 (4.4)	533 (5.5)	14 (5.8)	532 (5.1)	549 (6.0)	18 (6.8)	↔
Belgium (Flemish)	524 (8.7)	542 (9.0)	18 (12.6)	526 (4.6)	544 (7.2)	18 (10.3)	↔
Canada	508 (3.2)	521 (3.4) ▲	13 (3.8)	526 (3.2)	540 (2.4) ▲	14 (3.9)	↔
Cyprus	454 (2.9)	451 (2.4)	3 (3.4)	455 (3.1)	465 (3.0)	10 (3.9)	↔
Czech Republic	538 (5.7)	572 (4.8) ▲	34 (4.2)	523 (4.8)	557 (4.9) ▲	33 (4.8)	↔
England	522 (4.0)	543 (6.0)	21 (7.8)	522 (6.2)	554 (5.3) ▲	32 (6.6)	↔
Hong Kong, SAR	492 (6.5)	525 (6.3) ▲	33 (6.6)	522 (4.4)	537 (5.1)	14 (6.1)	↘
Hungary	525 (3.7)	549 (3.5) ▲	23 (3.7)	540 (4.0)	565 (4.5) ▲	25 (4.2)	↔
Iran, Islamic Rep.	448 (5.7)	475 (4.6) ▲	27 (7.6)	430 (5.7)	461 (4.4) ▲	31 (7.6)	↔
Italy	492 (4.5)	503 (3.8)	11 (4.4)	491 (5.1)	505 (6.4)	14 (6.2)	↔
Japan	544 (1.9)	564 (2.2) ▲	19 (2.4)	543 (2.8)	556 (3.6)	14 (4.6)	↔
Korea, Rep. of	530 (2.5)	559 (2.8) ▲	29 (4.2)	538 (4.0)	559 (3.2) ▲	21 (5.1)	↔
Latvia (LSS)	464 (3.8)	490 (4.3) ▲	26 (4.4)	495 (5.6)	510 (4.8) ▲	15 (4.0)	↔
Lithuania	452 (4.3)	477 (4.5) ▲	25 (3.5)	478 (4.4)	499 (5.0) ▲	21 (4.6)	↔
Netherlands	528 (5.7)	554 (7.4) ▲	26 (4.4)	536 (7.1)	554 (7.3) ▲	18 (4.1)	↔
New Zealand	497 (5.6)	524 (6.1) ▲	27 (6.6)	506 (5.4)	513 (7.0)	7 (7.8)	↔
Romania	464 (5.4)	478 (5.6) ▲	14 (3.7)	468 (6.4)	475 (6.5)	7 (5.4)	↔
Russian Federation	516 (4.5)	530 (5.1) ▲	15 (3.6)	519 (7.1)	540 (6.2) ▲	20 (3.9)	↔
Singapore	574 (6.7)	587 (7.0)	13 (8.1)	557 (7.9)	578 (9.7)	20 (7.9)	↔
Slovak Republic	520 (4.1)	545 (3.3) ▲	25 (3.3)	525 (3.4)	546 (4.5) ▲	21 (4.5)	↔
Slovenia	526 (3.3)	556 (3.3) ▲	30 (3.7)	527 (3.7)	540 (3.7) ▲	13 (3.7)	↘
United States	505 (5.4)	520 (6.1) ▲	14 (3.0)	505 (4.6)	524 (5.5) ▲	19 (4.1)	↔
<b>International Avg. §</b>	506 (1.0)	527 (1.1) ▲	21 (1.2)	512 (1.1)	531 (1.2) ▲	18 (1.2)	↘
<b>Countries with Unapproved Sampling Procedures at the Classroom Level in 1995</b>							
Israel	494 (6.9)	532 (6.8) ▲	38 (6.3)	476 (6.6)	492 (6.2)	16 (6.3)	↘
South Africa	243 (9.7)	283 (15.4) ▲	40 (13.0)	234 (9.2)	253 (7.7)	19 (6.7)	↔
Thailand	511 (5.4)	509 (4.9)	2 (4.6)	481 (4.6)	484 (4.4)	3 (4.3)	↔

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

▲ Significantly higher than other gender

Significance tests adjusted for multiple comparisons

Increased ↗

Decreased ↘

No change ↔

\* Indicates whether 1999 gender difference is significantly different than 1995 gender difference.

§ International average is for countries that participated and met sampling guidelines in both 1995 and 1999.

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Trend notes: Because coverage fell below 65% in 1995 and 1999, Latvia is annotated LSS for Latvian-Speaking Schools only. Lithuania tested later in 1999 than in 1995, at the beginning of the next school year. In 1995, Italy and Israel were unable to cover their International Desired Population; 1999 data are based on their comparable populations. Trends in gender data for Bulgaria are unavailable.

