

Chapter 1



International Student Achievement in Science

Korea and Singapore were the top-performing countries in science in TIMSS 2011 at the fourth grade, followed by Finland, Japan, the Russian Federation, and Chinese Taipei. At the eighth grade, Singapore had the highest average achievement, followed by Korea, Chinese Taipei, and Japan. Finland was the next highest-performing country.

Since 1995, fourth grade students have shown more improvement than reduction in science achievement (8 countries up vs. only 1 down), but improving eighth grade student achievement has been more difficult (11 up vs. 6 down).

Chapter 1 contains the science achievement results for the 52 countries and seven benchmarking participants in the fourth grade TIMSS 2011 assessment and the 45 countries and 14 benchmarking participants in the eighth grade TIMSS 2011 assessment. To summarize science achievement across the participants at fourth and eighth grades, the chapter provides:

- ◆ Averages (means) and distributions of science achievement;
- ◆ Trends in science achievement over time for participants in previous TIMSS assessments in 1995, 1999, 2003, and 2007;
- ◆ Trends across grades—Relative achievement of the 2007 fourth grade cohort as eighth grade students in 2011;
- ◆ Achievement differences by gender; and
- ◆ Trends in achievement differences by gender.

The results for percentages of students reaching the TIMSS International Benchmarks (Advanced, High, Intermediate, and Low) are presented in Chapter 2.

Science Achievement Across Countries

TIMSS 2011 Science Achievement

This section reports the TIMSS 2011 science results as average scores and distributions on the fourth and eighth grade TIMSS scales, each of which has a range of 0–1,000 (although student performance typically ranges between 300 and 700). The TIMSS science achievement scales were established in TIMSS 1995 based on the achievement distribution across all participating countries, treating each country equally. At each grade level, the scale centerpoint of 500 was set to correspond to the mean of the overall achievement distribution, and 100 points on the scale was set to correspond to the standard deviation. Achievement data from subsequent TIMSS assessment cycles were linked to these scales so that increases or decreases in average achievement may be monitored across assessments.¹ TIMSS uses the scale centerpoint as a point of reference that remains constant from assessment to assessment.

Exhibit 1.1 shows the distributions of student achievement for the participants in the TIMSS 2011 fourth grade assessment, including the average scale score with its 95 percent confidence interval and the ranges in performance for the middle half of the students (25th to 75th percentiles) as well as the extremes (5th and 95th percentiles). Similarly, Exhibit 1.2 shows the

¹ Please see *Methods and Procedures in TIMSS and PIRLS 2011* on the TIMSS and PIRLS website for further detail (<http://timssandpirls.bc.edu>).

distribution of science achievement for participants in the TIMSS 2011 eighth grade assessment.

The first page of Exhibit 1.1 presents the results for the 50 countries that assessed students at the TIMSS target population of the fourth grade. In particular, the TIMSS target population for the fourth grade assessment is the grade that represents four years of schooling, counting from the first year of ISCED Level 1.² Level 1 corresponds to primary education or the first stage of basic education, with the first year of Level 1 marking “systematic apprenticeship of reading, writing, and mathematics.” However, IEA has a policy that children should be at least 9 years old before being asked to participate in a paper-and-pencil assessment such as TIMSS. Thus, as a policy, TIMSS also tries to ensure that, at the time of testing, students do not fall under the minimum average age of 9.5 years old. So, England, Malta, and New Zealand, where students start school at a young age, were assessed in their fifth year of schooling, but still have among the youngest students and are reported together with the fourth grade countries. Exhibit C.1 in Appendix C shows the grades and average ages of the students tested across countries, together with information about the policies and practices related to age of entry to primary school across countries. The *TIMSS 2011 Encyclopedia* contains further details, such as countries’ policies about promotion and retention.

The second page of Exhibit 1.1 shows the results for three countries that assessed their sixth grade students. To meet the needs of the increasing number of developing countries wanting to participate in TIMSS 2011, the TIMSS & PIRLS International Study Center encouraged countries where the assessment was too difficult for fourth grade students to give the TIMSS fourth grade assessment at the sixth grade. The three countries that elected to assess sixth grade students were Botswana, Honduras, and Yemen (which also assessed its fourth grade students).

The second page of Exhibit 1.1 also presents the results for the TIMSS 2011 fourth grade benchmarking participants. The benchmarking participants followed the same procedures and met the same standards as the countries, the difference being that they are regional entities of countries. Benchmarking participants at the fourth grade included Florida and North Carolina (US states), Alberta, Ontario, and Québec (Canadian provinces), and Dubai and Abu Dhabi (emirates of the United Arab Emirates).

Following the same approach as Exhibit 1.1, the first page of Exhibit 1.2 presents the results for the 42 countries that assessed students at the TIMSS

2 ISCED stands for the International Standard Classification of Education developed by the UNESCO Institute for Statistics (OECD, 1999).

target population of the eighth grade, the grade that represents eight years of schooling. For the TIMSS eighth grade assessment, IEA has a policy that students should be at least 13 years old before being asked to participate. Thus, TIMSS tries to ensure that, at the time of testing, students do not fall under the minimum average age of 13.5 years old. So, England and New Zealand, where students start school at a young age, are reported together with the eighth grade countries. Exhibit C.1 in Appendix C shows the grades and average ages of students at the time of testing across countries, together with policies related to age of entry into school.

As with the fourth grade, the TIMSS & PIRLS International Study Center encouraged countries where the TIMSS eighth grade assessment was too difficult for eighth grade students to instead assess students at a higher grade. The second page of Exhibit 1.2 shows the results for three countries that assessed their ninth grade students—Botswana, Honduras, and South Africa.

The second page of Exhibit 1.2 also presents the results for the TIMSS 2011 eighth grade benchmarking participants. Benchmarking participants at the eighth grade included nine US states (Alabama, California, Colorado, Connecticut, Florida, Indiana, Massachusetts, Minnesota, and North Carolina), three Canadian provinces (Alberta, Ontario, and Québec), and two emirates (Dubai and Abu Dhabi).

For each section of Exhibit 1.1 and in Exhibit 1.2, participants are shown in decreasing order of average achievement. Also, there is a symbol by a participant's average scale score indicating if the average achievement is significantly higher (up arrow) or lower (down arrow) than the scale centerpoint of 500. TIMSS uses the centerpoint of the scale as a point of reference that remains constant from assessment to assessment. (In contrast, the international average, obtained by averaging across the mean scores for each of the participating countries, changes from assessment to assessment as the number and characteristics of the participating countries change.) Finally, several countries have annotations about 1) population coverage (detailed in Exhibit C.2); 2) sampling participation rates (explained in Exhibit C.8); and 3) the potential for bias in their achievement estimates (explained in the section after next).

Achievement in TIMSS 2011 at the Fourth Grade

The results in Exhibit 1.1 (first page) show that many countries performed well in TIMSS 2011 at the fourth grade, with 27 countries having higher achievement

than the scale centerpoint of 500 and several countries having average achievement above the High International Benchmark of 550. Because there are often relatively small differences between participants in average achievement, Exhibit 1.3 shows whether or not the differences in average achievement among the countries are statistically significant.

Korea and Singapore were the top-performing countries in TIMSS 2011 at the fourth grade. Looking at the results in Exhibit 1.1 and taking into account the information in Exhibit 1.3, it can be seen that these two countries performed similarly and had higher achievement than all of the other countries. The next highest-performing country was Finland, which had higher achievement than all countries except the two with the highest achievement, followed by Japan, the Russian Federation, and Chinese Taipei. The United States was the next highest performing country, with achievement higher than all countries except the six top performers. Also included in the top 14 high-achieving countries were the Czech Republic, Hong Kong SAR, Hungary, Sweden, the Slovak Republic, Austria, and the Netherlands. The benchmarking states of Florida and North Carolina and the province of Alberta had performance similar to these countries.

While there were small differences from country to country, there was a substantial range in performance from the top-performing to the lower-performing countries. Eighteen countries had average achievement below the TIMSS centerpoint of 500. For the most part, these countries had average achievement above the Low (400) International Benchmark.

Very Low Performance on TIMSS 2011

It is a well-known principle of educational measurement that the difficulty of the items used to assess student achievement should match the ability of the students taking the assessment. In the context of assessing science achievement, measurement is most efficient when there is a reasonable match between the science ability level of the student population being assessed and the difficulty of the assessment items. The greater the mismatch, the more difficult it becomes to achieve reliable measurement. In particular, when the assessment tasks are much too challenging for most students, to the extent that many students are responding at chance level, it is extremely difficult to achieve acceptable measurement quality.

Monitoring trends over time is particularly problematic for a country with a high degree of mismatch between assessment difficulty and student

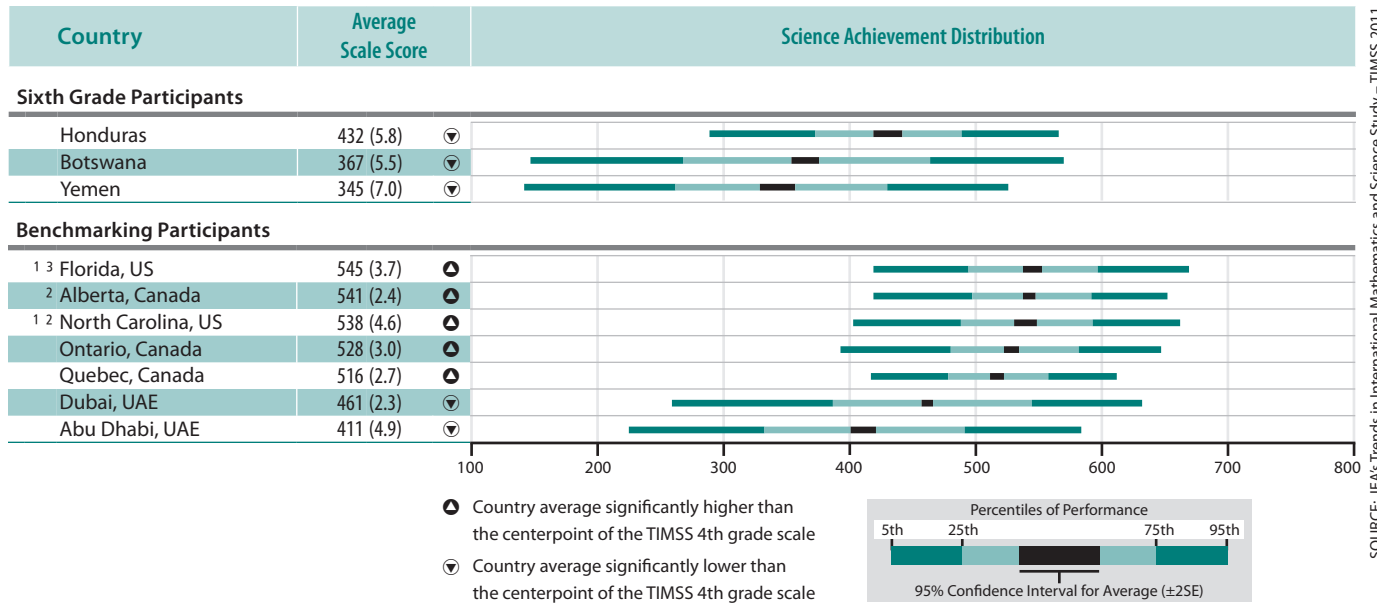
Exhibit 1.1: Distribution of Science Achievement



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

✱ Average achievement not reliably measured because the percentage of students with achievement too low for estimation exceeds 25%.
 Ψ Reservations about reliability of average achievement because the percentage of students with achievement too low for estimation does not exceed 25% but exceeds 15%.
 See Appendix C.2 for target population coverage notes 1, 2, and 3. See Appendix C.8 for sampling guidelines and sampling participation notes †, ‡, and §.
 () Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.1: Distribution of Science Achievement (Continued)



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

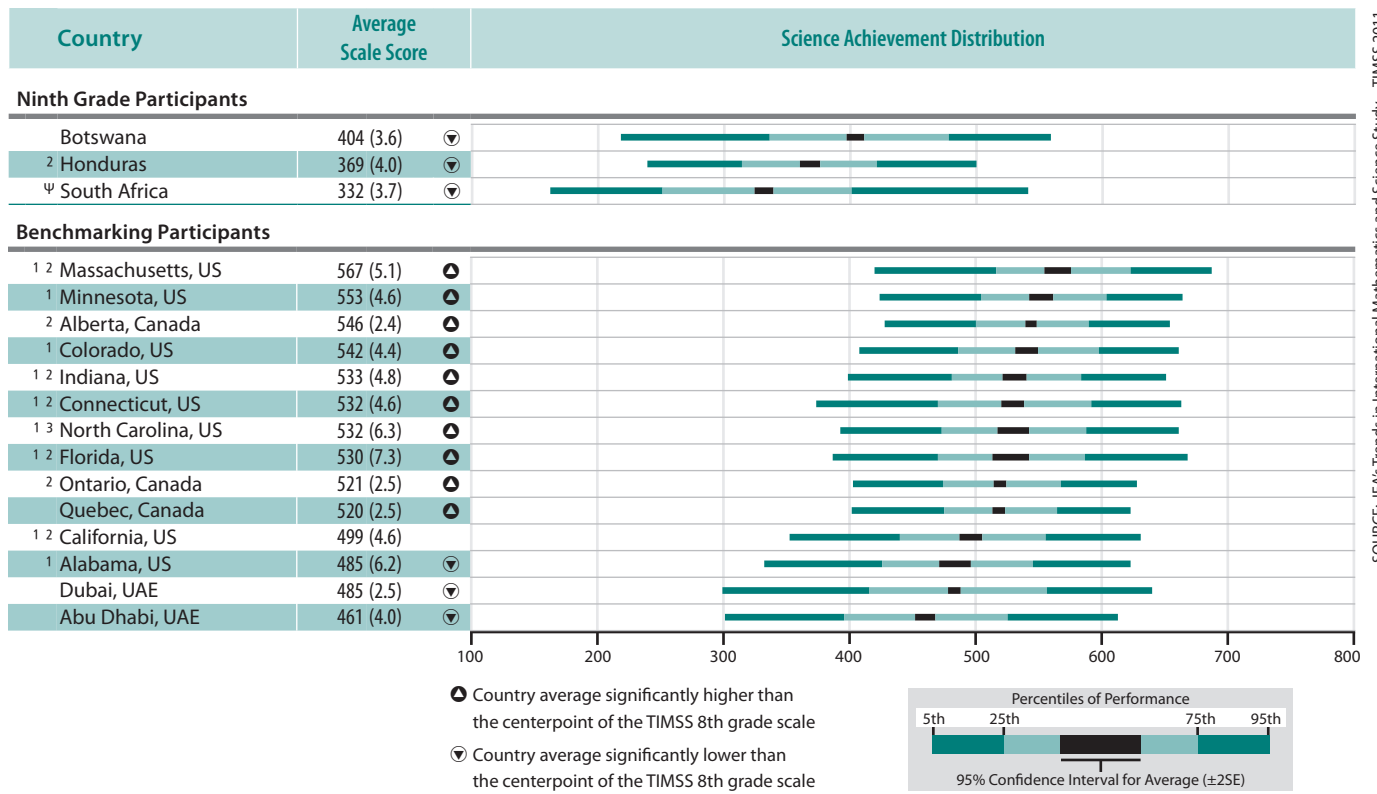
Exhibit 1.2: Distribution of Science Achievement



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Ψ Reservations about reliability of average achievement because the percentage of students with achievement too low for estimation does not exceed 25% but exceeds 15%. See Appendix C.3 for target population coverage notes 1, 2, and 3. See Appendix C.9 for sampling guidelines and sampling participation notes †, ‡, and §.
() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.2: Distribution of Science Achievement (Continued)



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 1.3: 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.

Country	Average Scale Score	Korea, Rep. of	Singapore	Finland	Japan	Russian Federation	Chinese Taipei	United States	Czech Republic	Hong Kong SAR	Hungary	Sweden	Slovak Republic	Austria	Netherlands	England	Denmark	Germany	Italy	Portugal	Slovenia	Northern Ireland	Ireland	Croatia	Australia	Serbia	Lithuania	Belgium (Flemish)	Romania	Spain	Poland
Korea, Rep. of	587 (2.0)																														
Singapore	583 (3.4)																														
Finland	570 (2.6)	▼	▼																												
Japan	559 (1.9)	▼	▼	▼																											
Russian Federation	552 (3.5)	▼	▼	▼	▼																										
Chinese Taipei	552 (2.2)	▼	▼	▼	▼																										
United States	544 (2.1)	▼	▼	▼	▼	▼																									
Czech Republic	536 (2.5)	▼	▼	▼	▼	▼	▼																								
Hong Kong SAR	535 (3.8)	▼	▼	▼	▼	▼	▼	▼																							
Hungary	534 (3.7)	▼	▼	▼	▼	▼	▼	▼	▼																						
Sweden	533 (2.7)	▼	▼	▼	▼	▼	▼	▼	▼	▼																					
Slovak Republic	532 (3.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼																				
Austria	532 (2.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼																			
Netherlands	531 (2.2)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼																		
England	529 (2.9)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼																	
Denmark	528 (2.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼																
Germany	528 (2.9)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼															
Italy	524 (2.7)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼														
Portugal	522 (3.9)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼													
Slovenia	520 (2.7)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼												
Northern Ireland	517 (2.6)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼											
Ireland	516 (3.4)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼										
Croatia	516 (2.1)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼									
Australia	516 (2.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼								
Serbia	516 (3.1)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼							
Lithuania	515 (2.4)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼						
Belgium (Flemish)	509 (2.0)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼					
Romania	505 (5.9)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼				
Spain	505 (3.0)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Poland	505 (2.6)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
New Zealand	497 (2.3)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Kazakhstan	495 (5.1)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Norway	494 (2.3)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Chile	480 (2.4)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Thailand	472 (5.6)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Turkey	463 (4.5)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Georgia	455 (3.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Iran, Islamic Rep. of	453 (3.7)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Bahrain	449 (3.5)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Malta	446 (1.9)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Azerbaijan	438 (5.6)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Saudi Arabia	429 (5.4)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
United Arab Emirates	428 (2.5)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Armenia	416 (3.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Qatar	394 (4.3)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Oman	377 (4.3)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Kuwait	347 (4.7)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Tunisia	346 (5.3)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Morocco	264 (4.5)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Yemen	209 (7.3)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Honduras (6)	432 (5.8)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Botswana (6)	367 (5.5)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Yemen (6)	345 (7.0)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Benchmarking Participants																															
Florida, US	545 (3.7)	▼	▼	▼	▼						▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	
Alberta, Canada	541 (2.4)	▼	▼	▼	▼	▼	▼					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	
North Carolina, US	538 (4.6)	▼	▼	▼	▼	▼	▼	▼																							
Ontario, Canada	528 (3.0)	▼	▼	▼	▼	▼	▼	▼	▼																						
Quebec, Canada	516 (2.7)	▼	▼	▼	▼	▼	▼	▼	▼	▼																					
Dubai, UAE	461 (2.3)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			
Abu Dhabi, UAE	411 (4.9)	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼			

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Significance tests were not adjusted for multiple comparisons. Five percent of the comparisons would be statistically significant by chance alone.
 () Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.4: Multiple Comparisons of Average Science Achievement (Continued)

Average achievement significantly higher than comparison country										Average achievement significantly lower than comparison country											
Benchmarking Participants										Benchmarking Participants										Average Scale Score	Country
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	590 (4.3)	Singapore
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	564 (2.3)	Chinese Taipei
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	560 (2.0)	Korea, Rep. of
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	558 (2.4)	Japan
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	552 (2.5)	Finland
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	543 (2.7)	Slovenia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	542 (3.2)	Russian Federation
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	535 (3.4)	Hong Kong SAR
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	533 (4.9)	England
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	525 (2.6)	United States
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	522 (3.1)	Hungary
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	519 (4.8)	Australia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	516 (4.0)	Israel
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	514 (2.6)	Lithuania
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	512 (4.6)	New Zealand
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	509 (2.5)	Sweden
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	501 (2.5)	Italy
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	501 (3.4)	Ukraine
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	494 (2.6)	Norway
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	490 (4.3)	Kazakhstan
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	483 (3.4)	Turkey
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	474 (4.0)	Iran, Islamic Rep. of
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	465 (3.5)	Romania
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	465 (2.4)	United Arab Emirates
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	461 (2.5)	Chile
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	452 (2.0)	Bahrain
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	451 (3.9)	Thailand
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	449 (4.0)	Jordan
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	439 (2.5)	Tunisia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	437 (3.1)	Armenia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	436 (3.9)	Saudi Arabia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	426 (6.3)	Malaysia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	426 (3.9)	Syrian Arab Republic
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	420 (3.2)	Palestinian Nat'l Auth.
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	420 (3.0)	Georgia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	420 (3.2)	Oman
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	419 (3.4)	Qatar
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	407 (5.4)	Macedonia, Rep. of
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	406 (4.9)	Lebanon
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	406 (4.5)	Indonesia
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	376 (2.2)	Morocco
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	306 (5.2)	Ghana
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	404 (3.6)	Botswana (9)
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	369 (4.0)	Honduras (9)
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	332 (3.7)	South Africa (9)
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	567 (5.1)	Massachusetts, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	553 (4.6)	Minnesota, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	546 (2.4)	Alberta, Canada
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	542 (4.4)	Colorado, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	533 (4.8)	Indiana, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	532 (4.6)	Connecticut, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	532 (6.3)	North Carolina, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	530 (7.3)	Florida, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	521 (2.5)	Ontario, Canada
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	520 (2.5)	Quebec, Canada
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	499 (4.6)	California, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	485 (6.2)	Alabama, US
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	485 (2.5)	Dubai, UAE
⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	⬆	461 (4.0)	Abu Dhabi, UAE

Significance tests were not adjusted for multiple comparisons. Five percent of the comparisons would be statistically significant by chance alone.

() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

achievement. If there are substantial numbers of students with very low scores, their achievement is likely to be overestimated and consequently the overall achievement distribution becomes biased upwards. Educators and policy makers may work hard and make real strides in improving education from this assessment cycle to the next. However, because the achievement distribution at the earlier cycle was overestimated to begin with, the country would not see evidence of this improvement in the assessment results. The apparently poor return for all of the effort could be very disheartening to those who worked so hard and could prove a disincentive to further investment and effort.

Having substantial numbers of students with very low scores in a country also makes it difficult to estimate performance separately for the science content and cognitive domains. The items comprising the science reasoning scale were particularly difficult for such countries.

To identify countries where performance is deemed too low to provide reliable measurement of achievement and meaningful trend comparisons, the TIMSS & PIRLS International Study Center conducted extensive investigations to detect when the quality of measurement erodes (Martin, Mullis, & Foy, in press). The proportion of students unable to respond to any items on the assessment was selected as the best indicator of degree of mismatch between students' skills and those demanded by the assessment. Although the absolute lower limit would be no items answered correctly, about half the items were in multiple-choice format and guessing on these was possible. Thus, beginning in 2011, the criterion for having achievement too low for estimation was established based on the percentage of the students having a score no higher than what a student would achieve by guessing on all the multiple-choice questions—essentially the percentage of students performing below chance.

For each country, Appendix D shows the percentage of students with achievement too low for estimation (Exhibit D.1 for the fourth grade and D.2 for the eighth grade). When, as in Morocco and Yemen at the fourth grade, the percentage of students with achievement too low for estimation exceeded 25 percent, the country was annotated with the symbol Ж. Achievement trends are not reported for these countries because of concerns about bias in the estimation of achievement for the student population. When, as in Kuwait and Tunisia, the percentage of students with achievement too low for estimation exceeded 15 percent but did not exceed 25 percent, the country was annotated with the symbol Ψ, indicating reservations about the reliability of the achievement estimates.

Achievement in TIMSS 2011 at the Sixth Grade

As a group, the countries assessing their sixth grade students had average achievement between 345 and 432. This level of achievement is comparable to that of most of the lower performing countries at the fourth grade. Despite the low average achievement of the sixth grade students in Yemen, it is noteworthy that it exceeded the average achievement of Yemen's fourth grade students by more than 130 points.

Achievement in TIMSS 2011 at the Eighth Grade

The results in Exhibit 1.2 (first page) show that 16 countries had higher achievement than the scale centerpoint of 500 and five countries had average achievement above the High International Benchmark of 550—Singapore, Chinese Taipei, Korea, Japan, and Finland.

Looking at the results in Exhibit 1.2 and taking into account the information in Exhibit 1.4, which shows whether or not the differences in average achievement among the countries are statistically significant, it can be seen that Singapore had the highest average achievement, and had higher achievement than all other countries. The next highest-performing countries—Chinese Taipei, Korea, and Japan—had higher achievement than all other countries except Singapore. Also included in the top nine high-achieving countries were Finland, Slovenia, the Russian Federation, Hong Kong SAR, and England.

Several benchmarking participants had average science achievement close to the High International Benchmark (550). The state of Massachusetts was outperformed only by the top-performing country of Singapore, and the state of Minnesota was outperformed only by Singapore and Chinese Taipei, although both had achievement similar to a number of other countries.

While there were small differences from country to country, there was a substantial range in performance from the top-performing to the lower-performing countries. Twenty-four countries had average achievement below the scale centerpoint, mostly falling above the Low (400) International Benchmark. Among the eighth grade participating countries, only Ghana had many low performing students, with a percentage of students with achievement too low for estimation between 15 and 25 percent.

Achievement in TIMSS 2011 at the Ninth Grade

As a group, the countries assessing their ninth grade students had average achievement between 332 and 404, at or below the Low International Benchmark (400) for eighth grade students. There was evidence of many very low performing ninth grade students in South Africa, with the percentage of students with achievement too low for estimation between 15 percent and 25 percent.

Trends in Science Achievement

Exhibits 1.5 and 1.6 display changes in average science achievement at the fourth and eighth grades, respectively, for the countries and benchmarking participants that have comparable data from previous TIMSS assessments. For the fourth grade, there are 29 countries and four benchmarking participants having data from 1995, 2003, or 2007 that can be compared to 2011.³ Twelve countries and two benchmarking participants have trend data from all four TIMSS fourth grade assessments. For the eighth grade (and Finland at the seventh grade), there are 35 countries and nine benchmarking participants having data from 1995, 1999, 2003, or 2007 that can be compared to 2011, including eleven countries and two benchmarking participants that have data from all five TIMSS eighth grade assessments. With the participants shown in alphabetical order, Exhibits 1.5 and 1.6 show average achievement for each assessment year, as well as achievement differences between years, with an indication of statistical significance. The science achievement distributions also are shown for each assessment year.

At the fourth grade, there are 17 countries and three benchmarking participants that have comparable data from 1995 and 2011 providing trends over the past 16 years. Exhibit 1.7 shows these countries ordered from most to least growth in achievement over this period, to focus on educational progress across the TIMSS assessment years and complement the complete detail in Exhibit 1.5. Exhibit 1.7 presents for the fourth grade a country-by-country graphical depiction of change in average science achievement from 1995 to 2011, with growth curves aligned country by country to facilitate comparisons of change from assessment to assessment. That is, the same scale is used for each country (10-point intervals), but the part of the scale shown differs according to each countries' average achievement. To complement Exhibit 1.6 and focus on long-term educational progress at the eighth grade, Exhibit 1.8 presents for the eighth grade a similar depiction of change in average achievement from 1995

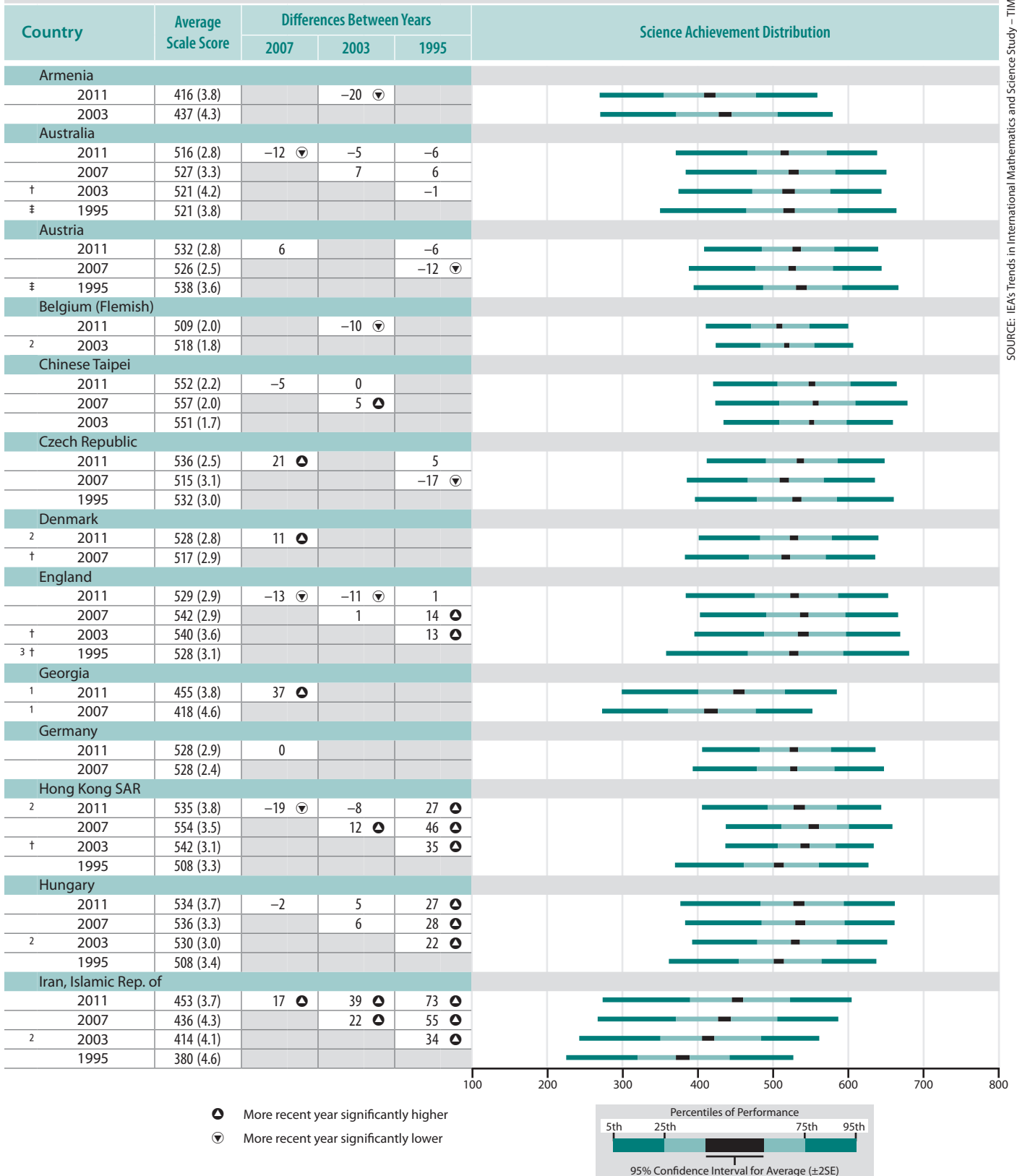
3 TIMSS 1999 did not include a fourth grade assessment.

(or 1999) to 2011 for the 25 countries and eight benchmarking participants that have comparable data from these assessment years.

It is particularly interesting to consider the TIMSS 2011 achievement results in light of the information countries provided in the *TIMSS 2011 Encyclopedia*. Many countries are engaged in implementing important structural, curricular, and instructional reforms and are using the TIMSS results across the assessment years to monitor the impact on achievement of these reforms. Looking at the trends in fourth grade science achievement during the 1995–2011 period, there have been more countries with increases than with decreases. Of the 17 countries and three benchmarking participants with data spanning this period (see Exhibit 1.7), eight countries and one benchmarking participant had increases in average achievement, one country and one benchmarking participant had decreases, and eight countries and one benchmarking participants had no difference. Among the countries with the greatest increase from 1995 to 2011 were Iran, Portugal, Singapore, and Slovenia, with average achievement increases of 56 points or more. Hong Kong SAR and Hungary also both had substantial increases. Among benchmarking participants, Ontario's average achievement increased, while the average achievement decreased in Québec.

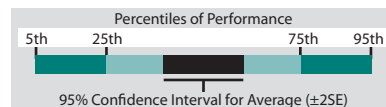
Exhibit 1.5: Trends in Science Achievement

Instructions: Read across the row to determine if the performance in the row year is significantly higher (▲) or significantly lower (▼) than the performance in the column year.



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

▲ More recent year significantly higher
▼ More recent year significantly lower



ψ Reservations about reliability of average achievement because the percentage of students with achievement too low for estimation does not exceed 25% but exceeds 15%. Such annotations in exhibits with trend data began in 2011, so data from assessments prior to 2011 are not annotated for reservations.

See Appendix C.2 for target population coverage notes 1, 2, and 3. See Appendix C.8 for sampling guidelines and sampling participation notes †, ‡, and §.

† Tested the same cohort of students as other countries, but later in the assessment year at the beginning of the next school year.

() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.5: Trends in Science Achievement (Continued)

Instructions: Read across the row to determine if the performance in the row year is significantly higher (▲) or significantly lower (▼) than the performance in the column year.

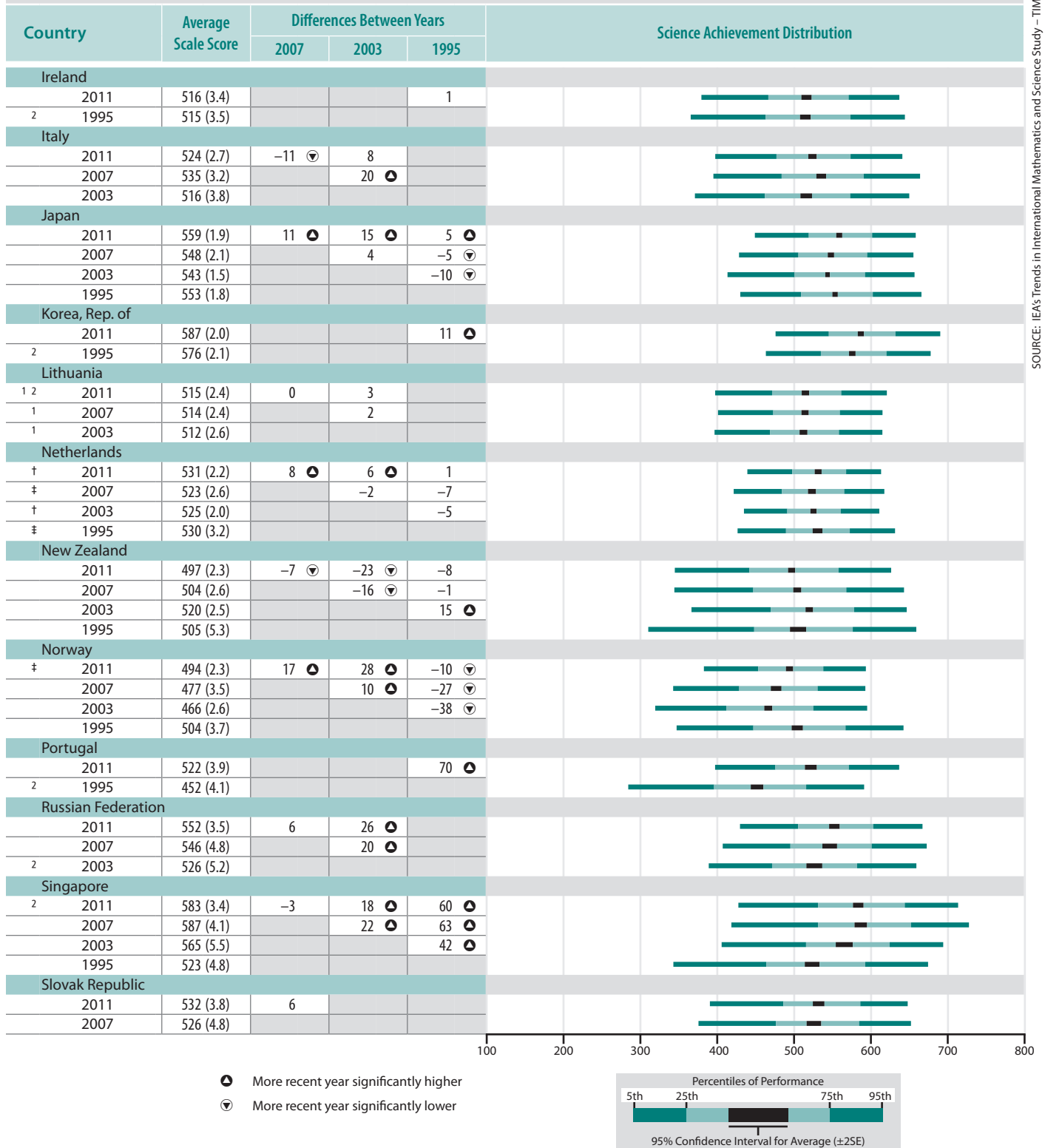
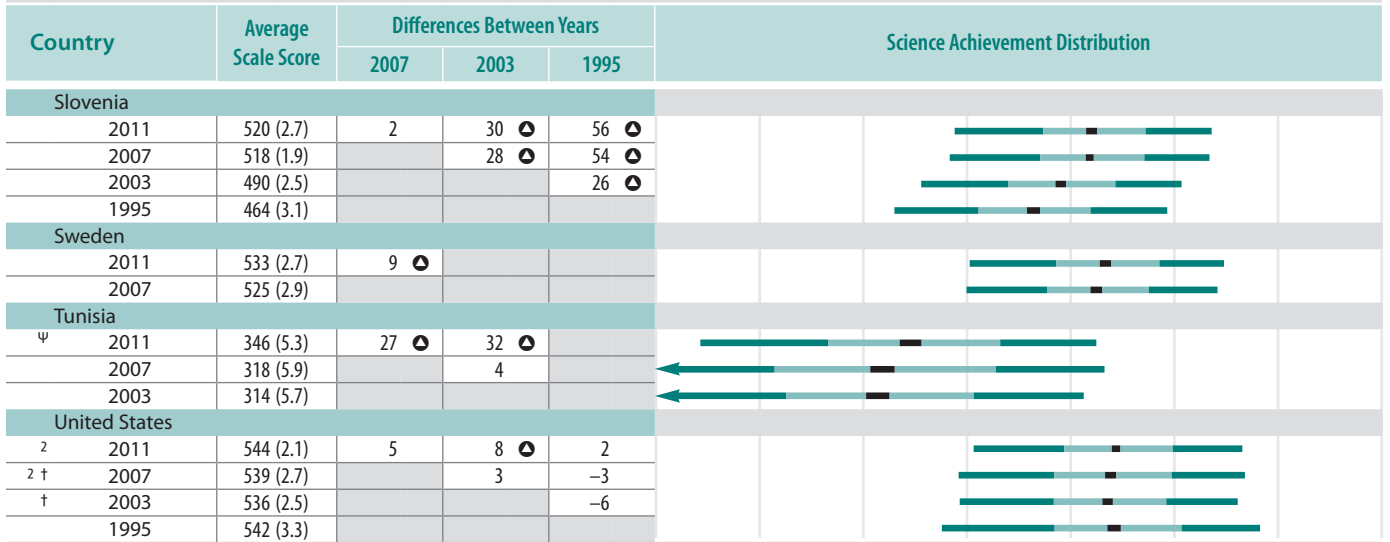


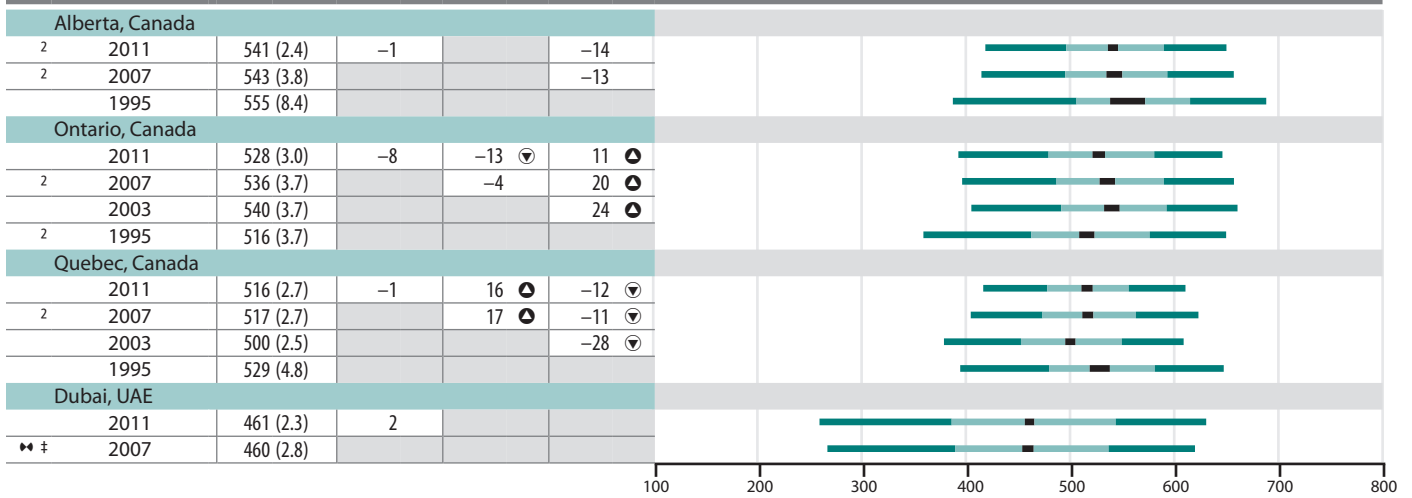
Exhibit 1.5: Trends in Science Achievement (Continued)

Instructions: Read across the row to determine if the performance in the row year is significantly higher (▲) or significantly lower (▼) than the performance in the column year.

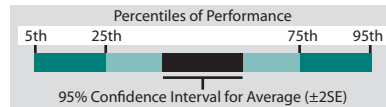


SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Benchmarking Participants



- ▲ More recent year significantly higher
- ▼ More recent year significantly lower

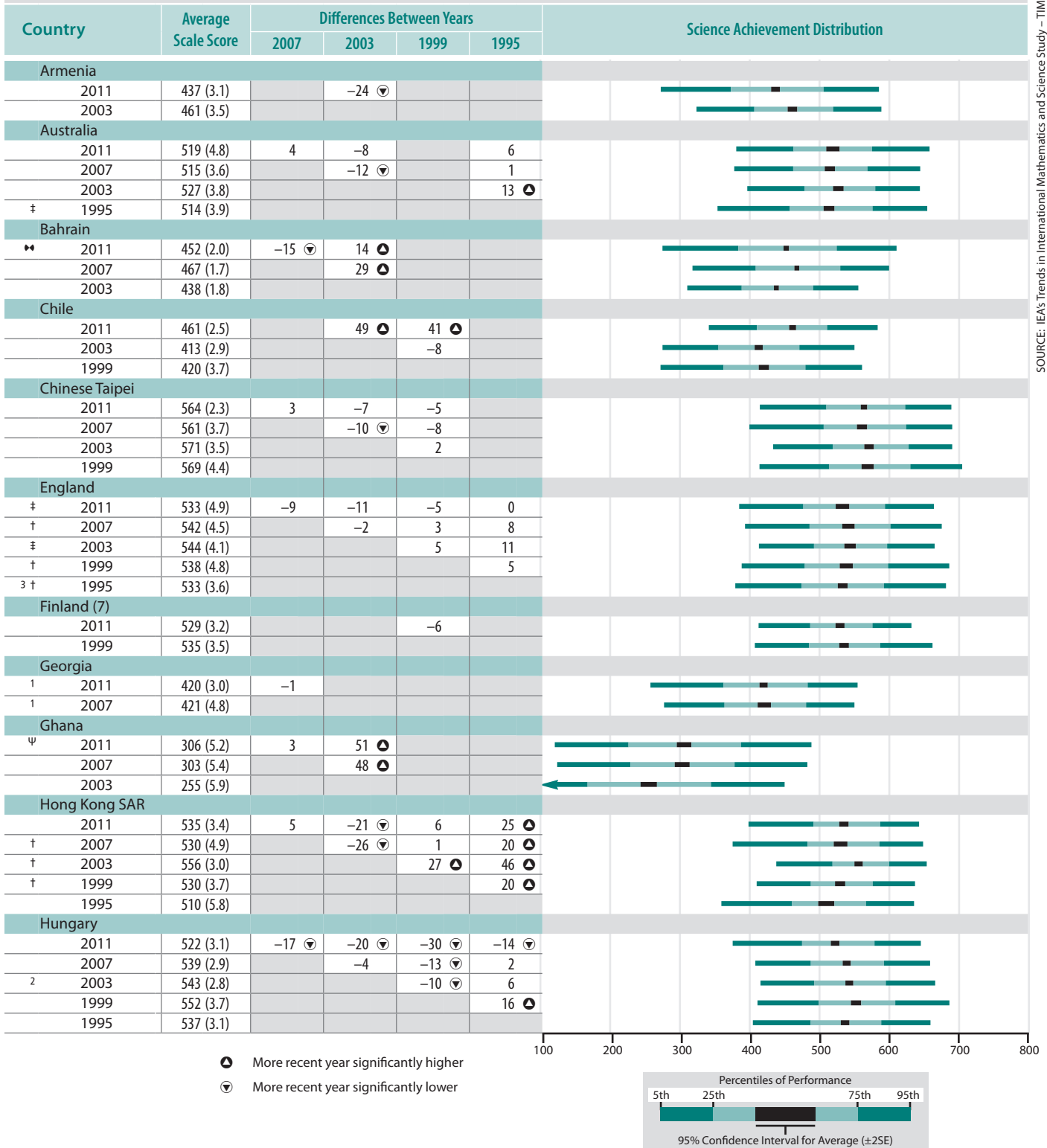


At the eighth grade, there was great variability among countries in changes in average science achievement between 1995 or 1999 and 2011, with some countries showing large improvements and others showing large declines. Of the 16 countries and four benchmarking participants with comparable data spanning the 1995 to 2011 period, seven countries and one benchmarking participant showed increases in average achievement, three countries and no benchmarking participants showed decreases, and six countries and three benchmarking participants showed no difference. The countries with the greatest increases in average science achievement between 1995 and 2011 at the eighth grade were Lithuania (50 points), Slovenia (29 points), Hong Kong SAR (25 points), and the Russian Federation (20 points); Ontario also had a similarly large increase during this period (25 points). Countries with the greatest decreases in average achievement between 1995 and 2011 were Sweden (43 points) and Norway (20 points).

For the nine countries and four benchmarking participants that did not participate in TIMSS 1995, but did participate in TIMSS 1999, two countries and two benchmarking participants showed an increase in average achievement from 1999 to 2011: Chile (41 points) and Tunisia (9 points), as well as the states of Massachusetts and North Carolina (34 and 24 points, respectively). Three countries showed a decrease in average achievement over this period, among which Macedonia and Malaysia showed the largest decreases (51 and 66 points, respectively). Four countries and two benchmarking participants showed no difference.

Exhibit 1.6: Trends in Science Achievement

Instructions: Read across the row to determine if the performance in the row year is significantly higher (▲) or significantly lower (▼) than the performance in the column year.



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

ψ Reservations about reliability of average achievement because the percentage of students with achievement too low for estimation does not exceed 25% but exceeds 15%. Such annotations in exhibits with trend data began in 2011, so data from assessments prior to 2011 are not annotated for reservations.

Trend Notes: Trend results for Finland are based on 7th grade data from 1999 and 2011, and so Finland's 2011 results differ from Exhibit 1.1.

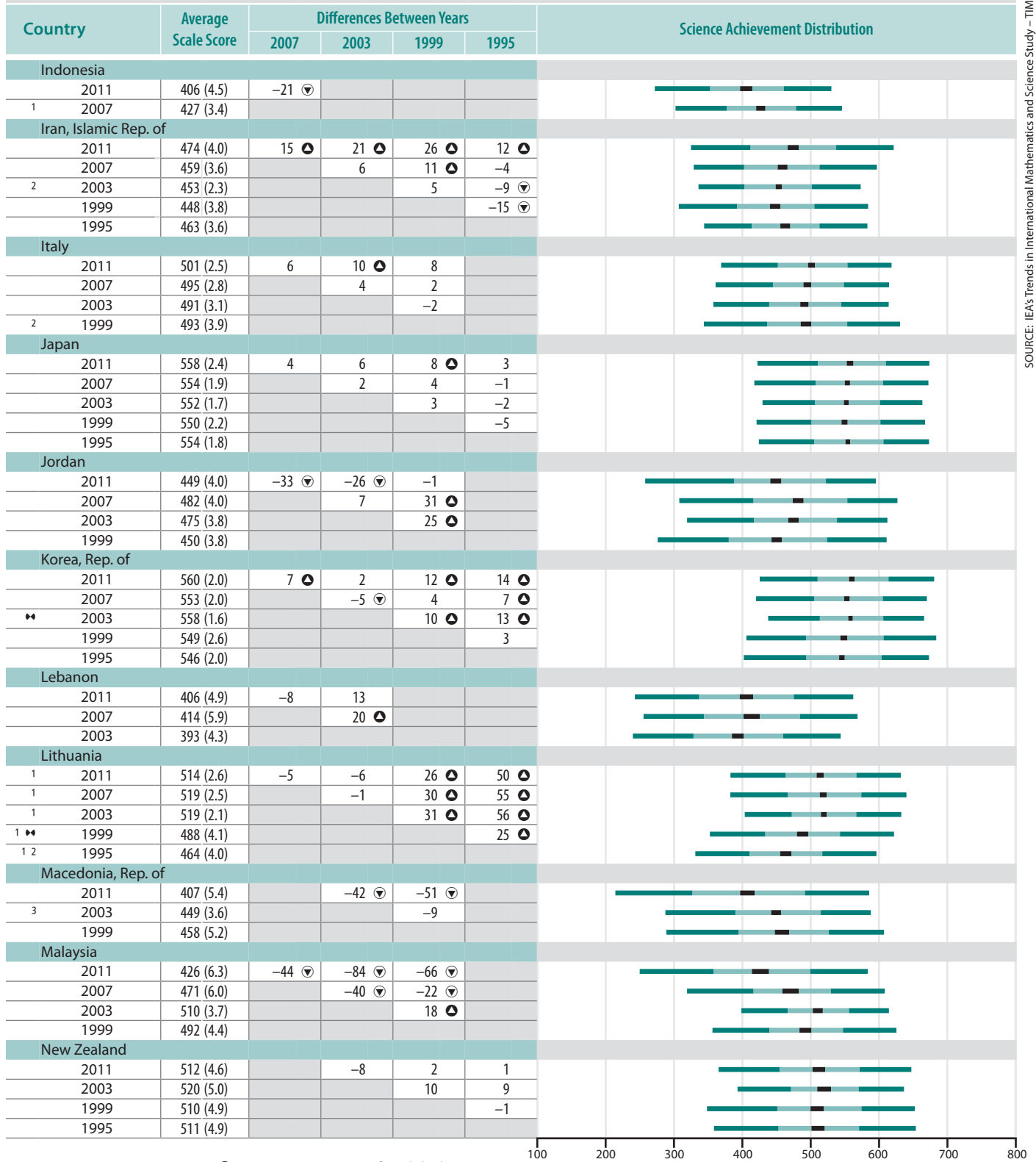
See Appendix C.3 for target population coverage notes 1, 2, and 3. See Appendix C.9 for sampling guidelines and sampling participation notes †, ‡, and §.

♦♦ Tested the same cohort of students as other countries, but later in the assessment year at the beginning of the next school year.

() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.6: Trends in Science Achievement (Continued)

Instructions: Read across the row to determine if the performance in the row year is significantly higher (▲) or significantly lower (▼) than the performance in the column year.



▲ More recent year significantly higher
▼ More recent year significantly lower

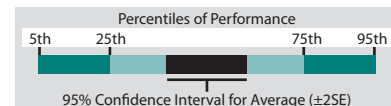
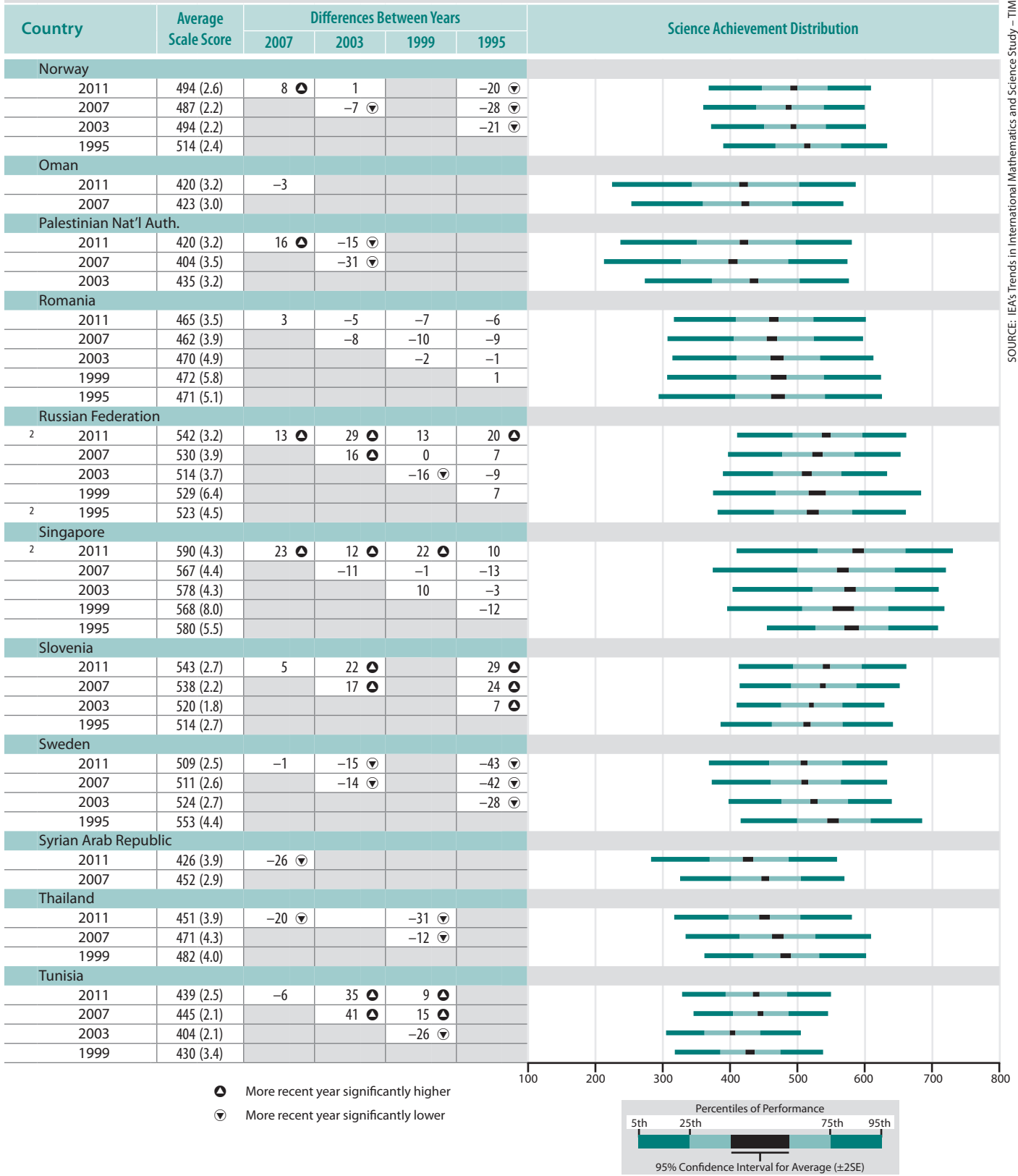


Exhibit 1.6: Trends in Science Achievement (Continued)

Instructions: Read across the row to determine if the performance in the row year is significantly higher (▲) or significantly lower (▼) than the performance in the column year.



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 1.6: Trends in Science Achievement (Continued)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

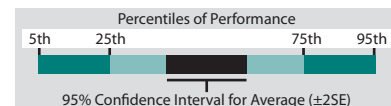
Instructions: Read across the row to determine if the performance in the row year is significantly higher (▲) or significantly lower (▼) than the performance in the column year.

Country	Average Scale Score	Differences Between Years				Science Achievement Distribution
		2007	2003	1999	1995	
Ukraine						
2011	501 (3.4)	16 ▲				
2007	485 (3.5)					
United States						
² 2011	525 (2.6)	5	-3	10	12 ▲	
^{2 †} 2007	520 (2.9)		-7	5	7	
[‡] 2003	527 (3.1)			12 ▲	15 ▲	
1999	515 (4.6)				2	
[†] 1995	513 (5.6)					

Benchmarking Participants

Alberta, Canada						
² 2011	546 (2.4)			-13	-4	
1999	559 (7.5)				9	
1995	550 (4.8)					
Ontario, Canada						
² 2011	521 (2.5)	-5	-12 ▼	3	25 ▲	
² 2007	526 (3.6)		-7	8	30 ▲	
² 2003	533 (2.7)			15 ▲	37 ▲	
1999	518 (3.1)				22 ▲	
1995	496 (3.7)					
Quebec, Canada						
2011	520 (2.5)	13 ▲	-11 ▼	-21 ▼	10	
³ 2007	507 (3.1)		-24 ▼	-34 ▼	-3	
2003	531 (3.0)			-9	21 ▲	
1999	540 (4.8)				30 ▲	
1995	510 (6.9)					
Dubai, UAE						
2011	485 (2.5)	-4				
^{♦ ‡} 2007	489 (2.8)					
Connecticut, US						
¹ 2011	532 (4.6)			2		
1999	529 (10.4)					
Indiana, US						
^{1 2} 2011	533 (4.8)		2	-1		
² 2003	531 (4.8)			-4		
^{2 †} 1999	534 (7.0)					
Massachusetts, US						
^{1 2} 2011	567 (5.1)	11		34 ▲		
² 2007	556 (4.6)			23 ▲		
1999	533 (7.4)					
Minnesota, US						
¹ 2011	553 (4.6)	15 ▲			10	
^{2 †} 2007	539 (4.8)				-5	
[†] 1995	544 (7.9)					
North Carolina, US						
^{1 3} 2011	532 (6.3)			24 ▲		
1999	508 (6.5)					

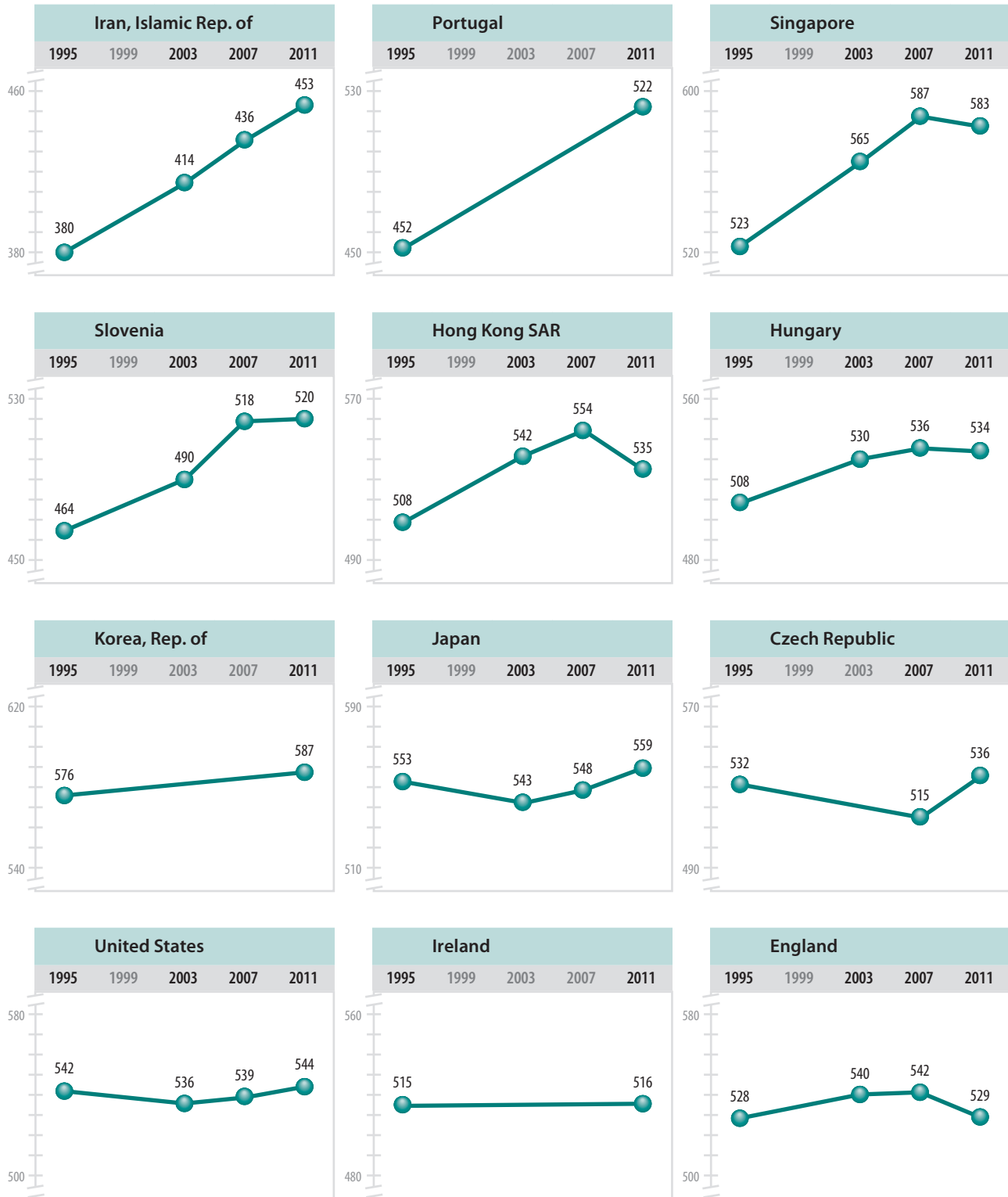
- ▲ More recent year significantly higher
- ▼ More recent year significantly lower



100 200 300 400 500 600 700 800

Exhibit 1.7: Trends in Science Achievement – 1995 Through 2011*

Includes only 2011 participants with comparable long term trend data beginning in 1995, ordered by most to least improvement in average achievement. Exhibit 1.5 provides details including statistical significance.

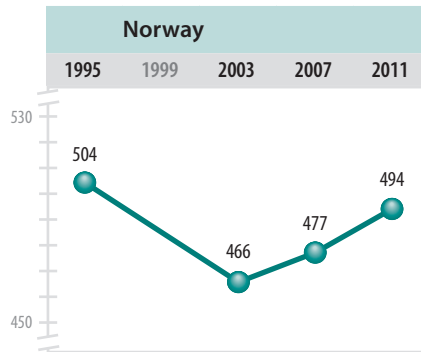
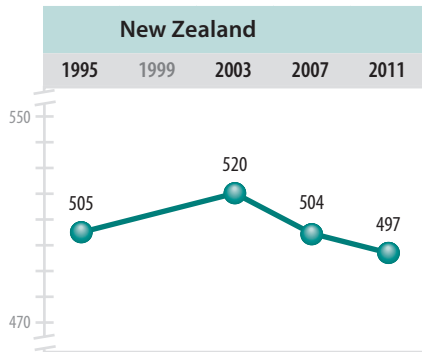
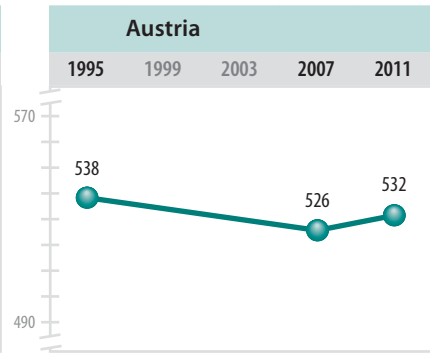
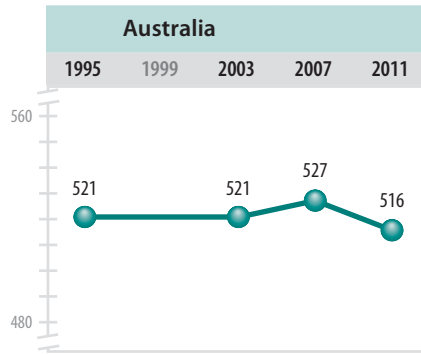


SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

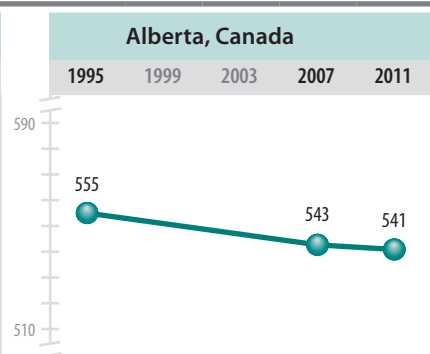
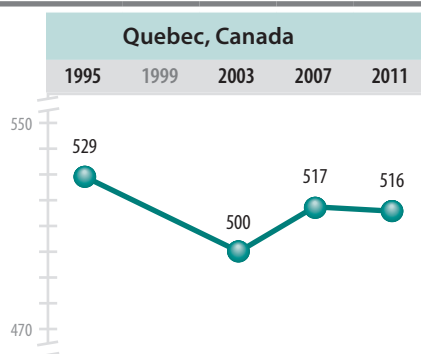
* No fourth-grade assessment in 1999.

Scale interval is 10 points for each country, but the part of the scale shown differs according to each country's average achievement.

Exhibit 1.7: Trends in Science Achievement – 1995 Through 2011* (Continued)



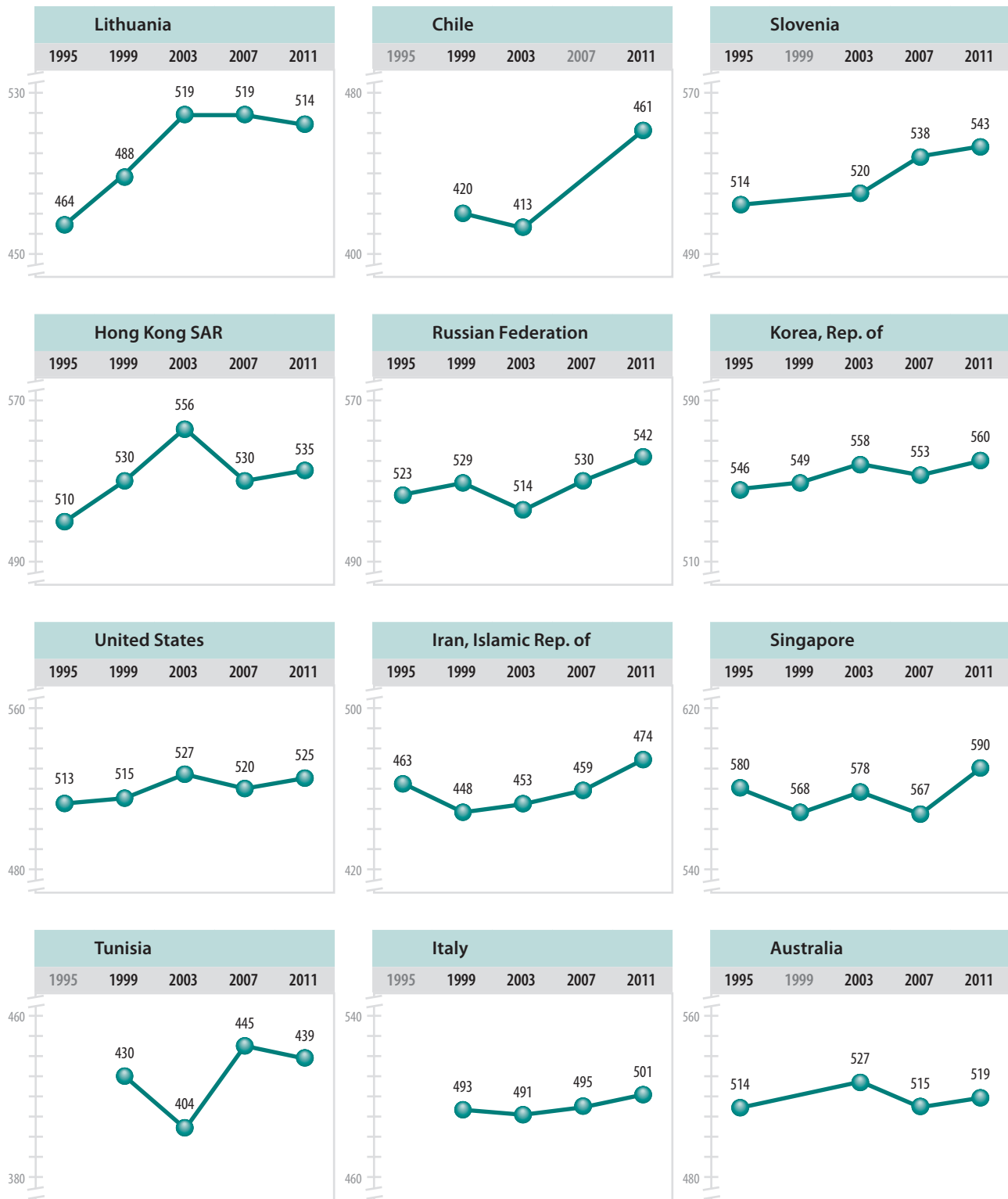
Benchmarking Participants



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 1.8: Trends in Science Achievement – 1995 Through 2011

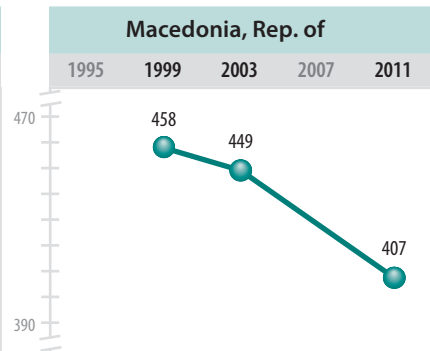
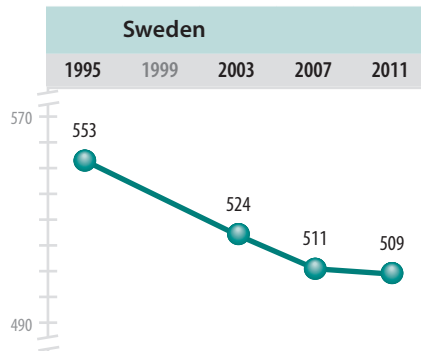
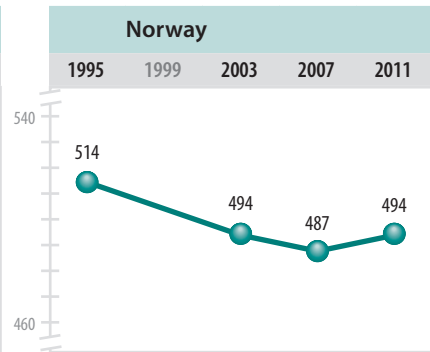
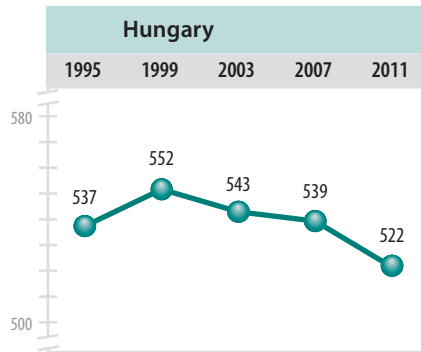
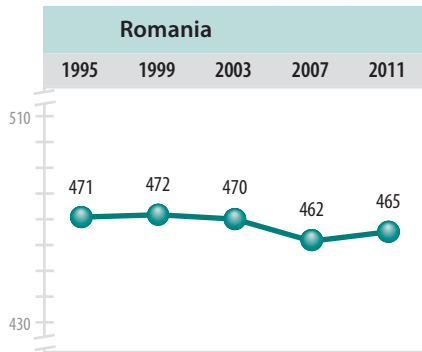
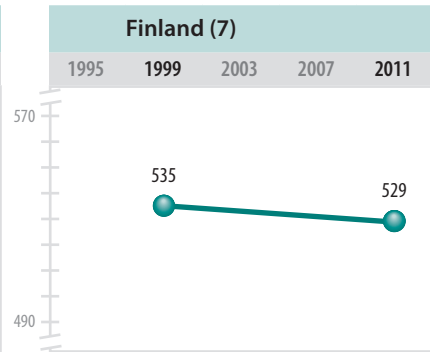
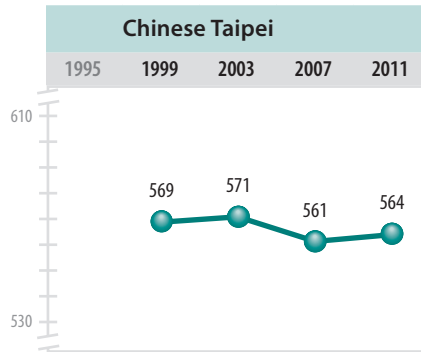
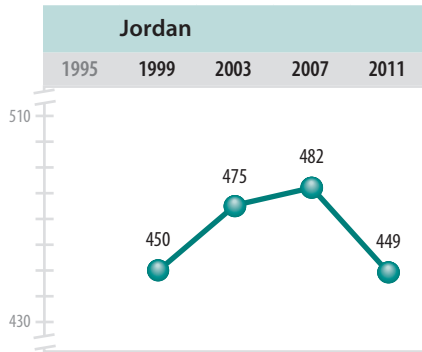
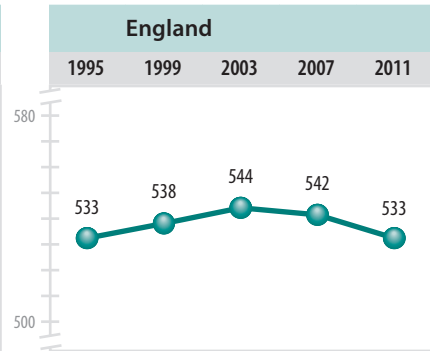
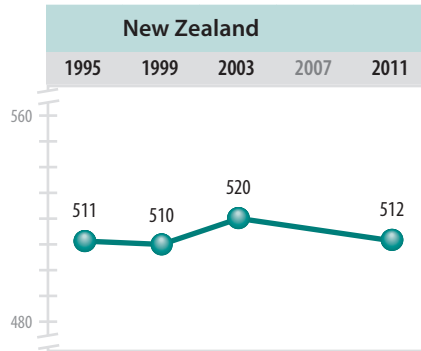
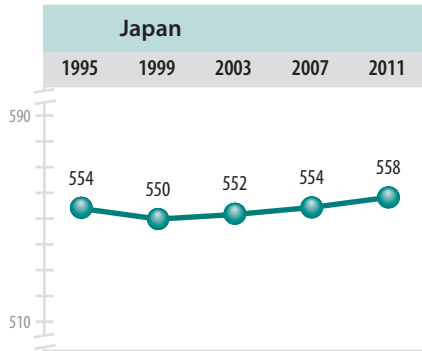
Includes only 2011 participants with comparable long term trend data beginning in either 1995 or 1999, ordered by most to least improvement in average achievement. Exhibit 1.6 provides details including statistical significance.



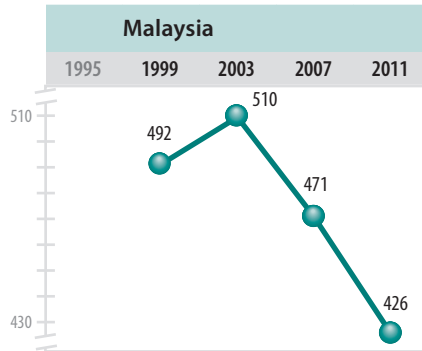
SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Scale interval is 10 points for each country, but the part of the scale shown differs according to each country's average achievement.

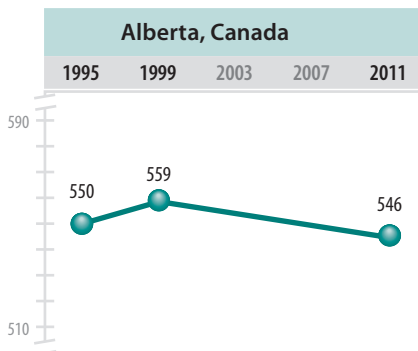
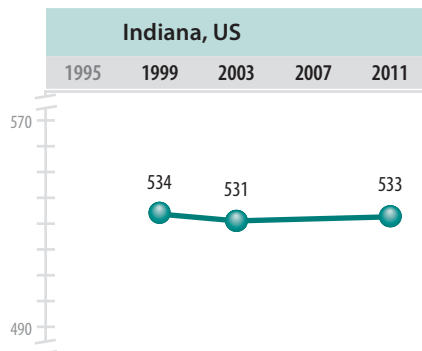
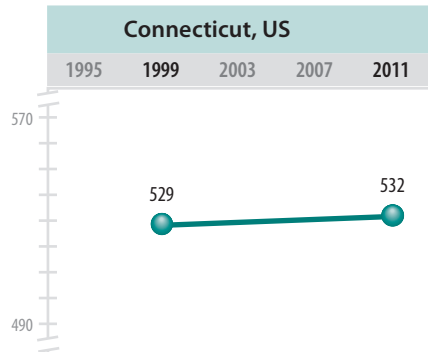
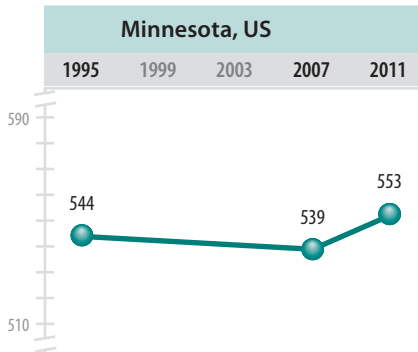
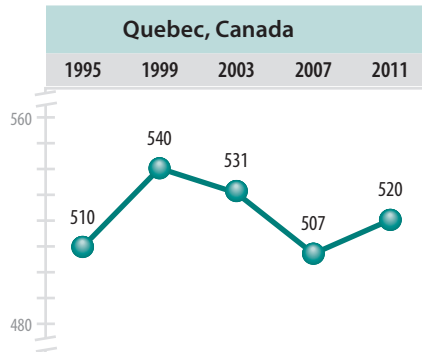
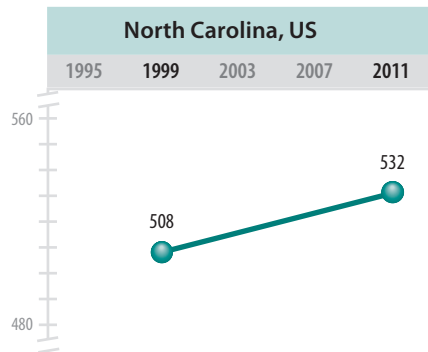
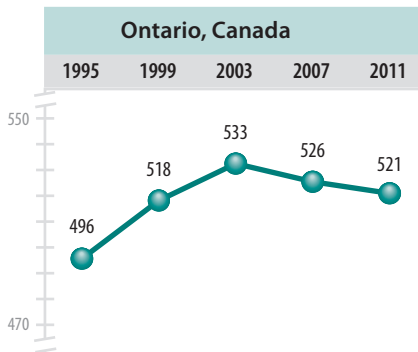
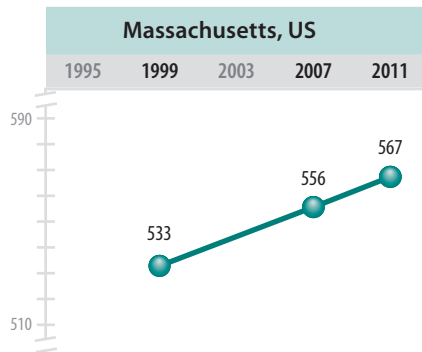
Exhibit 1.8: Trends in Science Achievement – 1995 Through 2011 (Continued)



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011



Benchmarking Participants



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Trends Across Grades: Fourth to Eighth Grade Cohort Analysis

Because TIMSS is conducted on a four-year cycle, the cohort of students that was assessed in the fourth grade in 2007 had reached the eighth grade by 2011, and thus was assessed at the eighth grade in 2011. This enables the 17 countries and three benchmarking participants that assessed both grades in both assessment years to examine how their performance relative to each other changed as the fourth grade students of 2007 became the eighth grade students of 2011. The results are presented in Exhibit 1.9, which shows average science achievement as a difference from the TIMSS scale centerpoint (500) for the fourth grade students in 2007 (upper-left panel) and in 2011 (upper-right panel). The exhibit also shows achievement for the eighth grade students in 2007 (lower-left panel) and in 2011 (lower-right panel). The trends for the fourth and eighth grades (indicated by the gray horizontal arrows), however, were presented more fully in Exhibits 1.5 and 1.6, respectively. The purpose of Exhibit 1.9 is to provide information about relative progress across grades as the cohort of students assessed at the fourth grade in 2007 moved to the eighth grade four years later in 2011. That is, to compare relative performance at the fourth grade in 2007 (upper-left panel) to relative performance at the eighth grade in 2011 (lower-right panel) as indicated by the darker arrow pointing diagonally downward.

Twelve countries—Singapore, Chinese Taipei, Hong Kong SAR, Japan, the Russian Federation, England, the United States, Hungary, Australia, Sweden, Slovenia, and Lithuania—and the Canadian provinces of Ontario and Québec performed above the scale centerpoint at the fourth grade in 2007 and again at the eighth grade in 2011 (although not in the same order of average achievement). Slovenia showed a particularly notable increase, with average achievement moving from 18 points above the centerpoint at the fourth grade in 2007 to 43 points above the centerpoint at the eighth grade in 2011. Norway, Iran, Georgia, Tunisia, and the UAE emirate of Dubai performed below the scale centerpoint in the fourth grade in 2007 and again at the eighth grade in 2011. Italy moved from above the centerpoint in the fourth grade in 2007 to the centerpoint at the eighth grade in 2011.

Gender Differences in Science Achievement

Previous TIMSS assessments have shown gender differences in science achievement to be smaller on average at the fourth grade than at the eighth grade, although the situation varies considerably from country to country.

Exhibit 1.9: Relative Achievement of 2007 Fourth Grade Cohort as Eighth Grade Students in 2011

2007 – Fourth Grade		
Country	Achievement Difference from TIMSS Scale Centerpoint (500)	
Singapore	87 (4.1)	▲
Chinese Taipei	57 (2.0)	▲
Hong Kong SAR	54 (3.5)	▲
Japan	48 (2.1)	▲
Russian Federation	46 (4.8)	▲
England	42 (2.9)	▲
United States	39 (2.7)	▲
Hungary	36 (3.3)	▲
Italy	35 (3.2)	▲
Australia	27 (3.3)	▲
Sweden	25 (2.9)	▲
Slovenia	18 (1.9)	▲
Lithuania	14 (2.4)	▲
Norway	-23 (3.5)	▼
Iran, Islamic Rep. of	-64 (4.3)	▼
Georgia	-82 (4.6)	▼
Tunisia	-182 (5.9)	▼
Benchmarking Participants		
Ontario, Canada	36 (3.7)	▲
Quebec, Canada	17 (2.7)	▲
Dubai, UAE	-40 (2.8)	▼

2011 – Fourth Grade		
Country	Achievement Difference from TIMSS Scale Centerpoint (500)	
Singapore	83 (3.4)	▲
Japan	59 (1.9)	▲
Russian Federation	52 (3.5)	▲
Chinese Taipei	52 (2.2)	▲
United States	44 (2.1)	▲
Hong Kong SAR	35 (3.8)	▲
Hungary	34 (3.7)	▲
Sweden	33 (2.7)	▲
England	29 (2.9)	▲
Italy	24 (2.7)	▲
Slovenia	20 (2.7)	▲
Australia	16 (2.8)	▲
Lithuania	15 (2.4)	▲
Norway	-6 (2.3)	▼
Georgia	-45 (3.8)	▼
Iran, Islamic Rep. of	-47 (3.7)	▼
Tunisia	-154 (5.3)	▼
Benchmarking Participants		
Ontario, Canada	28 (3.0)	▲
Quebec, Canada	16 (2.7)	▲
Dubai, UAE	-39 (2.3)	▼

2007 – Eighth Grade		
Country	Achievement Difference from TIMSS Scale Centerpoint (500)	
Singapore	67 (4.4)	▲
Chinese Taipei	61 (3.7)	▲
Japan	54 (1.9)	▲
England	42 (4.5)	▲
Hungary	39 (2.9)	▲
Slovenia	38 (2.2)	▲
Hong Kong SAR	30 (4.9)	▲
Russian Federation	30 (3.9)	▲
United States	20 (2.9)	▲
Lithuania	19 (2.5)	▲
Australia	15 (3.6)	▲
Sweden	11 (2.6)	▲
Italy	-5 (2.8)	
Norway	-13 (2.2)	▼
Iran, Islamic Rep. of	-41 (3.6)	▼
Tunisia	-55 (2.1)	▼
Georgia	-79 (4.8)	▼
Benchmarking Participants		
Ontario, Canada	26 (3.6)	▲
Quebec, Canada	7 (3.1)	▲
Dubai, UAE	-11 (2.8)	▼

2011 – Eighth Grade		
Country	Achievement Difference from TIMSS Scale Centerpoint (500)	
Singapore	90 (4.3)	▲
Chinese Taipei	64 (2.3)	▲
Japan	58 (2.4)	▲
Slovenia	43 (2.7)	▲
Russian Federation	42 (3.2)	▲
Hong Kong SAR	35 (3.4)	▲
England	33 (4.9)	▲
United States	25 (2.6)	▲
Hungary	22 (3.1)	▲
Australia	19 (4.8)	▲
Lithuania	14 (2.6)	▲
Sweden	9 (2.5)	▲
Italy	1 (2.5)	
Norway	-6 (2.6)	▼
Iran, Islamic Rep. of	-26 (4.0)	▼
Tunisia	-61 (2.5)	▼
Georgia	-80 (3.0)	▼
Benchmarking Participants		
Ontario, Canada	21 (2.5)	▲
Quebec, Canada	20 (2.5)	▲
Dubai, UAE	-15 (2.5)	▼

- ▲ Country average significantly higher than the centerpoint of the TIMSS scale
- ▼ Country average significantly lower than the centerpoint of the TIMSS scale

() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

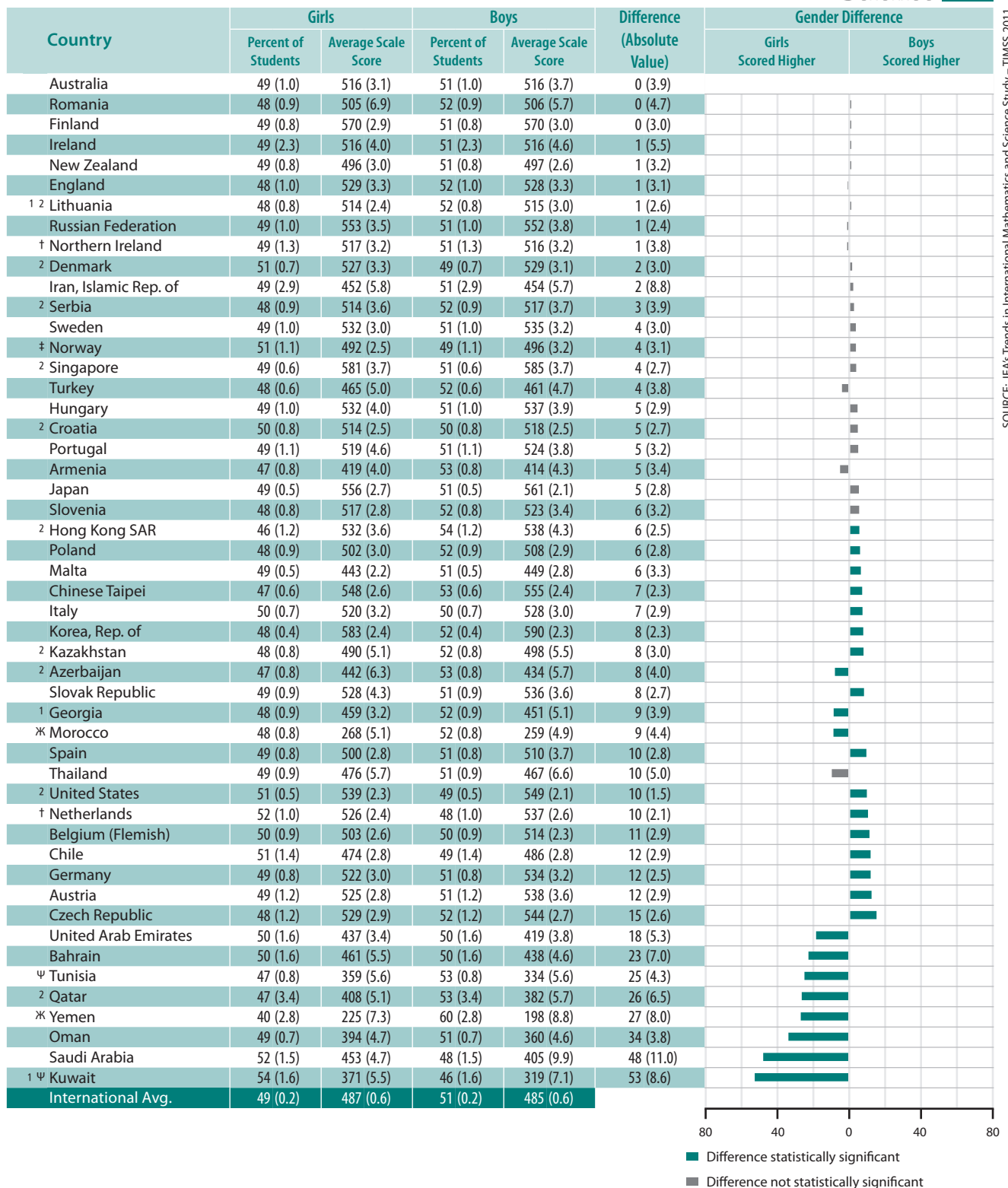
Exhibit 1.10 presents the TIMSS 2011 fourth grade results for gender differences in science achievement. For the TIMSS 2011 countries at the fourth grade, at sixth grade, and the benchmarking participants, it presents girls' average achievement, boys' average achievement, and the difference between the two averages. The bar graph shows the size of the achievement difference between boys and girls and whether that difference is statistically significant (as indicated by a darkened bar). International averages also are shown (averages across the mean scores for girls in each of the countries and the mean scores for boys in each of the countries). Exhibit 1.11 presents corresponding results for the TIMSS 2011 eighth grade assessment.

In each section of Exhibit 1.10, participants are shown in order by the increasing size of the difference between girls and boys in average science achievement. Overall, there was little achievement difference between girls and boys (International Average: 487 vs. 485). Of the 50 countries at the fourth grade, 23 had no significant gender difference in science achievement. Of the 27 remaining countries, 16 had relatively small differences favoring boys, and three had relatively small differences favoring girls. Eight countries had relatively larger differences favoring girls (the United Arab Emirates, Bahrain, Tunisia, Qatar, Yemen, Oman, Saudi Arabia, and Kuwait).

At the sixth grade, there was a significant achievement difference favoring girls in Botswana. Among benchmarking participants, boys had higher average science achievement than girls in Québec, Alberta, Florida, and North Carolina. Girls had higher achievement than boys in Abu Dhabi, UAE.

As shown in Exhibit 1.11, gender differences in science achievement at the eighth grade were larger, on average, than at the fourth grade, with the difference favoring girls (International Average: 480 vs. 474). Similar to the fourth grade, the gender difference varied across countries, with no difference in 17 of the 42 eighth grade countries, a difference favoring boys in ten countries, and a difference favoring girls in the remaining 15 countries. As at the fourth grade, and consistent with findings from TIMSS 2007, the largest achievement differences favoring girls at the eighth grade were in Arabic-speaking countries from the Middle East (the United Arab Emirates, Qatar, Saudi Arabia, the Palestinian National Authority, Jordan, Bahrain, and Oman). Among countries that assessed their ninth grade students, girls had higher science achievement than boys in Botswana and boys had higher achievement than girls in Honduras. Among the 14 benchmarking participants, boys performed better than girls in six US states and the Canadian province of Alberta and girls performed better than boys (by a large margin) in Dubai, UAE.

Exhibit 1.10: Average Science Achievement by Gender



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

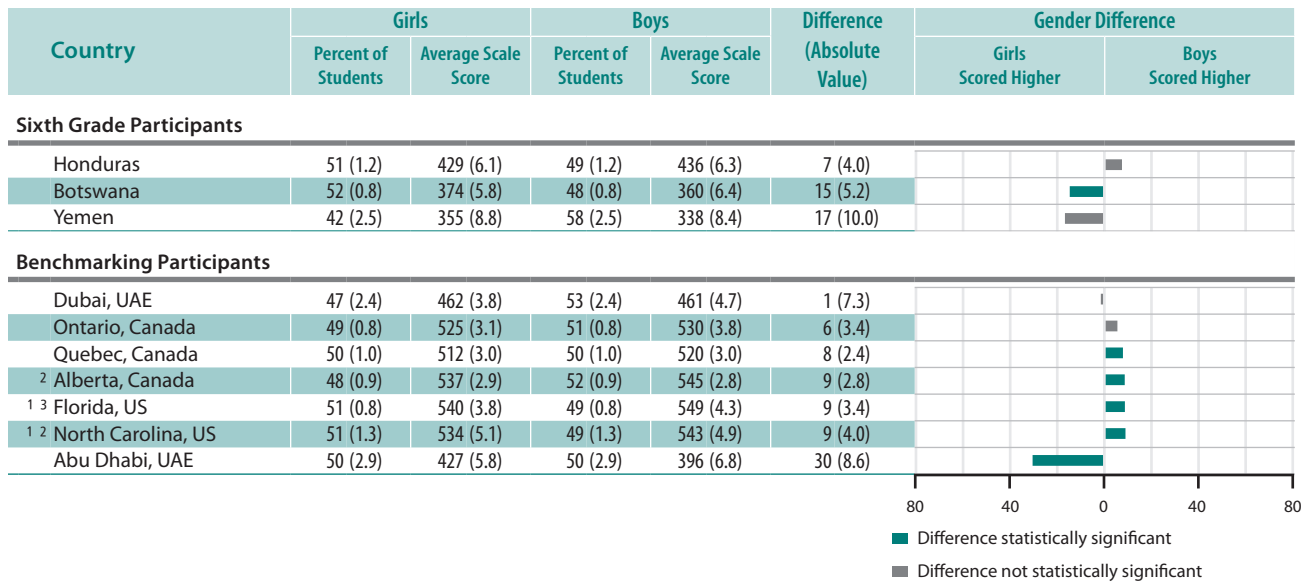
* Average achievement not reliably measured because the percentage of students with achievement too low for estimation exceeds 25%.

‡ Reservations about reliability of average achievement because the percentage of students with achievement too low for estimation does not exceed 25% but exceeds 15%.

See Appendix C.2 for target population coverage notes 1, 2, and 3. See Appendix C.8 for sampling guidelines and sampling participation notes †, ‡, and §.

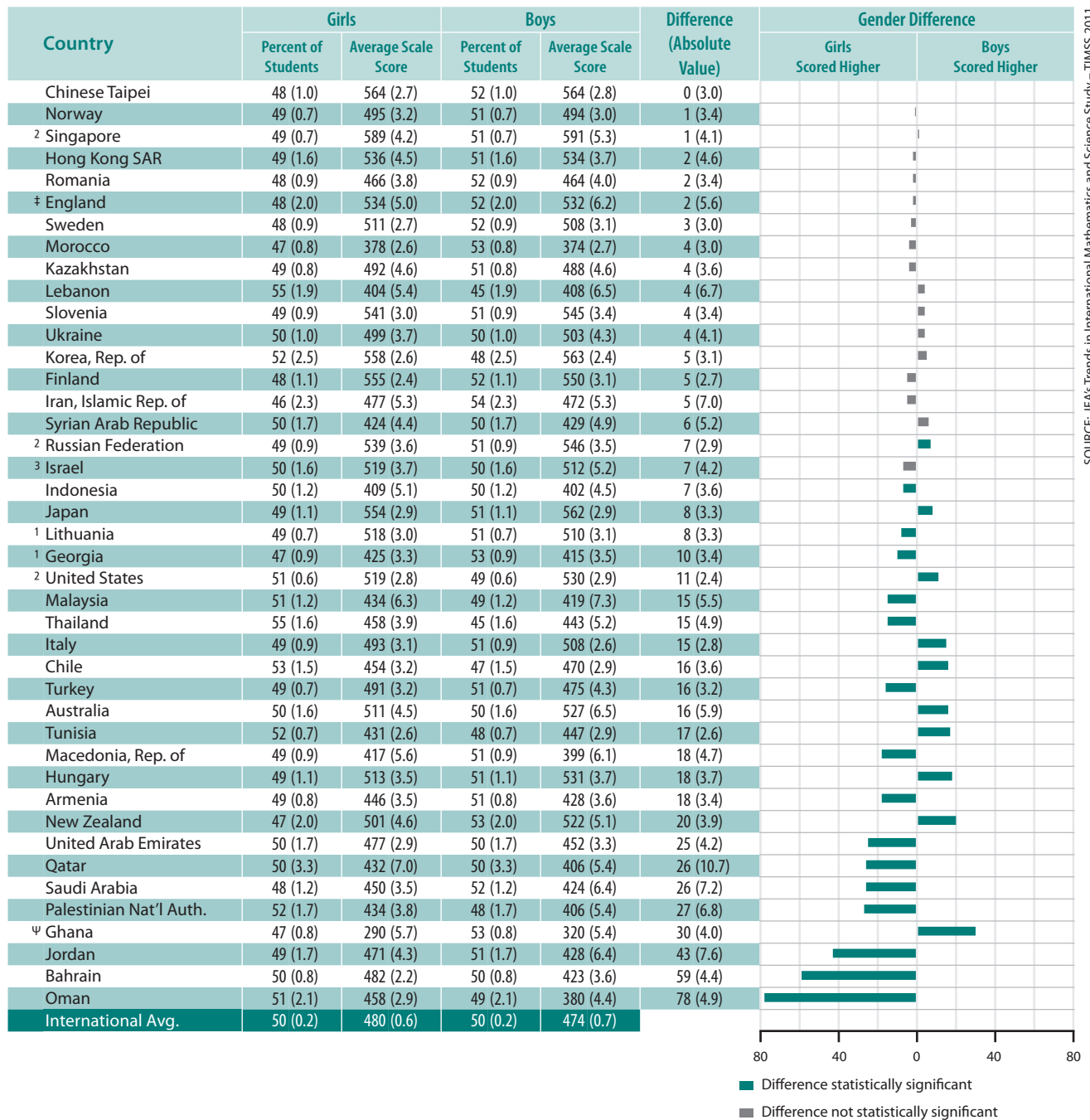
() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.10: Average Science Achievement by Gender (Continued)



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

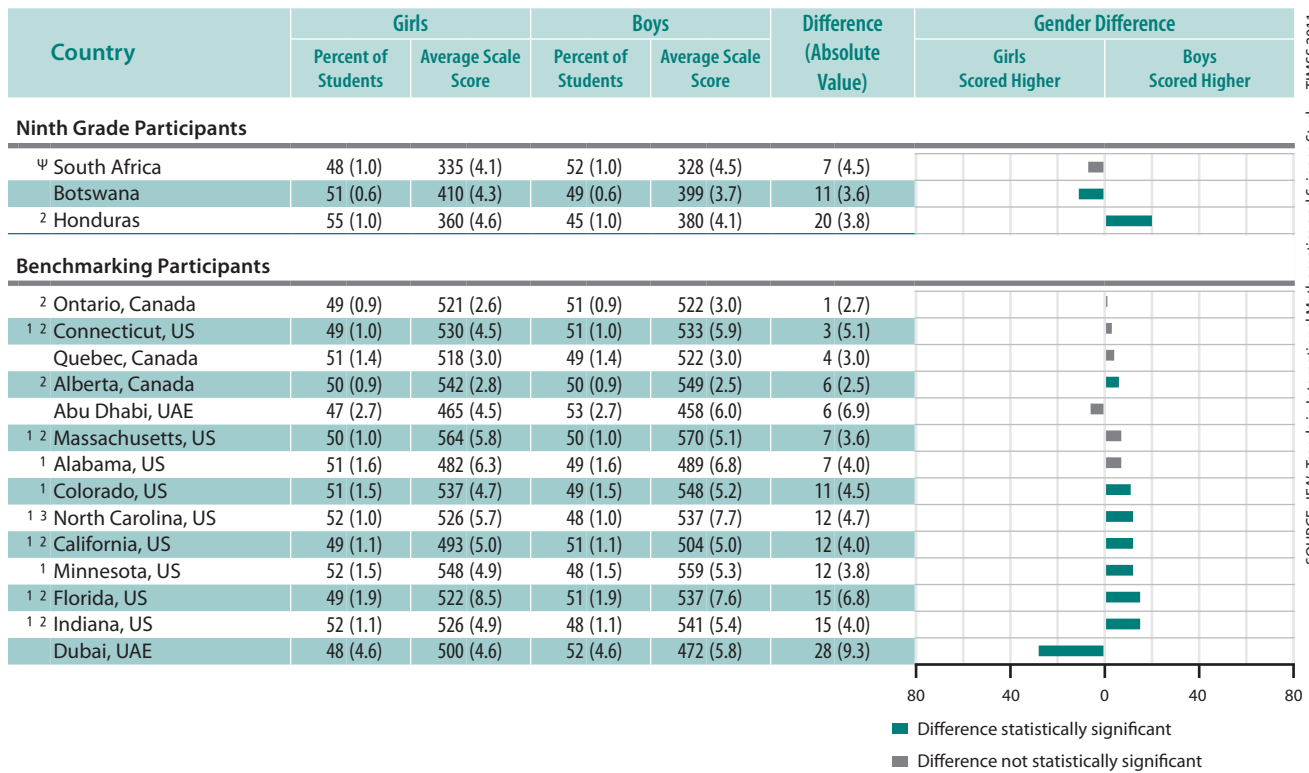
Exhibit 1.11: Average Science Achievement by Gender



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

^ψ Reservations about reliability of average achievement because the percentage of students with achievement too low for estimation does not exceed 25% but exceeds 15%. See Appendix C.3 for target population coverage notes 1, 2, and 3. See Appendix C.9 for sampling guidelines and sampling participation notes †, ‡, and §.
() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.11: Average Science Achievement by Gender (Continued)



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Trends in Science Achievement by Gender

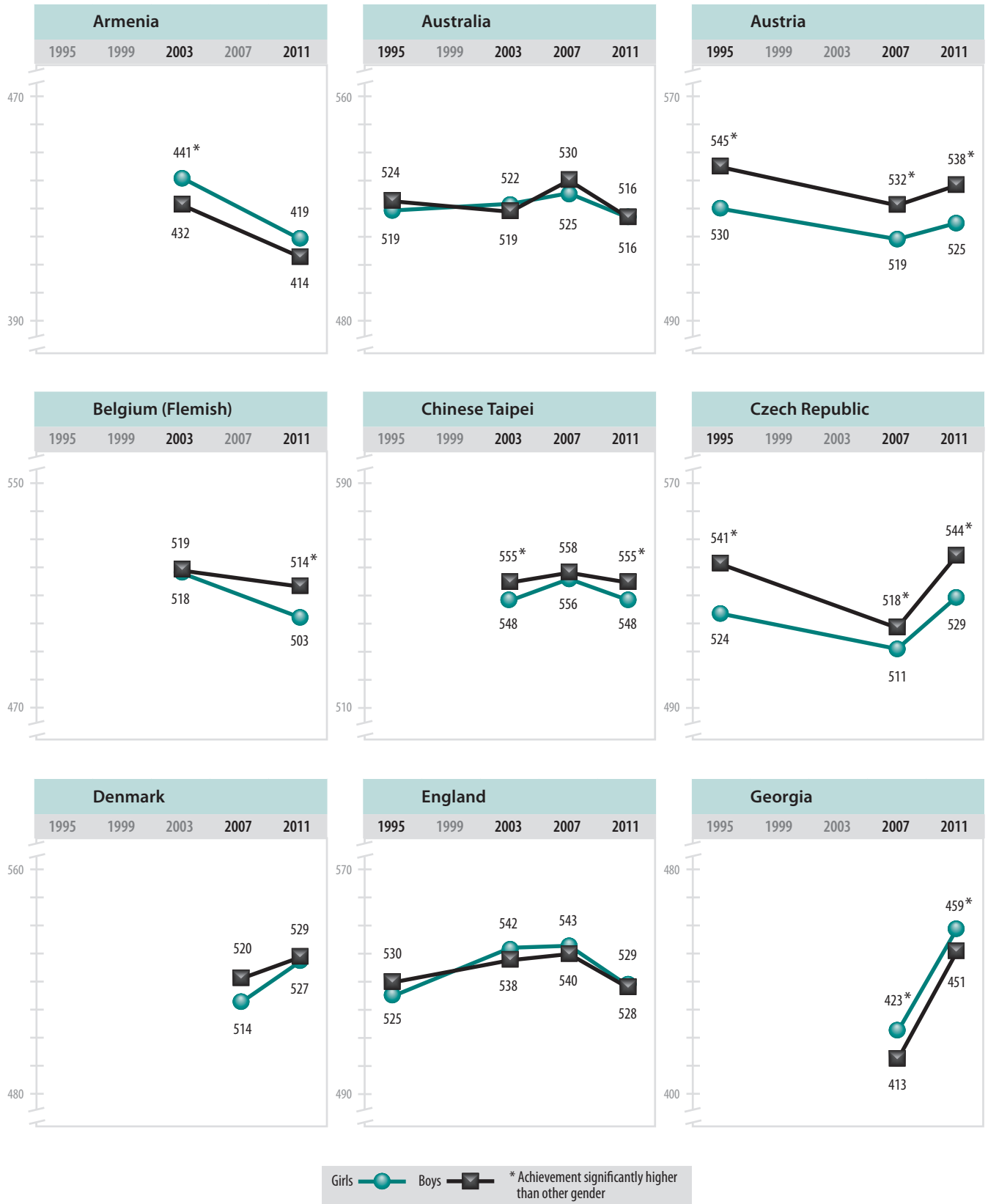
Exhibits 1.12 and 1.13 show graphic representations across the TIMSS assessments in science achievement of boys and girls for the fourth and eighth grades, respectively. For each country that participated in one or more of the previous TIMSS assessments, these displays show how trends in science achievement have been influenced by differential performance by boys and girls. Because there are many different patterns across countries, the countries are presented in alphabetical order. The scale interval is the same for each country (10 points) to permit comparisons, although the part of scale shown differs according to each country's average achievement. For countries with gender differences in science achievement, the displays reveal progress (or lack thereof) over time toward gender equity.

As described in the previous section, at the fourth grade, there is already gender equity in science achievement in many countries, but there are also countries where overall achievement is less than it might be if both boys and girls performed at the same high level. Countries in which fourth grade girls performed consistently below boys (i.e., in 2011 and on at least two other TIMSS assessments) include Austria, the Czech Republic, the Netherlands, and the United States. In Germany, Italy, and the Slovak Republic, boys had higher average achievement than girls on the two most recent assessments (2007 and 2011); while in Georgia and Tunisia, girls had higher achievement than boys on these two assessments. Armenia, Hungary, Iran, Japan, New Zealand, Norway, and Slovenia had gender differences in earlier assessments but not in TIMSS 2011.

At the eighth grade, with greater gender differences among countries than at the fourth grade, and data from across five TIMSS assessments, trends in science achievement for boys and girls follow a variety of paths. Boys consistently had higher achievement than girls (i.e., in 2011 and in one or more previous assessments) in nine countries: Australia, Chile, Ghana, Hungary, Italy, Japan, the Russian Federation, Tunisia, and the United States. A similar pattern of boys outperforming girls also occurred in four benchmarking participants: the states of Indiana, Minnesota, and North Carolina, and the province of Alberta. Conversely, in eight countries—Armenia, Bahrain, Georgia, Jordan, Macedonia, Oman, the Palestinian National Authority, and Thailand—girls had higher achievement than boys in 2011 and in one or more prior assessments; and twelve countries and four benchmarking participants showed no gender difference in 2011, despite having gender differences on one or more previous assessments. Gender differences in average science achievement followed a different pattern in Lithuania and Malaysia; in these countries, girls outperformed boys in 2011 although boys had outperformed girls in at least one prior assessment.

Exhibit 1.12: Trends in Science Achievement by Gender[◇]

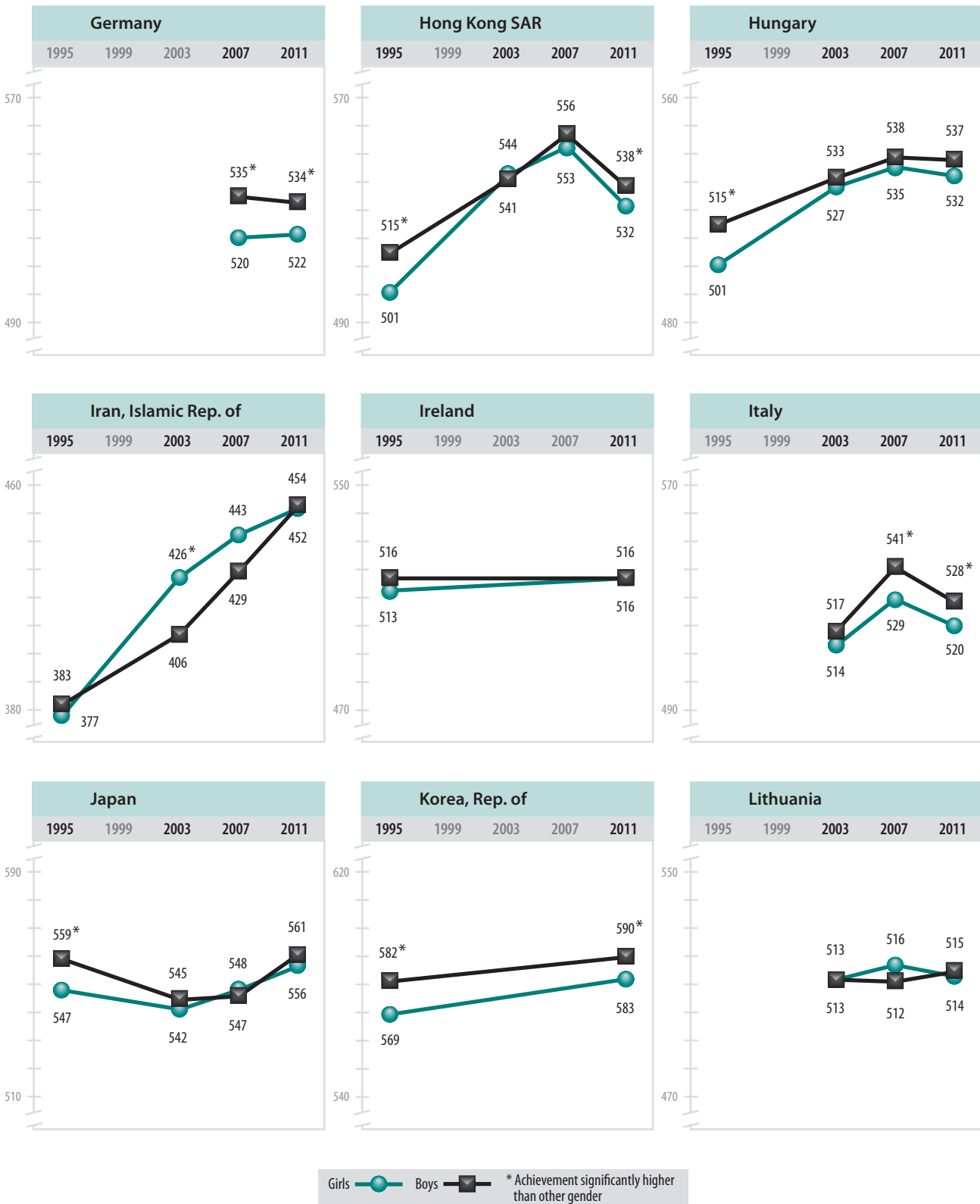
SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011



[◇] No fourth-grade assessment in 1999.

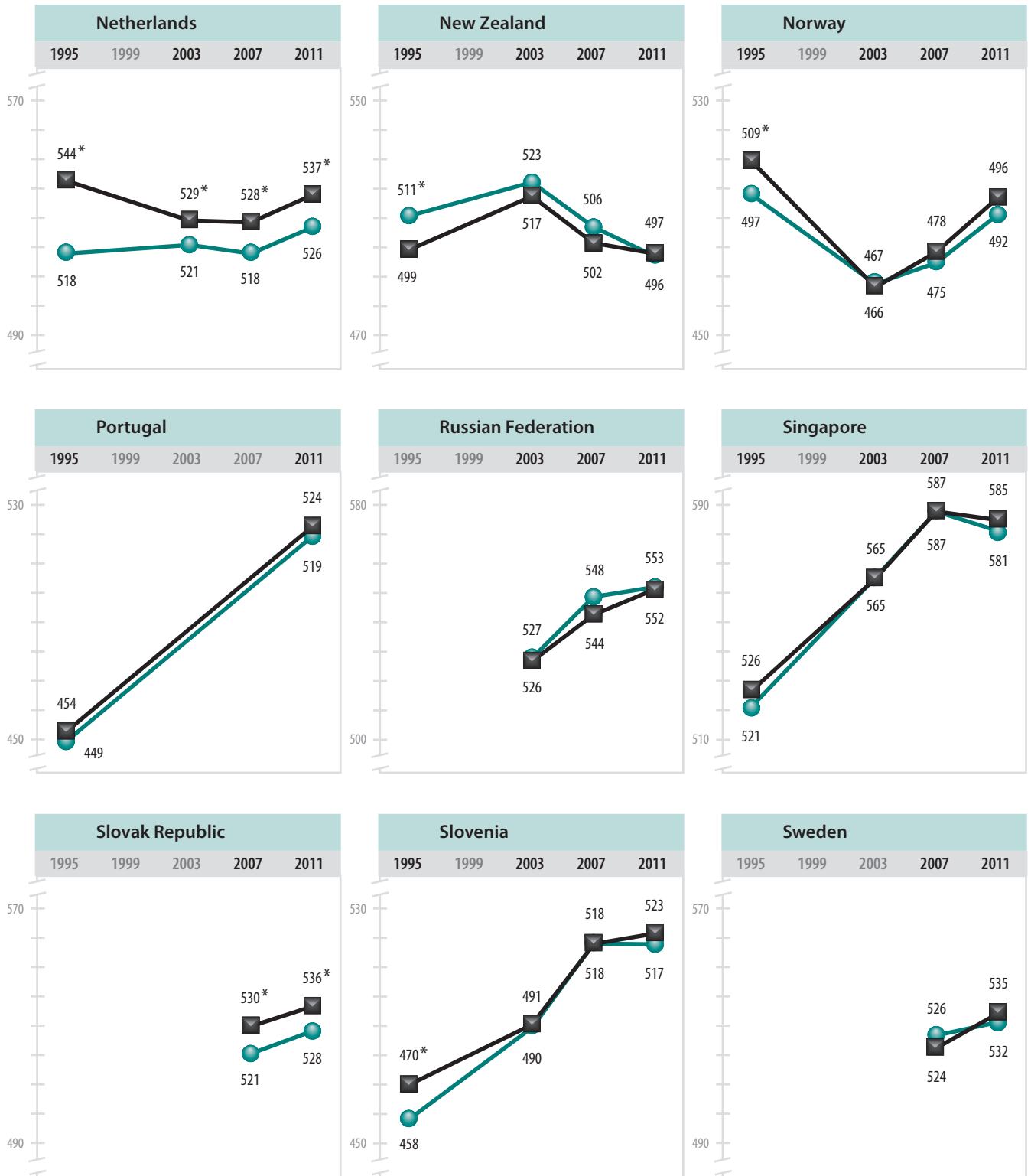
Scale interval is 10 points for each country, but the part of the scale shown differs according to each country's average achievement.

Exhibit 1.12: Trends in Science Achievement by Gender^o (Continued)



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 1.12: Trends in Science Achievement by Gender^o (Continued)

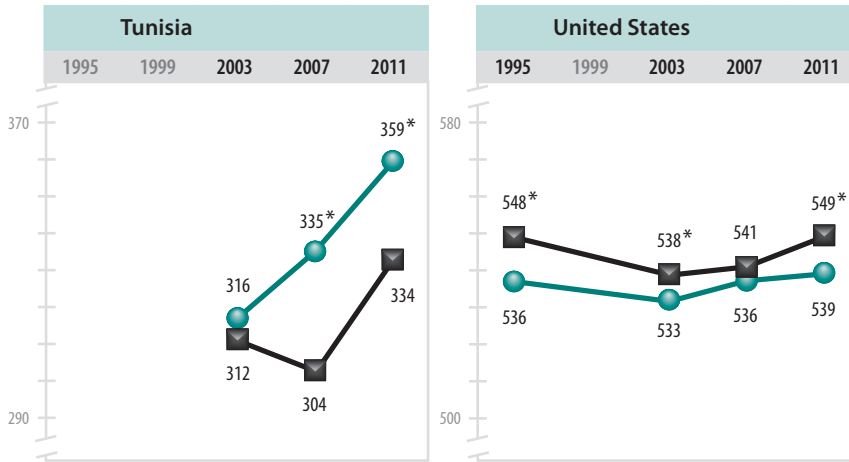


SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

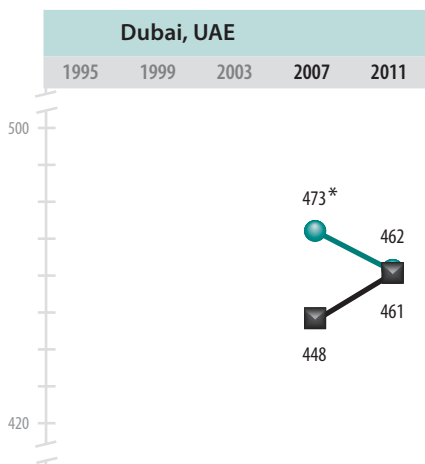
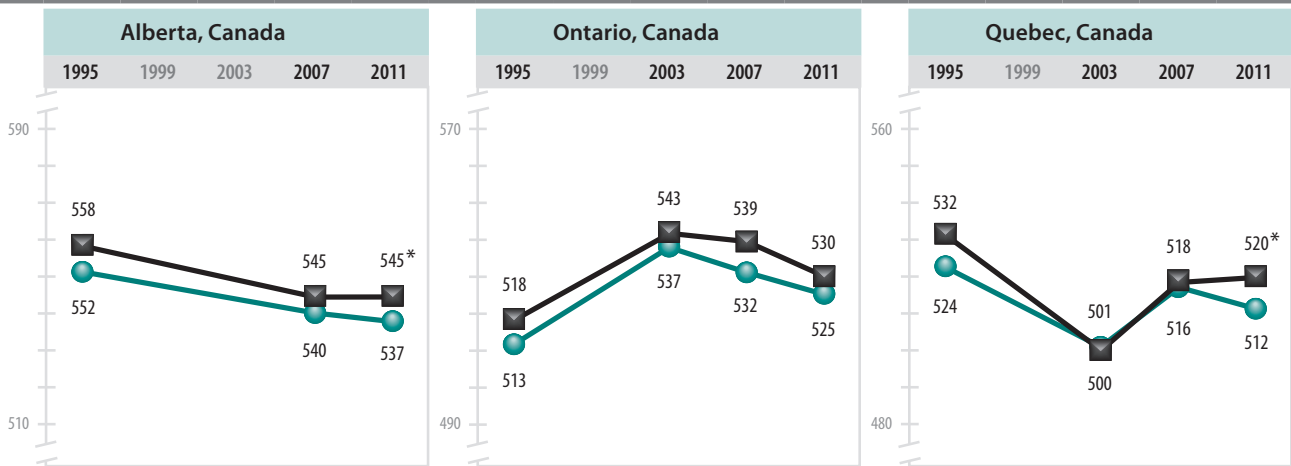
Girls ● Boys ■ * Achievement significantly higher than other gender

Exhibit 1.12: Trends in Science Achievement by Gender^a (Continued)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011



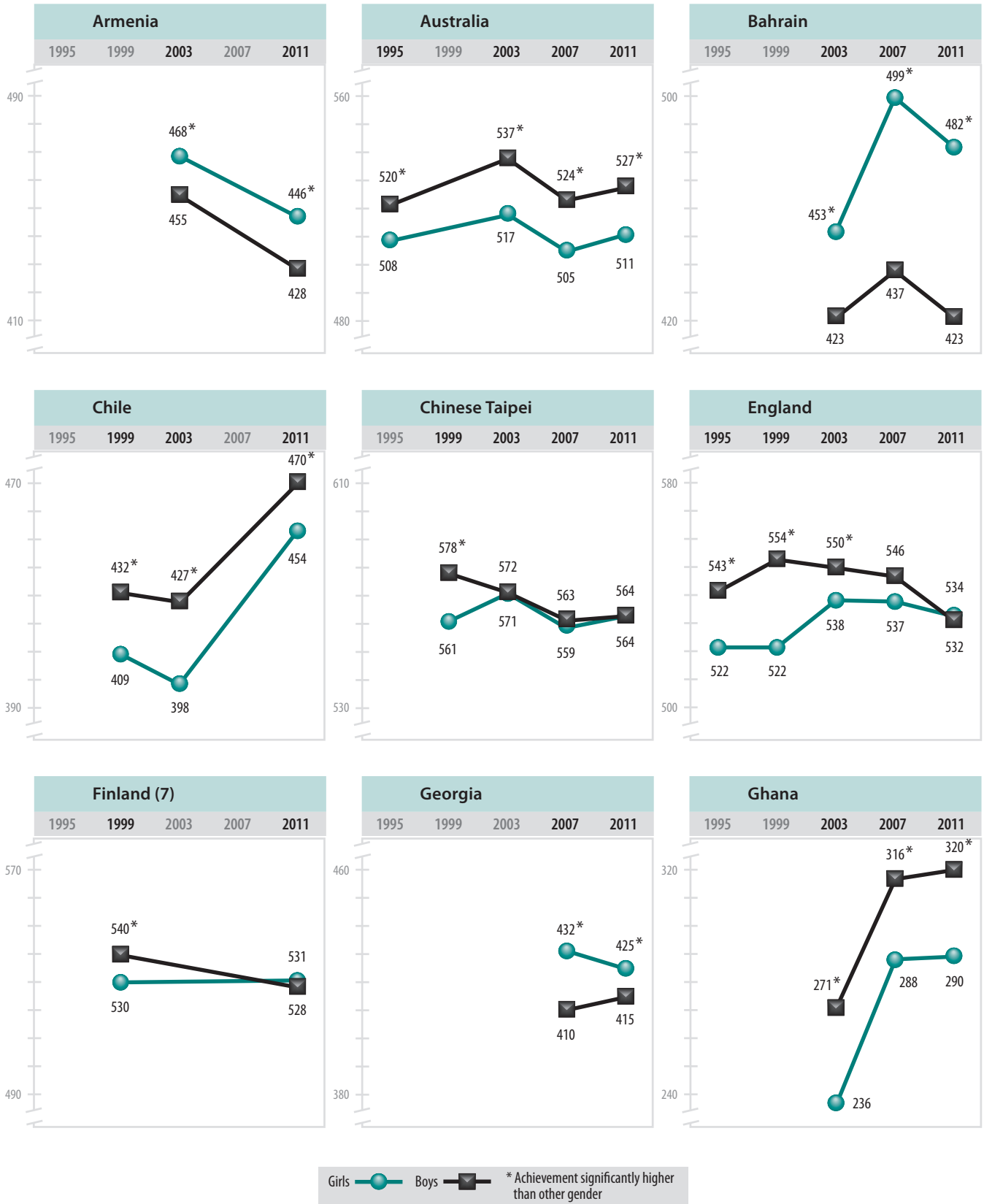
Benchmarking Participants



Girls —●— Boys —■— * Achievement significantly higher than other gender

Exhibit 1.13: Trends in Science Achievement by Gender

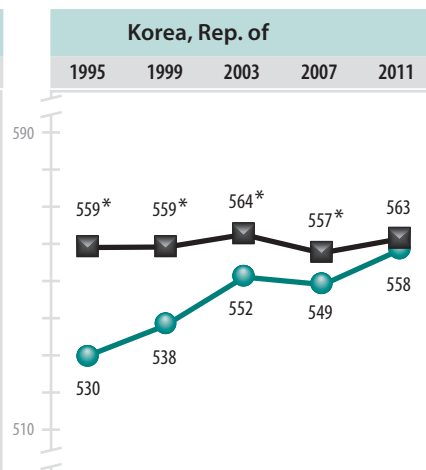
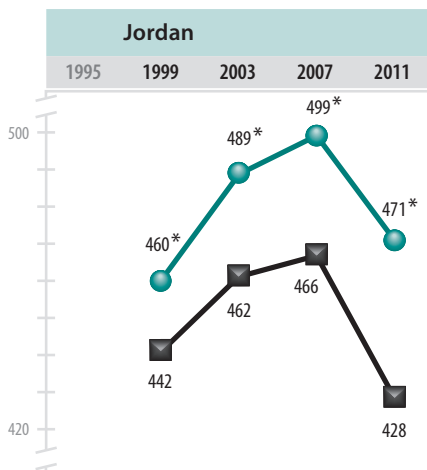
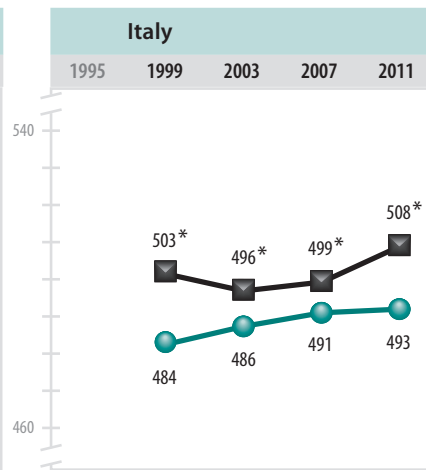
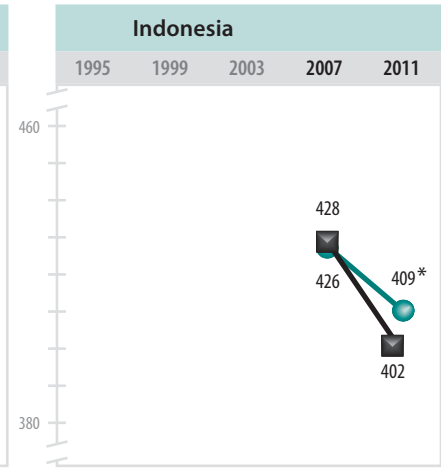
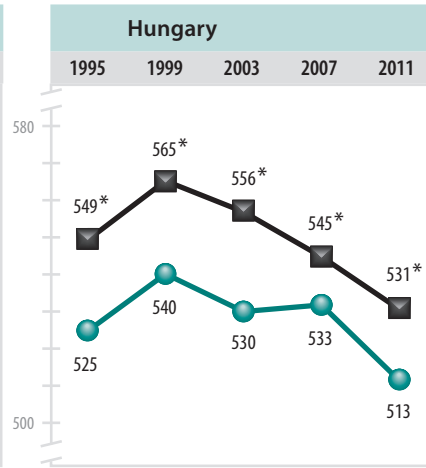
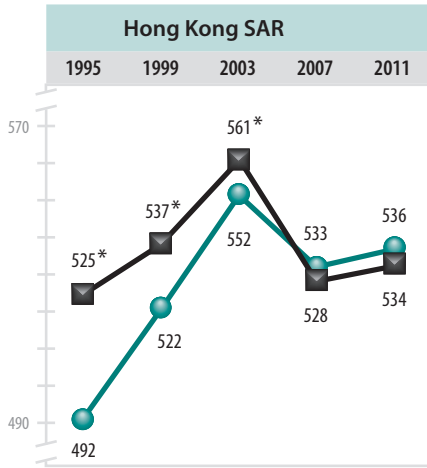
SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011



Scale interval is 10 points for each country, but the part of the scale shown differs according to each country's average achievement.

Exhibit 1.13: Trends in Science Achievement by Gender (Continued)

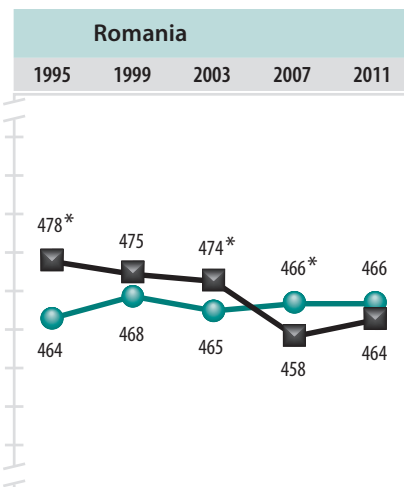
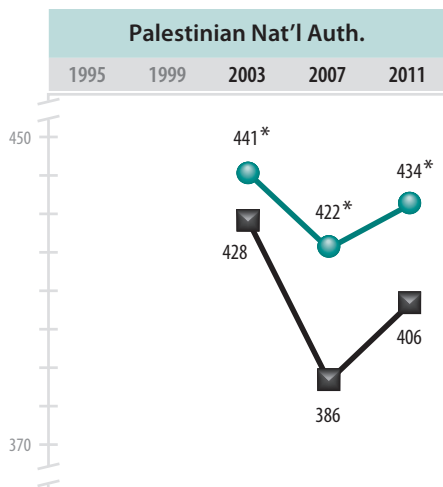
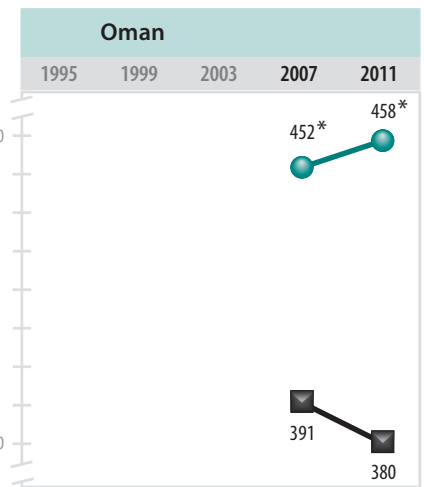
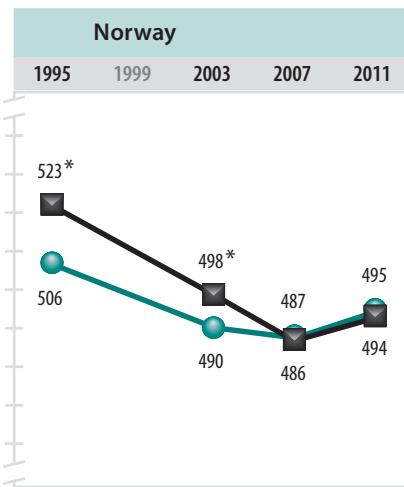
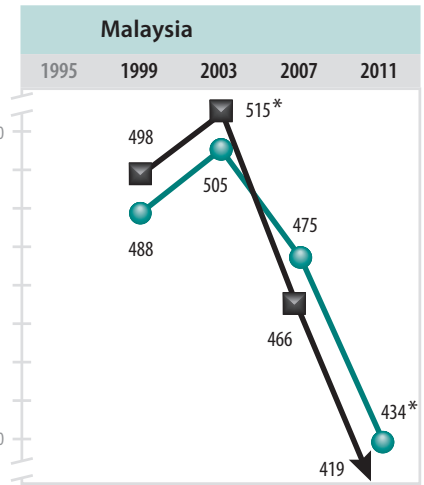
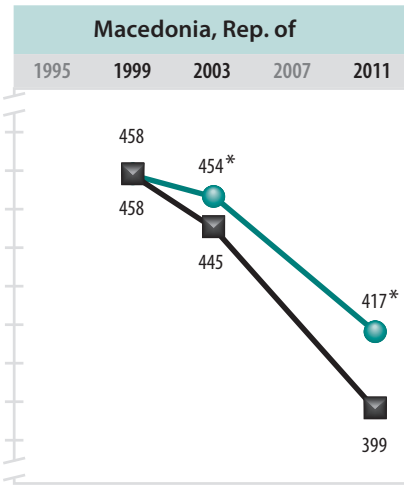
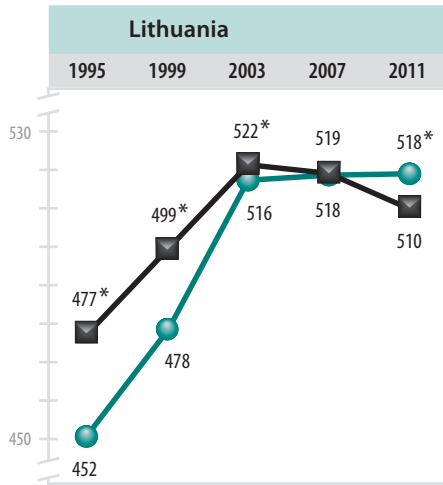
SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011



Girls —●— Boys —■— * Achievement significantly higher than other gender

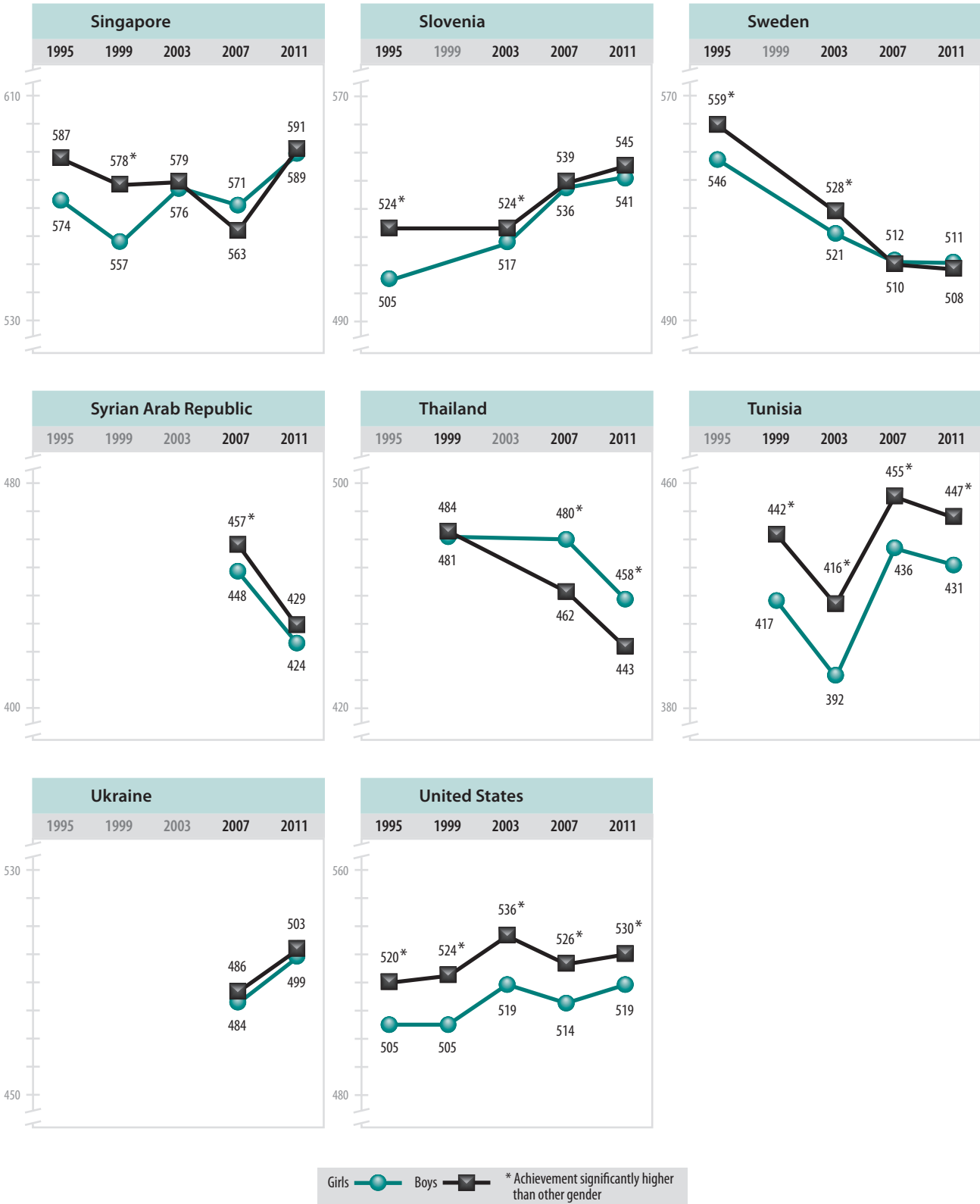
Exhibit 1.13: Trends in Science Achievement by Gender (Continued)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011



Girls —●— Boys —■— * Achievement significantly higher than other gender

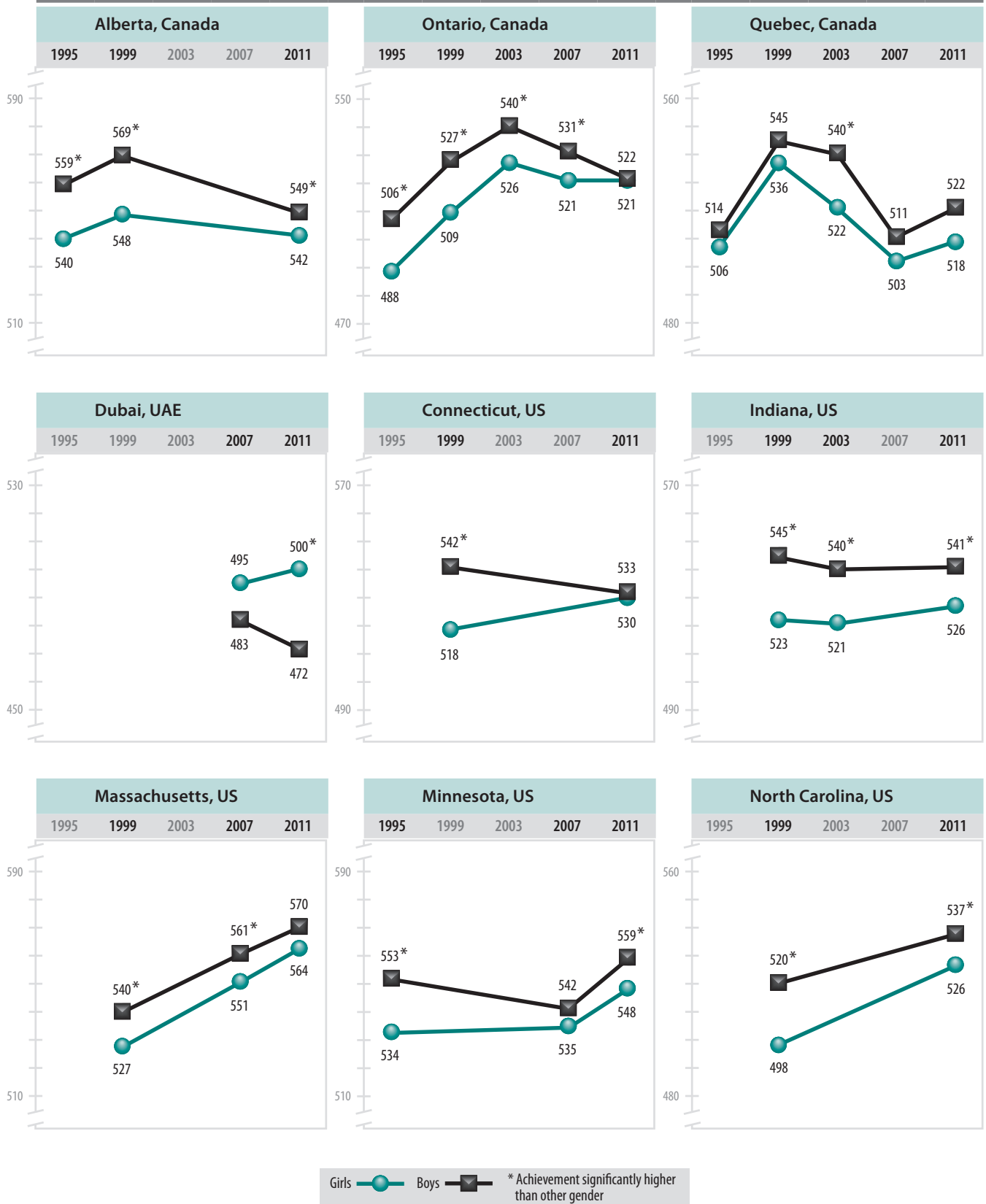
Exhibit 1.13: Trends in Science Achievement by Gender (Continued)



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

Exhibit 1.13: Trends in Science Achievement by Gender (Continued)

Benchmarking Participants



SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

