

Chapter 1

International Student Achievement in Mathematics

East Asian countries continue to lead the world in mathematics achievement. Singapore, Korea, and Hong Kong SAR, followed by Chinese Taipei and Japan, were the top-performing countries in TIMSS 2011 at the fourth grade. Similarly, at the eighth grade, Korea, Singapore, and Chinese Taipei outperformed all countries, followed by Hong Kong SAR and Japan.

Since 1995, fourth grade students have shown more improvement than reduction in mathematics achievement (12 countries up vs. only 3 down), but improving eighth grade student achievement has been more difficult (9 up vs. 11 down).

Chapter 1 contains the mathematics 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 mathematics achievement across the participants at fourth and eighth grades, the chapter provides:

- ◆ Averages (means) and distributions of mathematics achievement;
- ◆ Trends in mathematics achievement over time for participants in previous TIMSS assessments in 1995, 1999, 2003 and 2007;
- ◆ Trends across grades—relative achievement of 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.

Mathematics Achievement Across Countries

TIMSS 2011 Mathematics Achievement

This section reports the TIMSS 2011 mathematics 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 mathematics 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 distribution

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

of mathematics 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. 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 several 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. Three countries elected to assess sixth grade students, including 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 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, for this assessment, 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 24 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.

Singapore, Korea, and Hong Kong SAR 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 three countries performed similarly and had higher achievement than all of the other countries. The next highest-performing country was Chinese Taipei, which had higher achievement than all countries except the three with the highest achievement, followed by Japan, which had average achievement higher than all countries except Chinese Taipei and the three top performers. Also included in the top ten high-achieving countries were Northern Ireland, Belgium (Flemish), Finland, England, and the Russian Federation. The benchmarking states of Florida and North Carolina had performance similar to these countries.

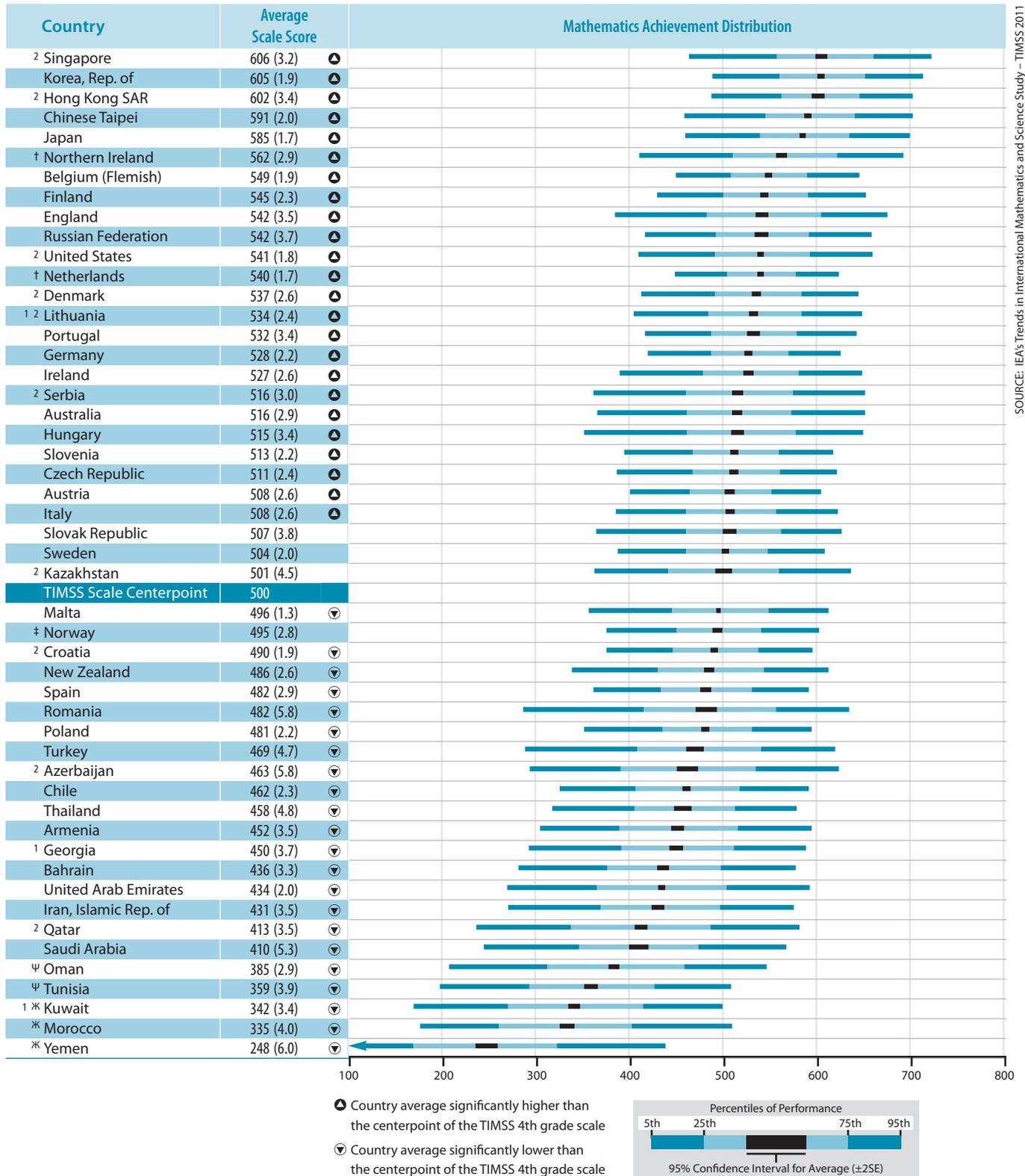
While there were small differences from country to country, there was a substantial range in performance from the top-performing to the lower-performing countries. Twenty-two countries had average achievement below the TIMSS centerpoint of 500. For the most part, these countries had average achievement 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 mathematics achievement, measurement is most efficient when there is a reasonable match between the mathematics 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 achievement. If there are substantial numbers of students with very low scores, their achievement is likely to be overestimated and consequently the overall

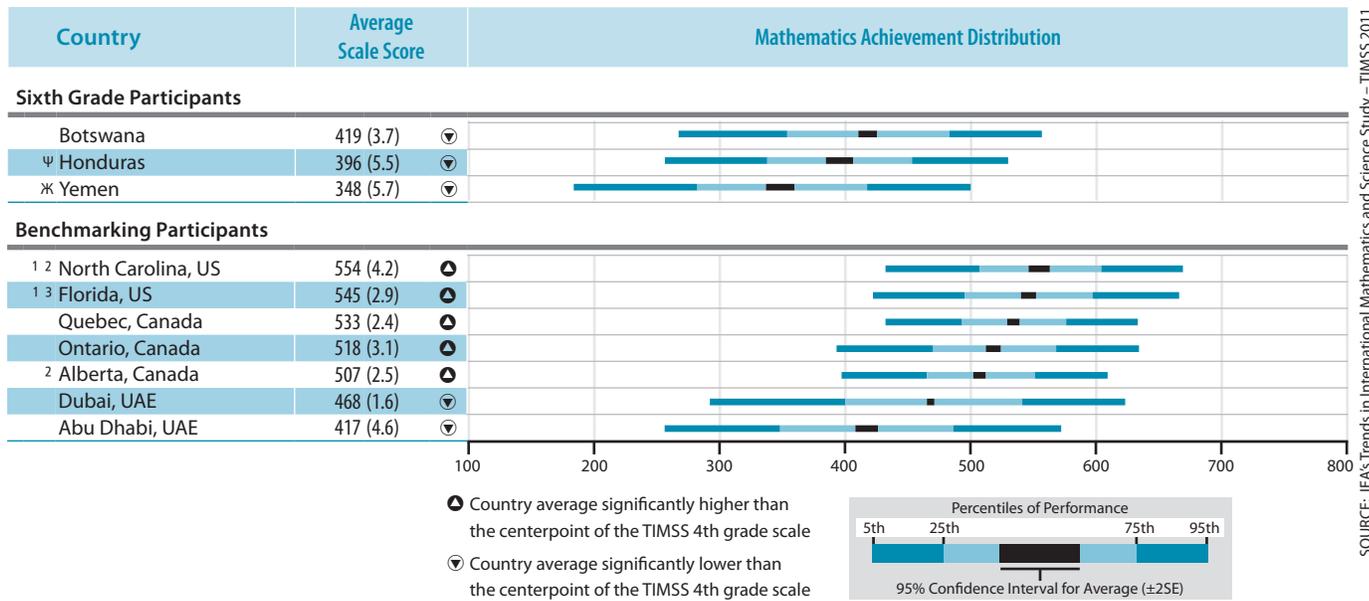
Exhibit 1.1: Distribution of Mathematics Achievement



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

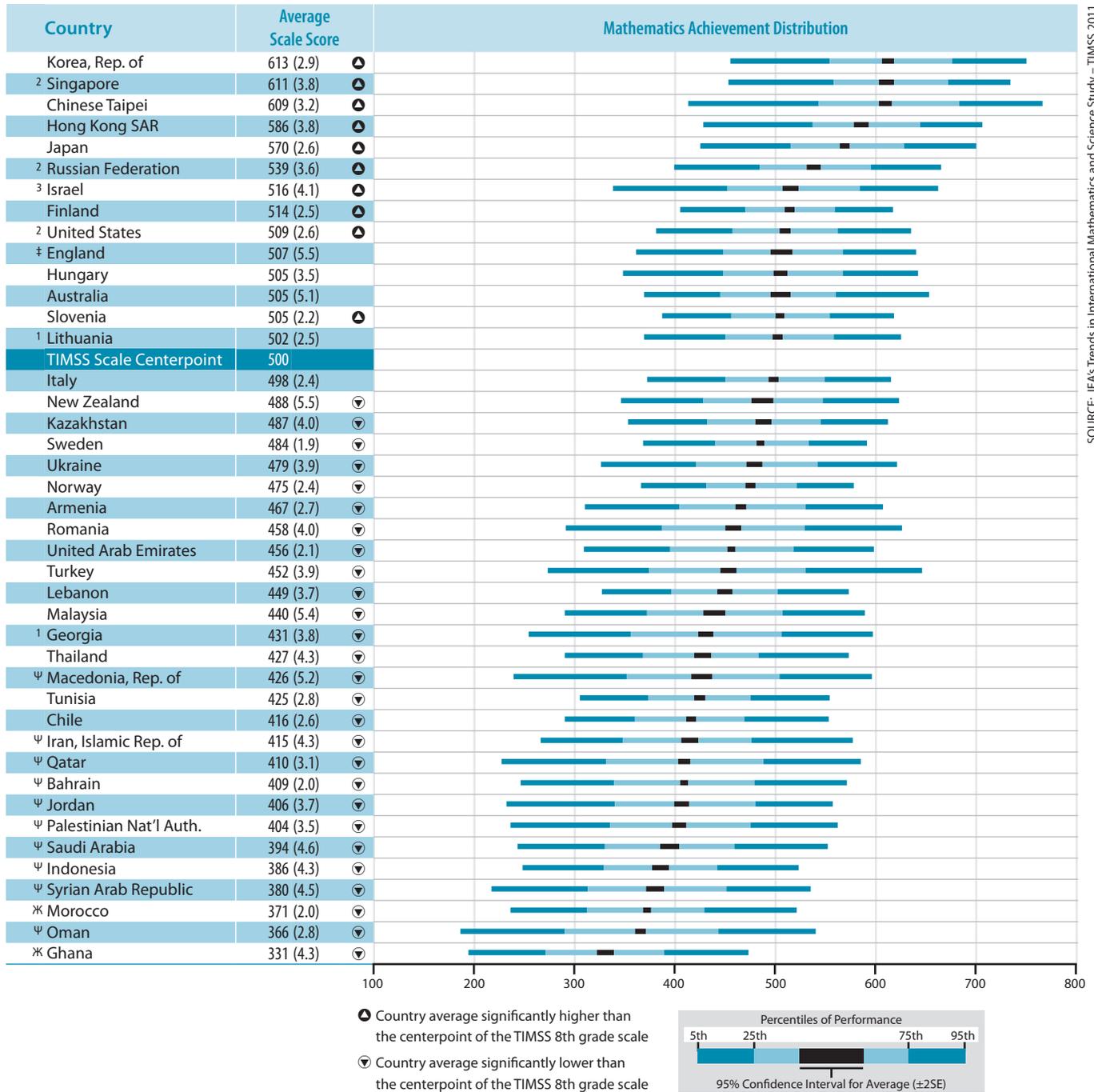
* Average achievement not reliably measured because the percentage of students with achievement too low for estimation exceeds 25%.
Ψ Reservations about reliability of average achievement because the percentage of students with achievement too low for estimation does not exceed 25% but exceeds 15%.
See 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 Mathematics Achievement (Continued)



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

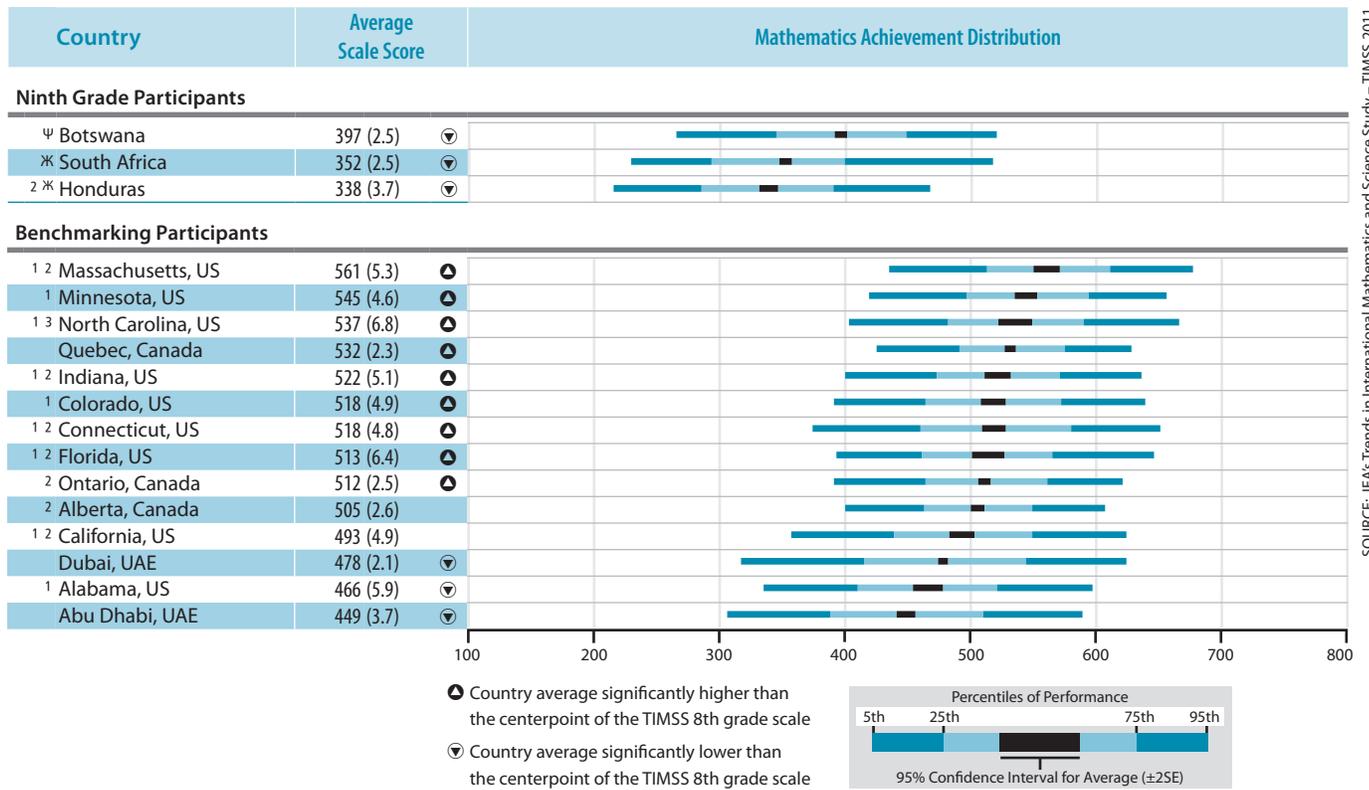
Exhibit 1.2: Distribution of Mathematics Achievement



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.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 Mathematics Achievement (Continued)



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

Exhibit 1.4: Multiple Comparisons of Average Mathematics 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	Chinese Taipei	Hong Kong SAR	Japan	Russian Federation	Israel	Finland	United States	England	Hungary	Australia	Slovenia	Lithuania	Italy	New Zealand	Kazakhstan	Sweden	Ukraine	Norway	Armenia	Romania	United Arab Emirates	Turkey	Lebanon	Malaysia	Georgia	Thailand	Macedonia, Rep. of	Tunisia
Korea, Rep. of	613 (2.9)																														
Singapore	611 (3.8)																														
Chinese Taipei	609 (3.2)																														
Hong Kong SAR	586 (3.8)																														
Japan	570 (2.6)																														
Russian Federation	539 (3.6)																														
Israel	516 (4.1)																														
Finland	514 (2.5)																														
United States	509 (2.6)																														
England	507 (5.5)																														
Hungary	505 (3.5)																														
Australia	505 (5.1)																														
Slovenia	505 (2.2)																														
Lithuania	502 (2.5)																														
Italy	498 (2.4)																														
New Zealand	488 (5.5)																														
Kazakhstan	487 (4.0)																														
Sweden	484 (1.9)																														
Ukraine	479 (3.9)																														
Norway	475 (2.4)																														
Armenia	467 (2.7)																														
Romania	458 (4.0)																														
United Arab Emirates	456 (2.1)																														
Turkey	452 (3.9)																														
Lebanon	449 (3.7)																														
Malaysia	440 (5.4)																														
Georgia	431 (3.8)																														
Thailand	427 (4.3)																														
Macedonia, Rep. of	426 (5.2)																														
Tunisia	425 (2.8)																														
Chile	416 (2.6)																														
Iran, Islamic Rep. of	415 (4.3)																														
Qatar	410 (3.1)																														
Bahrain	409 (2.0)																														
Jordan	406 (3.7)																														
Palestinian Nat'l Auth.	404 (3.5)																														
Saudi Arabia	394 (4.6)																														
Indonesia	386 (4.3)																														
Syrian Arab Republic	380 (4.5)																														
Morocco	371 (2.0)																														
Oman	366 (2.8)																														
Ghana	331 (4.3)																														
Botswana (9)	397 (2.5)																														
South Africa (9)	352 (2.5)																														
Honduras (9)	338 (3.7)																														
Benchmarking Participants																															
Massachusetts, US	561 (5.3)																														
Minnesota, US	545 (4.6)																														
North Carolina, US	537 (6.8)																														
Quebec, Canada	532 (2.3)																														
Indiana, US	522 (5.1)																														
Colorado, US	518 (4.9)																														
Connecticut, US	518 (4.8)																														
Florida, US	513 (6.4)																														
Ontario, Canada	512 (2.5)																														
Alberta, Canada	505 (2.6)																														
California, US	493 (4.9)																														
Dubai, UAE	478 (2.1)																														
Alabama, US	466 (5.9)																														
Abu Dhabi, UAE	449 (3.7)																														

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

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

Average achievement significantly higher than comparison country										Average achievement significantly lower than comparison country										Average Scale Score		Country																																																																																																																																												
Chile	Iran, Islamic Rep. of	Qatar	Bahrain	Jordan	Palestinian Nat'l Auth.	Saudi Arabia	Indonesia	Syrian Arab Republic	Morocco	Oman	Ghana	Botswana (9)	South Africa (9)	Honduras (9)	Benchmarking Participants	Massachusetts, US	Minnesota, US	North Carolina, US	Quebec, Canada	Indiana, US	Colorado, US	Connecticut, US	Florida, US	Ontario, Canada	Alberta, Canada	California, US	Dubai, UAE	Alabama, US	Abu Dhabi, UAE	613 (2.9)	611 (3.8)	609 (3.2)	586 (3.8)	570 (2.6)	539 (3.6)	516 (4.1)	514 (2.5)	509 (2.6)	507 (5.5)	505 (3.5)	505 (5.1)	505 (2.2)	502 (2.5)	498 (2.4)	488 (5.5)	487 (4.0)	484 (1.9)	479 (3.9)	475 (2.4)	467 (2.7)	458 (4.0)	456 (2.1)	452 (3.9)	449 (3.7)	440 (5.4)	431 (3.8)	427 (4.3)	426 (5.2)	425 (2.8)	416 (2.6)	415 (4.3)	410 (3.1)	409 (2.0)	406 (3.7)	404 (3.5)	394 (4.6)	386 (4.3)	380 (4.5)	371 (2.0)	366 (2.8)	331 (4.3)	397 (2.5)	352 (2.5)	338 (3.7)	Korea, Rep. of	Singapore	Chinese Taipei	Hong Kong SAR	Japan	Russian Federation	Israel	Finland	United States	England	Hungary	Australia	Slovenia	Lithuania	Italy	New Zealand	Kazakhstan	Sweden	Ukraine	Norway	Armenia	Romania	United Arab Emirates	Turkey	Lebanon	Malaysia	Georgia	Thailand	Macedonia, Rep. of	Tunisia	Chile	Iran, Islamic Rep. of	Qatar	Bahrain	Jordan	Palestinian Nat'l Auth.	Saudi Arabia	Indonesia	Syrian Arab Republic	Morocco	Oman	Ghana	Botswana (9)	South Africa (9)	Honduras (9)	Benchmarking Participants	Massachusetts, US	Minnesota, US	North Carolina, US	Quebec, Canada	Indiana, US	Colorado, US	Connecticut, US	Florida, US	Ontario, Canada	Alberta, Canada	California, US	Dubai, UAE	Alabama, US	Abu Dhabi, UAE	561 (5.3)	545 (4.6)	537 (6.8)	532 (2.3)	522 (5.1)	518 (4.9)	518 (4.8)	513 (6.4)	512 (2.5)	505 (2.6)	493 (4.9)	478 (2.1)	466 (5.9)	449 (3.7)	Massachusetts, US	Minnesota, US	North Carolina, US	Quebec, Canada	Indiana, US	Colorado, US	Connecticut, US	Florida, US	Ontario, Canada	Alberta, Canada	California, US	Dubai, UAE	Alabama, US	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 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 mathematics content and cognitive domains. The items comprising the mathematics 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 Kuwait, 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 \mathcal{X} . 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 Oman 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 348 and 419, falling at or below the Low International Benchmarks (400). This level of achievement is comparable to that of most of the lower-performing countries at the fourth grade. In addition, there was evidence of many very low-performing sixth grade students in Yemen (annotated with the symbol \mathcal{K} , indicating that the percentage of students with achievement too low for estimation exceeded 25%) and to a lesser extent in Honduras (annotated with the symbol Ψ , indicating that the percentage of students with achievement too low for estimation exceeded 15% but did not exceed 25%). 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 100 points.

Achievement in TIMSS 2011 at the Eighth Grade

The results in Exhibit 1.2 (first page) show that Korea, Singapore, Chinese Taipei, Hong Kong SAR, and Japan, the five Asian countries with the highest average mathematics achievement at fourth grade, also have the highest achievement at eighth grade, with average achievement above the High International Benchmark of 550 in each case. In addition to these countries, the Russian Federation, Israel, Finland, the United States, and Slovenia had higher achievement than the scale centerpoint of 500.

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 Korea, Singapore, and Chinese Taipei performed similarly and had higher achievement than all of the other countries. The next highest-performing country was Hong Kong SAR, which had higher achievement than all countries except the three with the highest achievement, followed by Japan, which had average achievement higher than all countries except Hong Kong SAR and the three top performers. Also included in the top ten high-achieving countries were the Russian Federation, Israel, Finland, the United States, and England. Among benchmarking participants, the state of Massachusetts was outperformed only by the four highest achieving Asian countries, while the states of Minnesota and North Carolina and the Canadian province of Québec were outperformed only by the top five 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. Although ten countries had average achievement above the TIMSS centerpoint of 500, twenty-seven countries had average achievement below this point, mostly falling above the Low (400) International Benchmark.

Similar to the fourth grade, a number of eighth grade participants had significant percentages of very low performing students, including Morocco and Ghana (percentage of students with achievement too low for estimation exceeded 25%), and Macedonia, Iran, Qatar, Bahrain, Jordan, Palestinian National Authority, Saudi Arabia, Indonesia, the Syrian Arab Republic, and Oman (percentage of students with achievement too low for estimation exceeded 15% but did not exceed 25%).

Achievement in TIMSS 2011 at the Ninth Grade

As a group, the countries assessing their sixth grade students had average achievement between 338 and 397, below the Low (400) International Benchmark for eighth grade students. In addition, there was evidence of many very low performing ninth grade students in all three countries, with the percentage of students with achievement too low for estimation exceeding 25 percent in South Africa and Honduras and between 15 percent and 25 percent in Botswana.

Trends in Mathematics Achievement

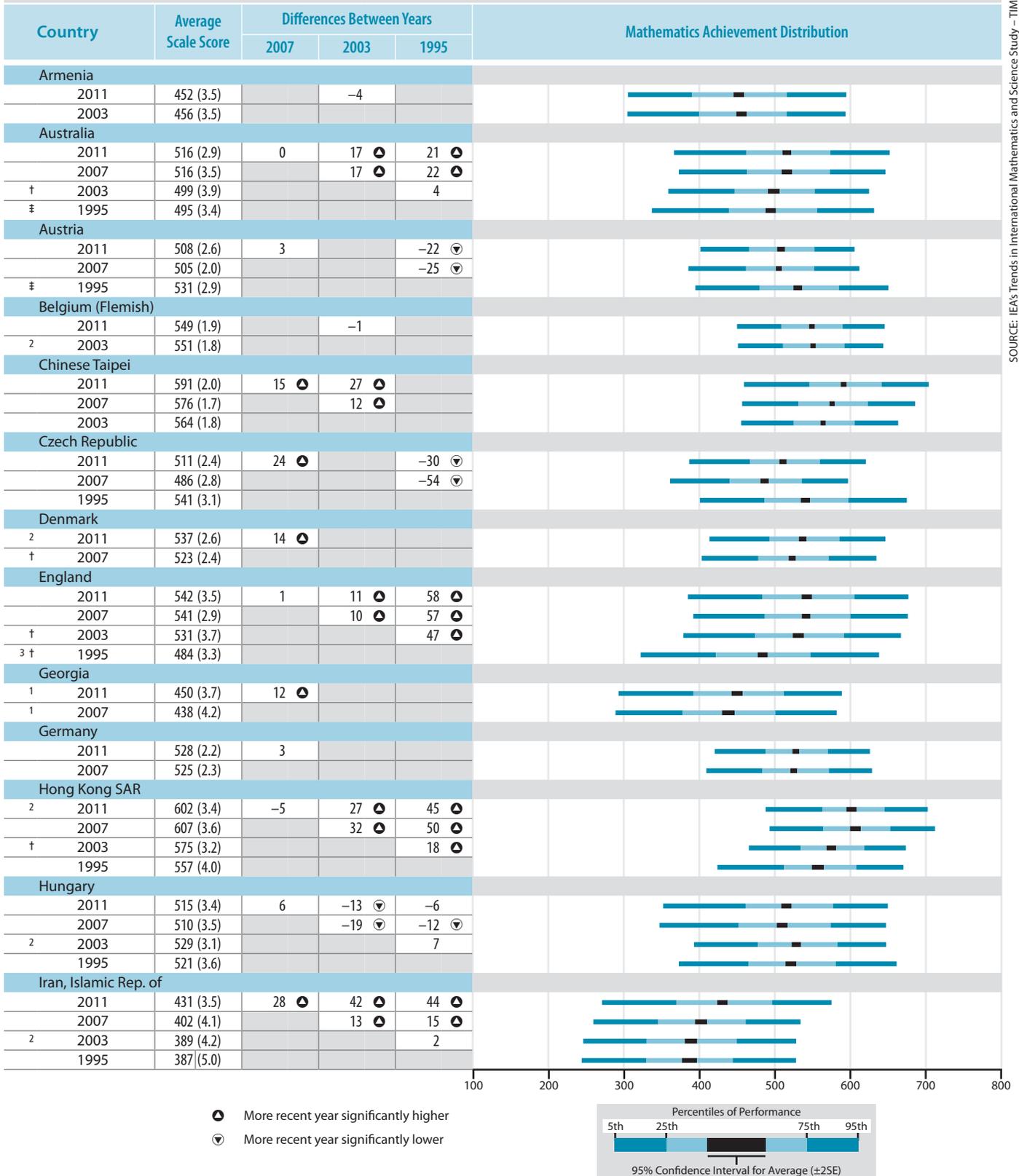
Exhibits 1.5 and 1.6 display changes in average mathematics achievement at fourth and eighth grades, respectively, for the countries and benchmarking participants that have comparable data from previous TIMSS assessments. For the fourth grade, there are 29 countries and four benchmarking participants 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 34 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 mathematics 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 mathematics 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 country's average achievement. To complement Exhibit 1.6 and focus on long-term educational progress at the eighth grade, Exhibit 1.8 presents a similar depiction for the 25 countries and eight benchmarking participants that have comparable data at the eighth grade from the 1995 or 1999 and 2011 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 mathematics achievement during the 1995–2011 period, there have been more countries with increases

3 TIMSS 1999 did not include a fourth grade assessment.

Exhibit 1.5: Trends in Mathematics 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.
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 Mathematics 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.

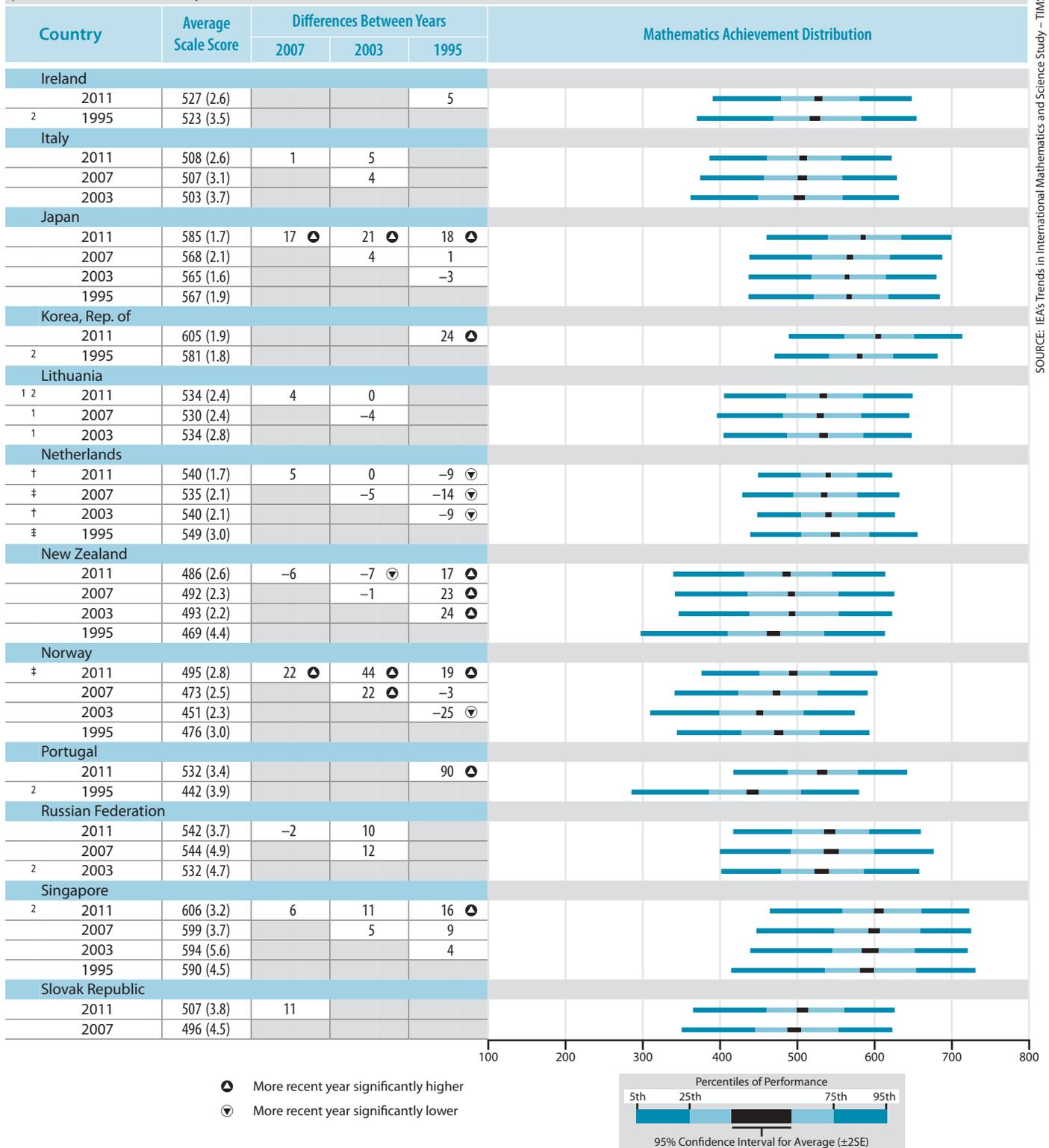


Exhibit 1.5: Trends in Mathematics Achievement (Continued)

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.

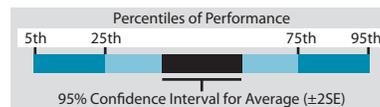
Country	Average Scale Score	Differences Between Years			Mathematics Achievement Distribution
		2007	2003	1995	
Slovenia					
2011	513 (2.2)	11 ▲	34 ▲	51 ▲	
2007	502 (1.8)		23 ▲	40 ▲	
2003	479 (2.6)			17 ▲	
1995	462 (3.1)				
Sweden					
2011	504 (2.0)	1			
2007	503 (2.5)				
Tunisia					
ψ 2011	359 (3.9)	32 ▲	20 ▲		
2007	327 (4.5)		-12		
2003	339 (4.7)				
United States					
2 2011	541 (1.8)	12 ▲	22 ▲	23 ▲	
2 † 2007	529 (2.4)		11 ▲	11 ▲	
† 2003	518 (2.4)			0	
1995	518 (2.9)				

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

Benchmarking Participants

Alberta, Canada					
2 2011	507 (2.5)	1		-17	
2 2007	505 (3.0)			-18 ▼	
1995	523 (8.3)				
Ontario, Canada					
2011	518 (3.1)	6	7	29 ▲	
2 2007	512 (3.1)		0	23 ▲	
2003	511 (3.8)			23 ▲	
2 1995	489 (3.5)				
Quebec, Canada					
2011	533 (2.4)	14 ▲	27 ▲	-17 ▼	
2 2007	519 (3.0)		13 ▲	-31 ▼	
2003	506 (2.4)			-44 ▼	
1995	550 (4.2)				
Dubai, UAE					
2011	468 (1.6)	24 ▲			
♦ ‡ 2007	444 (2.1)				

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

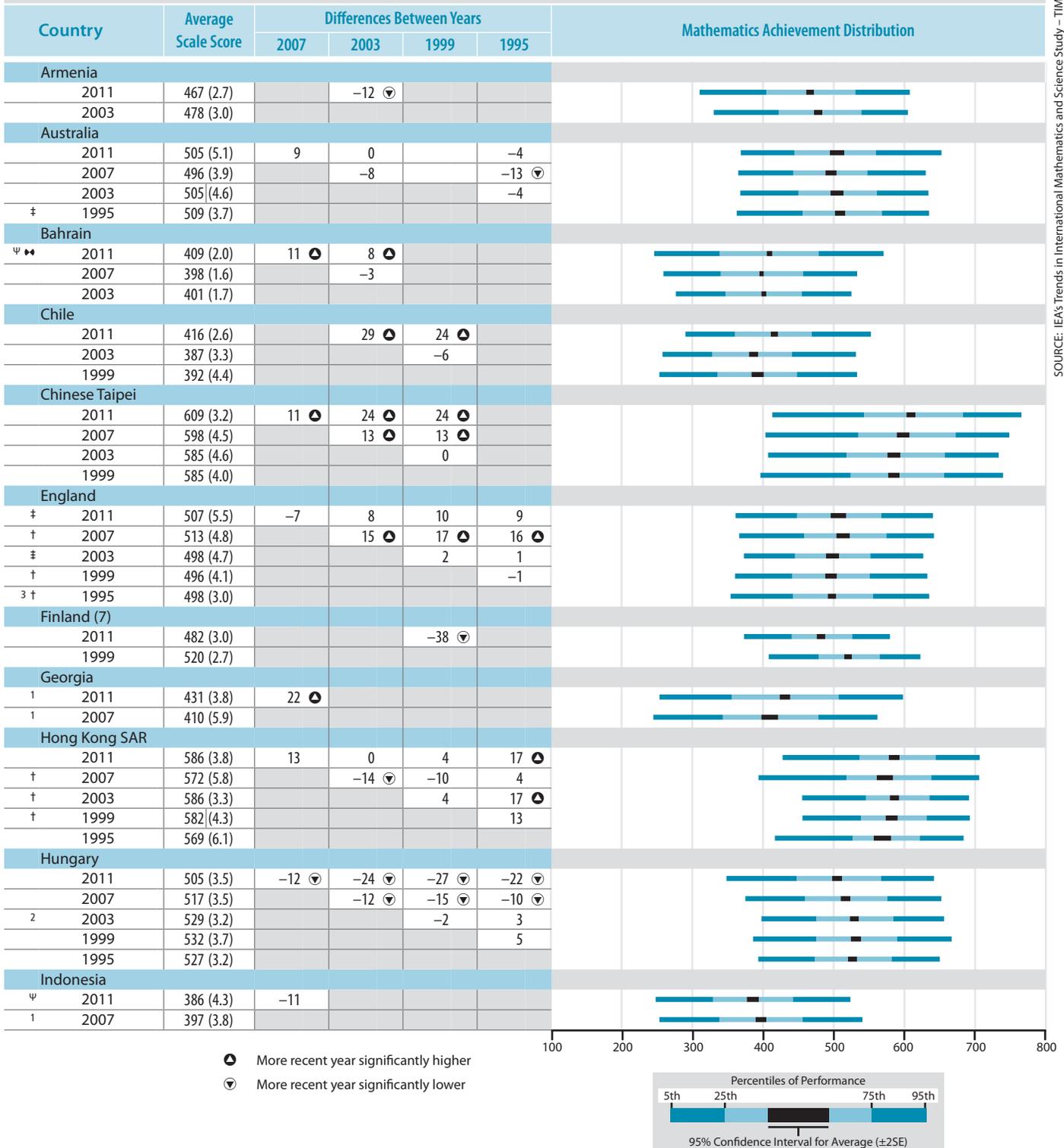


than with decreases. Of the 17 countries and three benchmarking participants with data spanning this period (see Exhibit 1.7), twelve countries and one benchmarking participant had increases in average achievement, three countries and one benchmarking participant had decreases, and two countries and one benchmarking participant had no difference. Among the countries with the greatest increase from 1995 to 2011 were Portugal, England, Slovenia, Hong Kong SAR, and Iran, with average achievement increases of more than 40 points. Australia, Korea, and the United States increased more than 20 points, as did the province of Ontario.

At the eighth grade, there was more balance between mathematics achievement growth and decline among countries. Of the 25 countries and eight benchmarking participants with comparable data spanning the 1995 or 1999 to 2011 period, nine countries and four benchmarking participants had increased achievement, eleven countries and two benchmarking participants had decreased achievement, and five countries and two benchmarking participants showed no difference. The countries with the greatest increases in average mathematics achievement at the eighth grade included Korea, Lithuania, Chinese Taipei, and Chile (more than 20 points), as well as Italy, the United States, Hong Kong SAR, the Russian Federation, and Slovenia (10–20 points). Countries with the greatest decreases included Thailand, Sweden, and Malaysia (40 points or more), and Macedonia, Jordan, Hungary, Tunisia, and Norway (20–40 points). Also, Finland had a comparable decrease (38 points) for its seventh grade students. Among benchmarking participants, there were increases in Massachusetts (47 points), North Carolina (42 points), Minnesota (26 points), and Ontario (11 points). Alberta and Québec had decreased achievement over the period (22 and 25 points, respectively).

Exhibit 1.6: Trends in Mathematics 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.



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.

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

Exhibit 1.6: Trends in Mathematics 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.

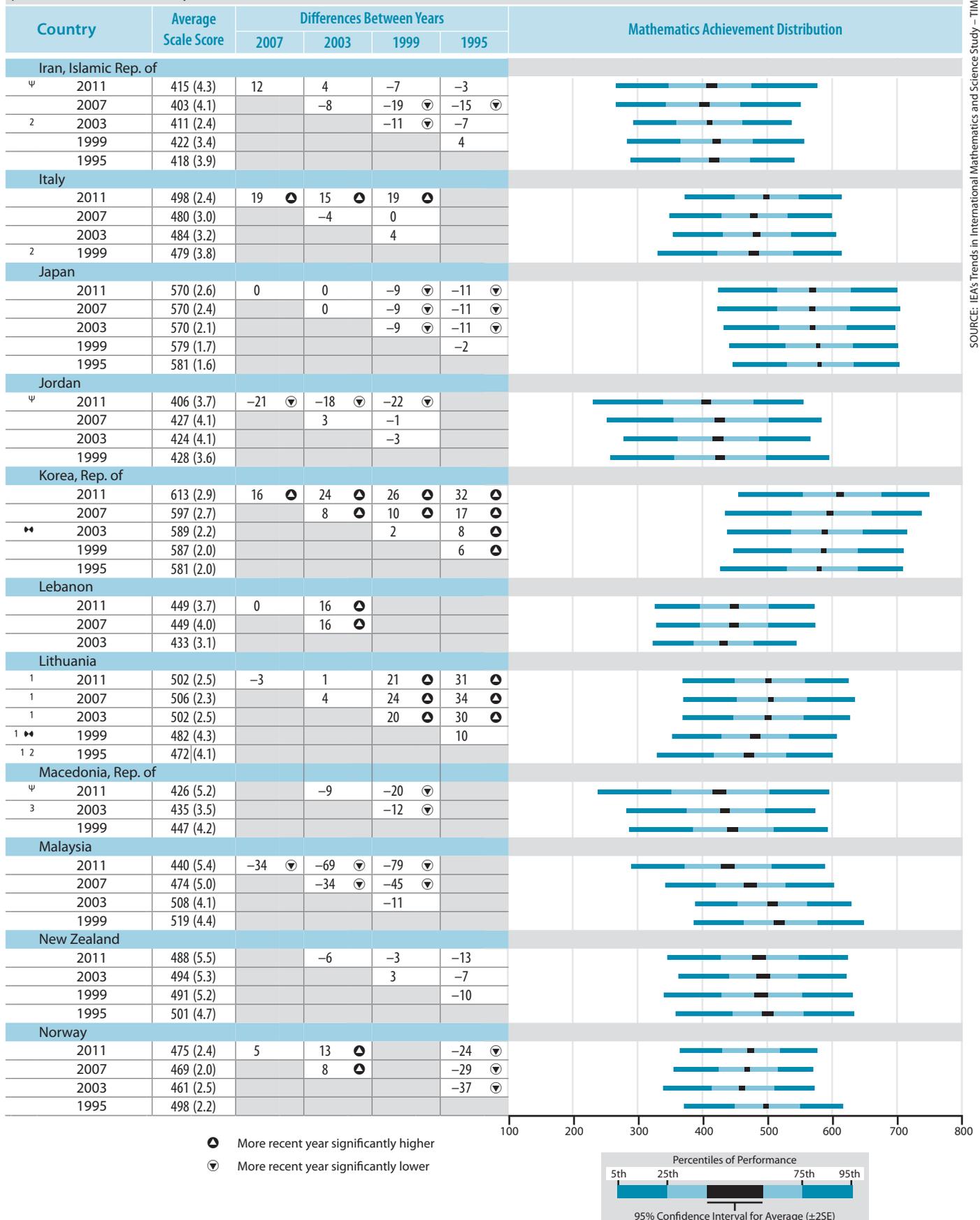
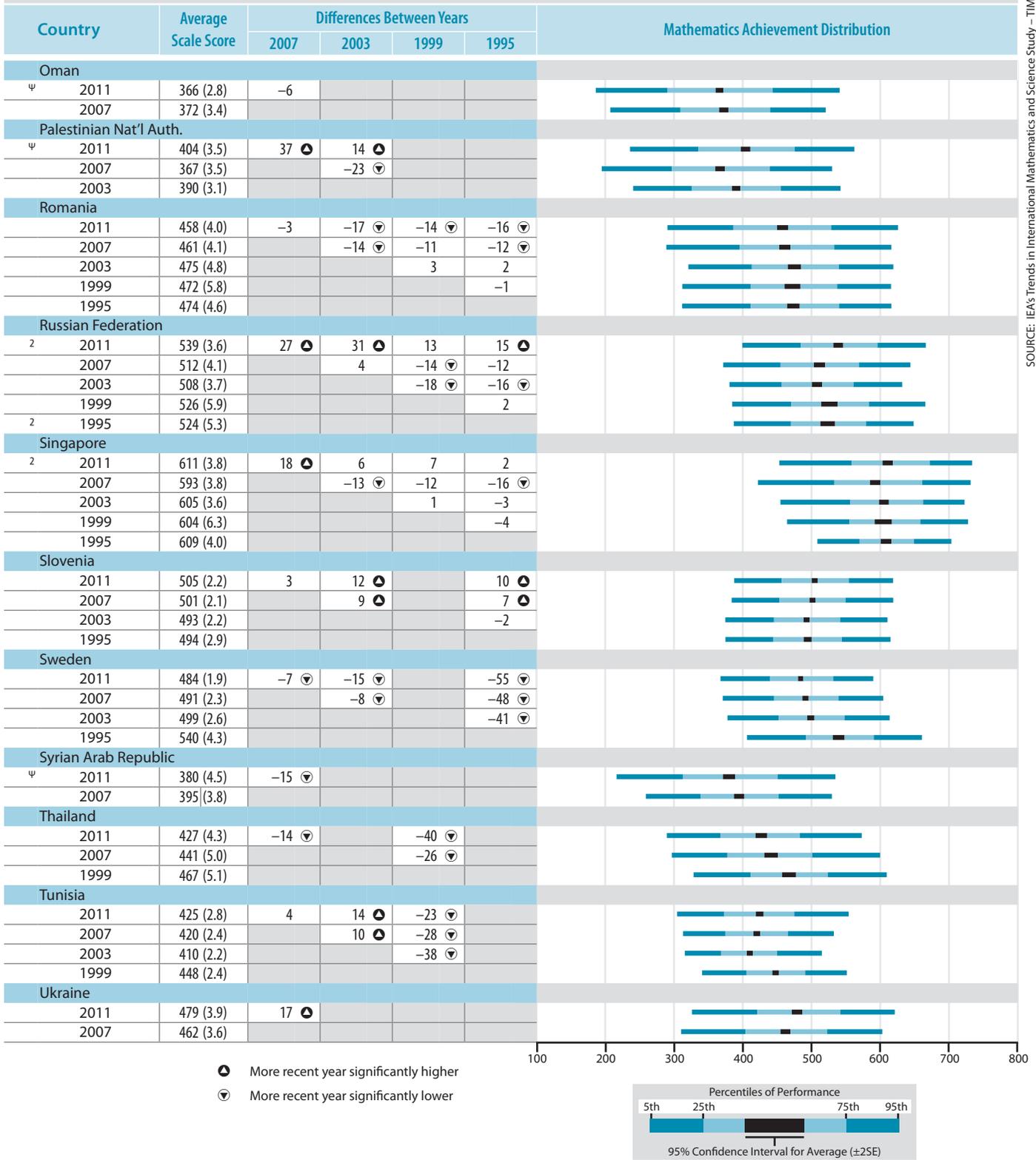


Exhibit 1.6: Trends in Mathematics Achievement (Continued)

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 Mathematics 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.

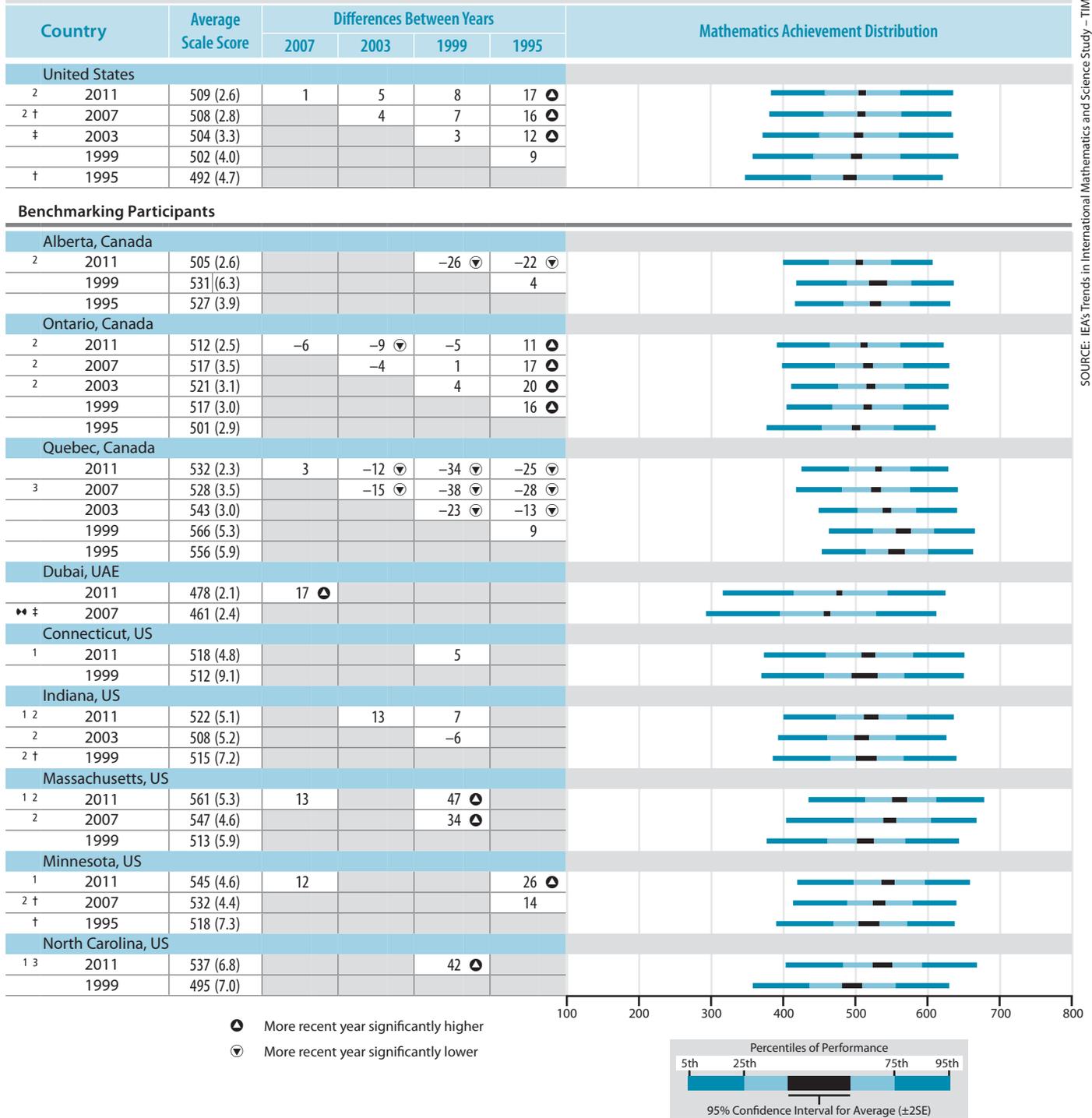
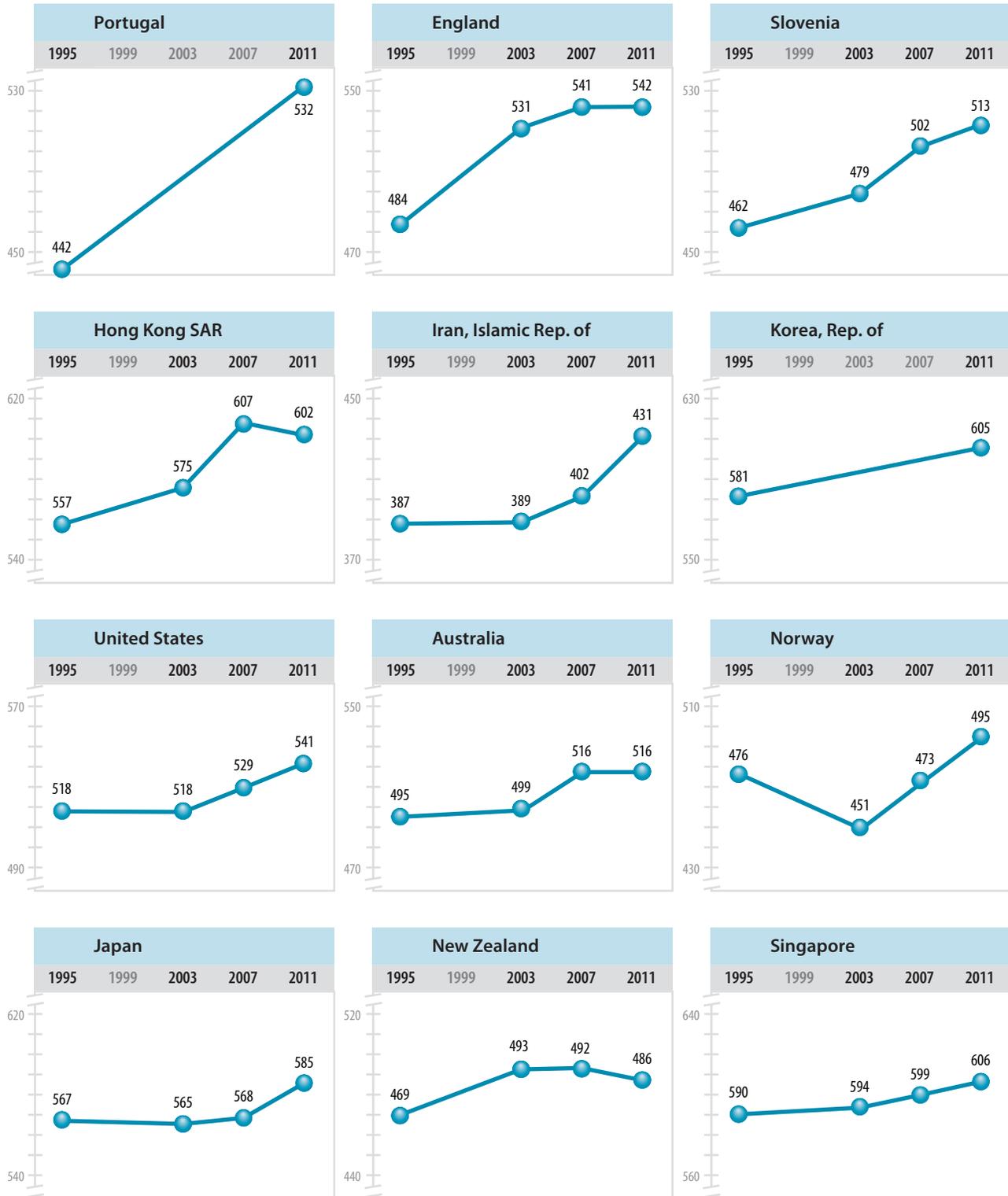


Exhibit 1.7: Trends in Mathematics Achievement – 1995 Through 2011*

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.

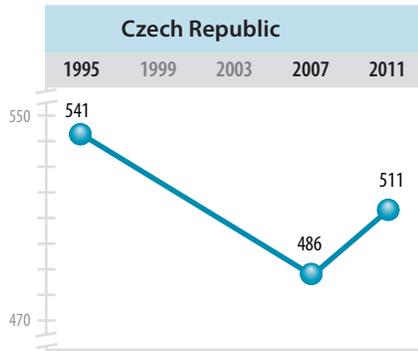
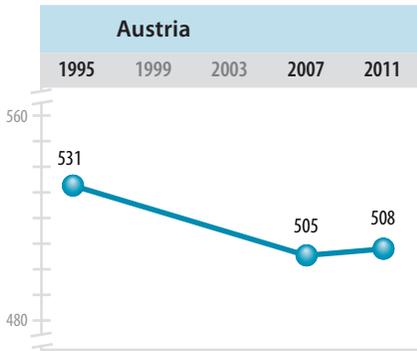
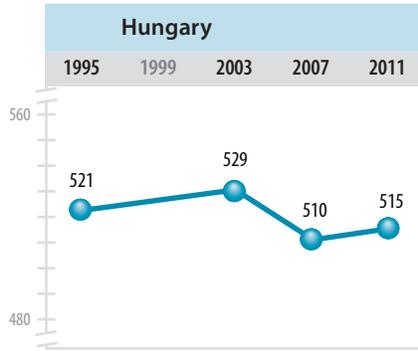


SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

* 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 Mathematics Achievement – 1995 Through 2011* (Continued)



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

Benchmarking Participants

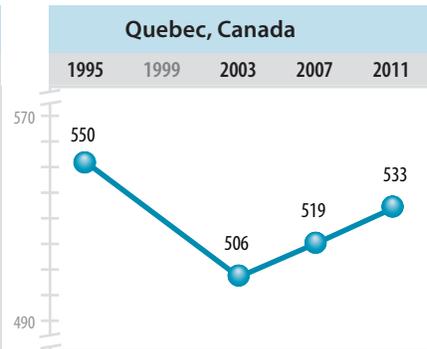
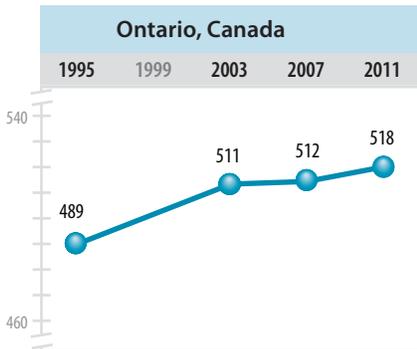
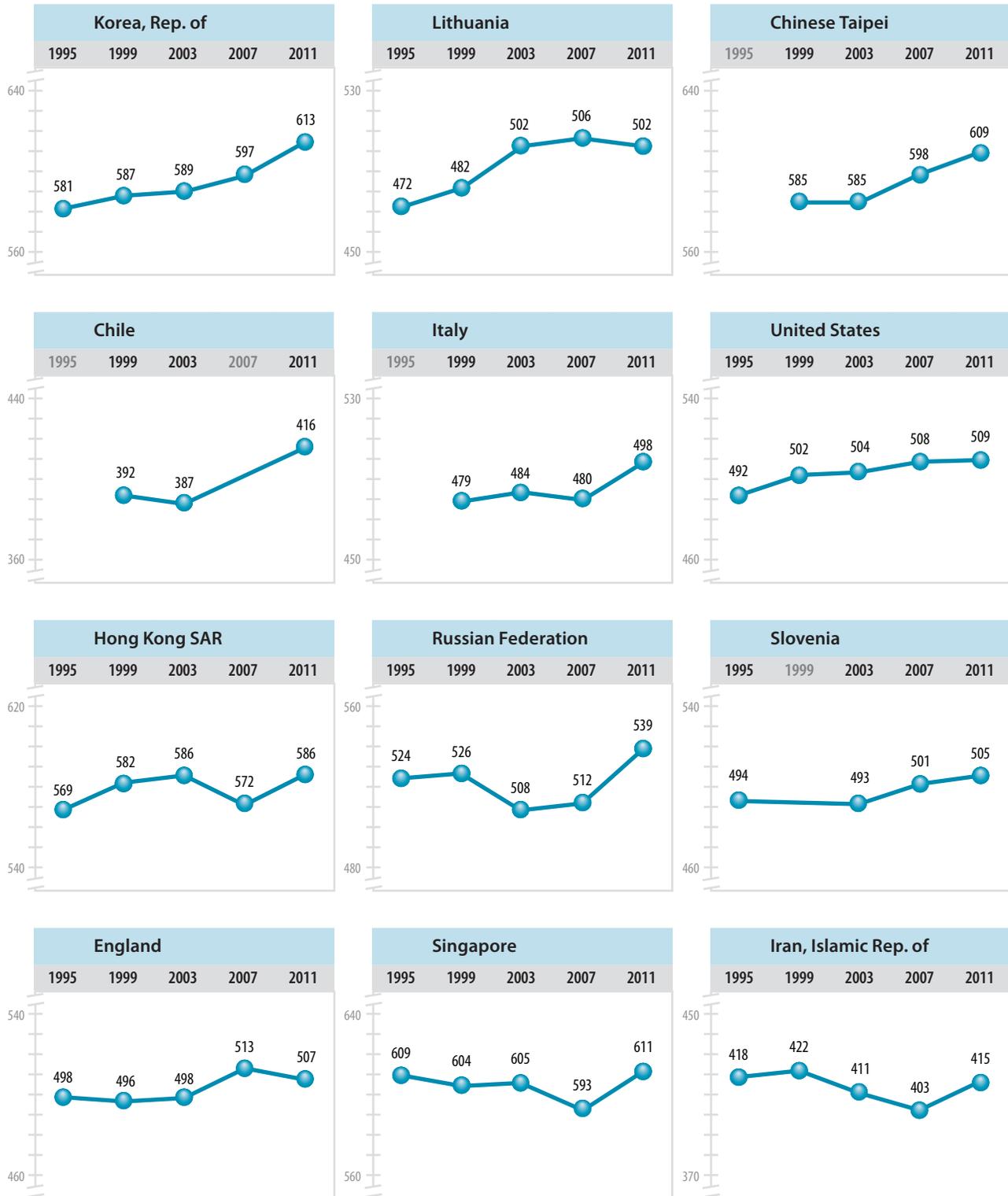


Exhibit 1.8: Trends in Mathematics Achievement – 1995 Through 2011

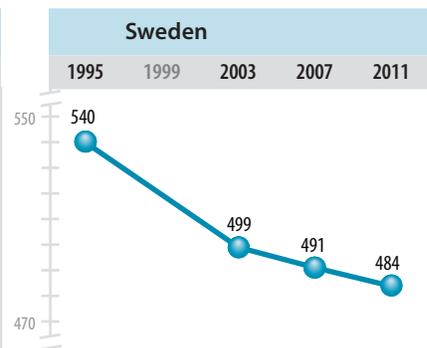
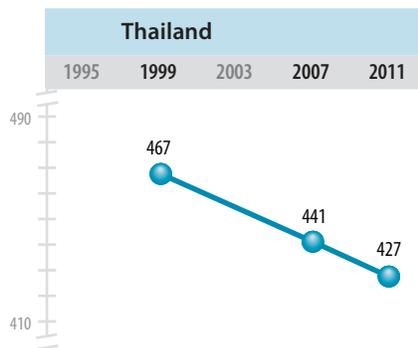
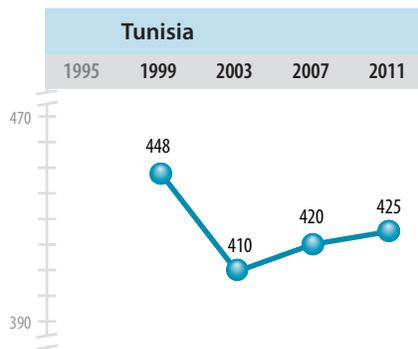
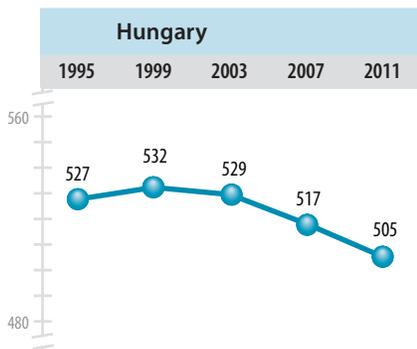
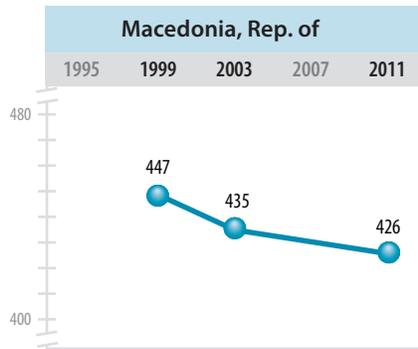
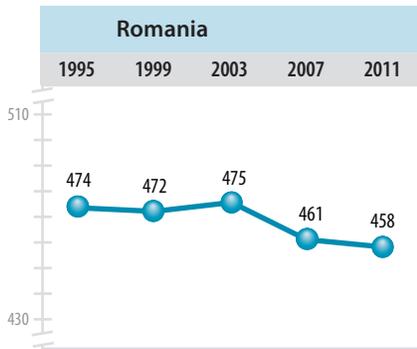
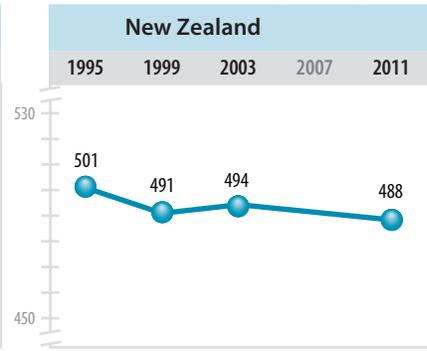
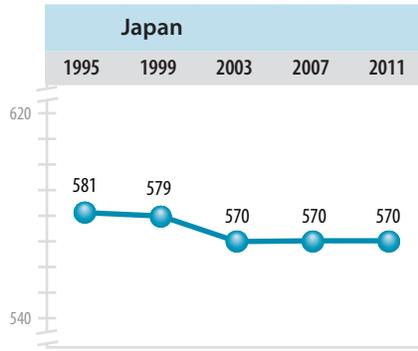
Includes only 2011 participants with comparable long term trend data beginning in either 1995 or 1999, ordered by most to least improvement in average achievement. Exhibit 1.6 provides details including statistical significance.



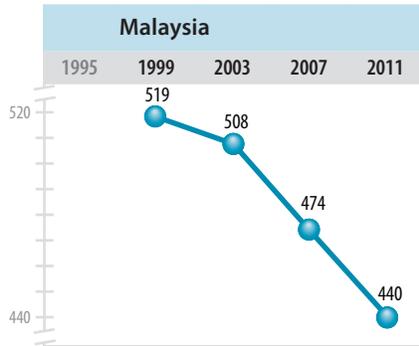
SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

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

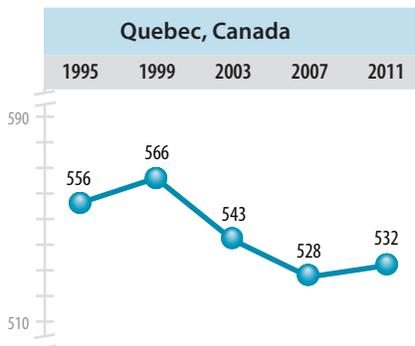
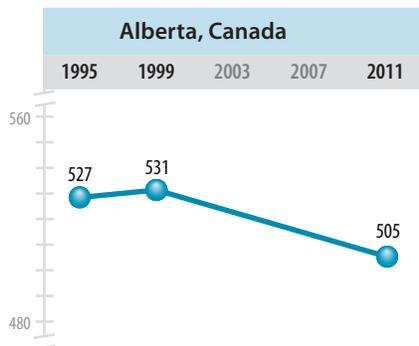
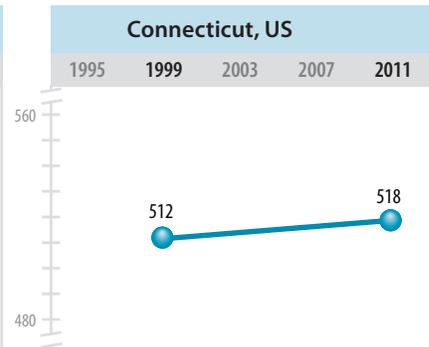
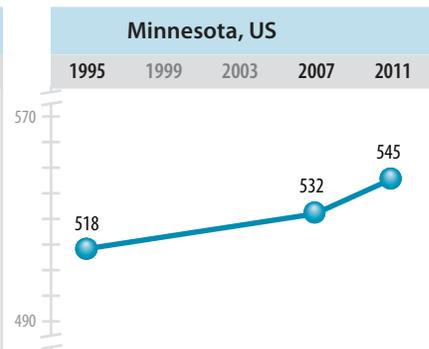
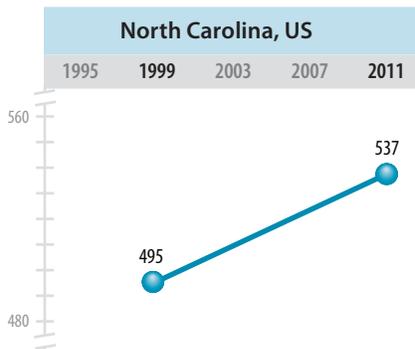
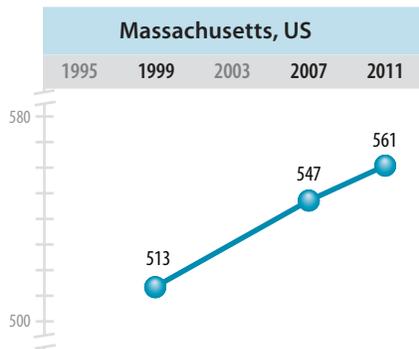
Exhibit 1.8: Trends in Mathematics Achievement – 1995 Through 2011 (Continued)



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 mathematics 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.

Six countries, including Hong Kong SAR, Singapore, Chinese Taipei, Japan, the Russian Federation, and the United States as well as the two 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). Norway, Georgia, Iran, Tunisia and Dubai, UAE also retained the same relative positions, performing below the scale centerpoint in the fourth grade in 2007 and again at the eighth grade in 2011. However, six countries had a relative decline in achievement from fourth to eighth grades, with England, Lithuania, Australia, Hungary, and Italy

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)	
Hong Kong SAR	107 (3.6)	▲
Singapore	99 (3.7)	▲
Chinese Taipei	76 (1.7)	▲
Japan	68 (2.1)	▲
Russian Federation	44 (4.9)	▲
England	41 (2.9)	▲
Lithuania	30 (2.4)	▲
United States	29 (2.4)	▲
Australia	16 (3.5)	▲
Hungary	10 (3.5)	▲
Italy	7 (3.1)	▲
Sweden	3 (2.5)	
Slovenia	2 (1.8)	
Norway	-27 (2.5)	▼
Georgia	-62 (4.2)	▼
Iran, Islamic Rep. of	-98 (4.1)	▼
Tunisia	-173 (4.5)	▼
Benchmarking Participants		
Quebec, Canada	19 (3.0)	▲
Ontario, Canada	12 (3.1)	▲
Dubai, UAE	-56 (2.1)	▼

2011 - Fourth Grade		
Country	Achievement Difference from TIMSS Scale Centerpoint (500)	
Singapore	106 (3.2)	▲
Hong Kong SAR	102 (3.4)	▲
Chinese Taipei	91 (2.0)	▲
Japan	85 (1.7)	▲
England	42 (3.5)	▲
Russian Federation	42 (3.7)	▲
United States	41 (1.8)	▲
Lithuania	34 (2.4)	▲
Australia	16 (2.9)	▲
Hungary	15 (3.4)	▲
Slovenia	13 (2.2)	▲
Italy	8 (2.6)	▲
Sweden	4 (2.0)	
Norway	-5 (2.8)	▼
Georgia	-50 (3.7)	▼
Iran, Islamic Rep. of	-69 (3.5)	▼
Tunisia	-141 (3.9)	▼
Benchmarking Participants		
Quebec, Canada	33 (2.4)	▲
Ontario, Canada	18 (3.1)	▲
Dubai, UAE	-32 (1.6)	▼

2007 - Eighth Grade		
Country	Achievement Difference from TIMSS Scale Centerpoint (500)	
Chinese Taipei	98 (4.5)	▲
Singapore	93 (3.8)	▲
Hong Kong SAR	72 (5.8)	▲
Japan	70 (2.4)	▲
Hungary	17 (3.5)	▲
England	13 (4.8)	▲
Russian Federation	12 (4.1)	▲
United States	8 (2.8)	▲
Lithuania	6 (2.3)	▲
Slovenia	1 (2.1)	
Australia	-4 (3.9)	
Sweden	-9 (2.3)	▼
Italy	-20 (3.0)	▼
Norway	-31 (2.0)	▼
Tunisia	-80 (2.4)	▼
Georgia	-90 (5.9)	▼
Iran, Islamic Rep. of	-97 (4.1)	▼
Benchmarking Participants		
Quebec, Canada	28 (3.5)	▲
Ontario, Canada	17 (3.5)	▲
Dubai, UAE	-39 (2.4)	▼

2011 - Eighth Grade		
Country	Achievement Difference from TIMSS Scale Centerpoint (500)	
Singapore	111 (3.8)	▲
Chinese Taipei	109 (3.2)	▲
Hong Kong SAR	86 (3.8)	▲
Japan	70 (2.6)	▲
Russian Federation	39 (3.6)	▲
United States	9 (2.6)	▲
England	7 (5.5)	
Hungary	5 (3.5)	
Australia	5 (5.1)	
Slovenia	5 (2.2)	▲
Lithuania	2 (2.5)	
Italy	-2 (2.4)	
Sweden	-16 (1.9)	▼
Norway	-25 (2.4)	▼
Georgia	-69 (3.8)	▼
Tunisia	-75 (2.8)	▼
Iran, Islamic Rep. of	-85 (4.3)	▼
Benchmarking Participants		
Quebec, Canada	32 (2.3)	▲
Ontario, Canada	12 (2.5)	▲
Dubai, UAE	-22 (2.1)	▼

- ▲ 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

moving from above the centerpoint at the fourth grade in 2007 to close to the centerpoint at the eighth grade in 2011, and Sweden moving from near the centerpoint to below the centerpoint in 2011. In comparison, Slovenia was the only country to show relative improvement, moving from about the centerpoint at the fourth grade in 2007 to just above it at the eighth grade in 2011.

Gender Differences in Mathematics Achievement

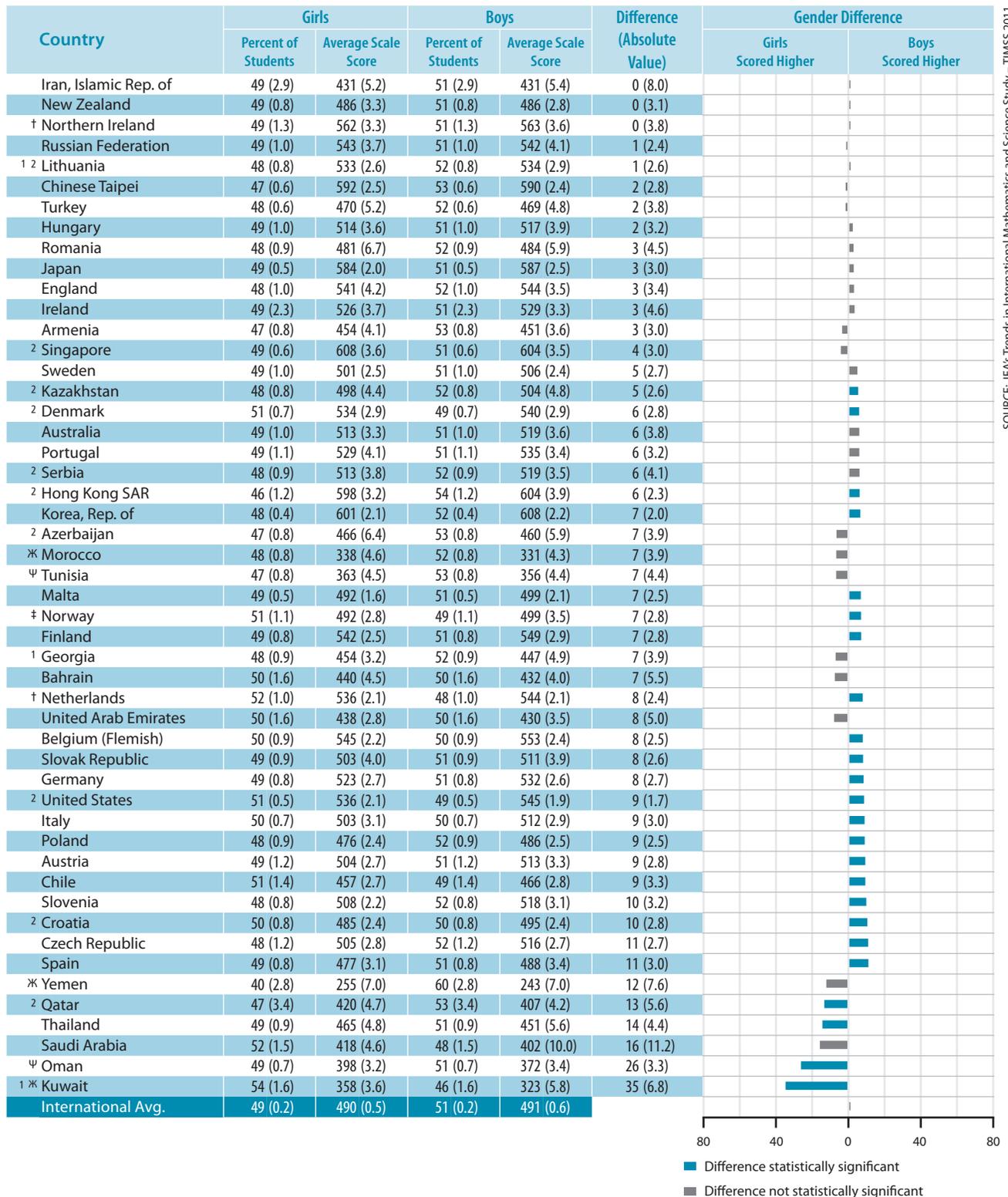
Previous TIMSS assessments have shown gender differences in mathematics achievement to be small on average at the fourth grade and somewhat larger in favor of girls at the eighth grade, although the situation varies considerably from country to country.

Exhibit 1.10 presents the TIMSS 2011 fourth grade results for gender differences in mathematics achievement. For the TIMSS 2011 countries at fourth grade, at sixth grade, and the benchmarking participants, it shows girls' average achievement, boys' average achievement, and the difference between the two averages. The bar graph shows the size of the achievement difference and whether that difference is statistically significant (as indicated by a darkened bar). For countries participating at the fourth grade, 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 mathematics achievement. Averaging mathematics achievement across countries, it is clear that there was little achievement difference between girls and boys (International Average: 490 vs. 491). Of the 50 countries at the fourth grade, 26 had no significant gender difference in mathematics achievement. Of the 24 remaining countries, 20 had small differences favoring boys, and four had relatively larger differences favoring girls (Qatar, Thailand, Oman, and Kuwait). At the sixth grade, there was a significant achievement difference favoring girls in Botswana and favoring boys in Honduras. Boys also had higher average mathematics achievement than girls in each of the benchmarking entities, except Dubai where there was no difference and Abu Dhabi where girls had higher achievement than boys.

As shown in Exhibit 1.11, gender differences in mathematics achievement at the eighth grade were larger, on average, than at fourth grade, with the difference favoring girls (International Average: 469 vs. 465). Similar to the

Exhibit 1.10: Average Mathematics Achievement by Gender



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

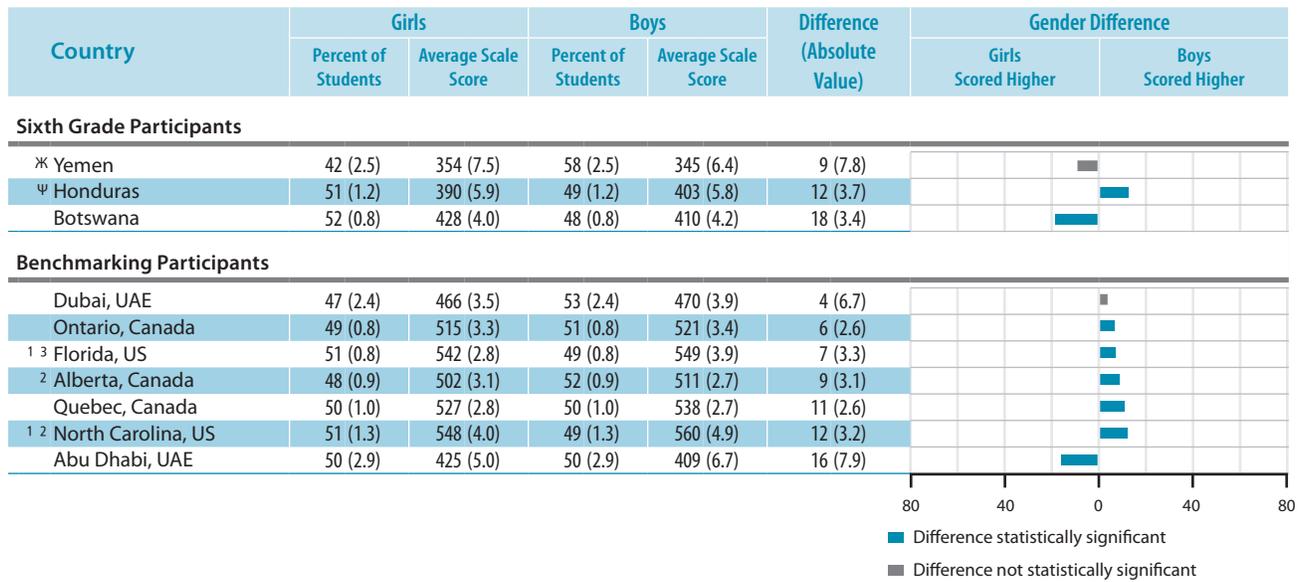
* Average achievement not reliably measured because the percentage of students with achievement too low for estimation exceeds 25%.

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 §.

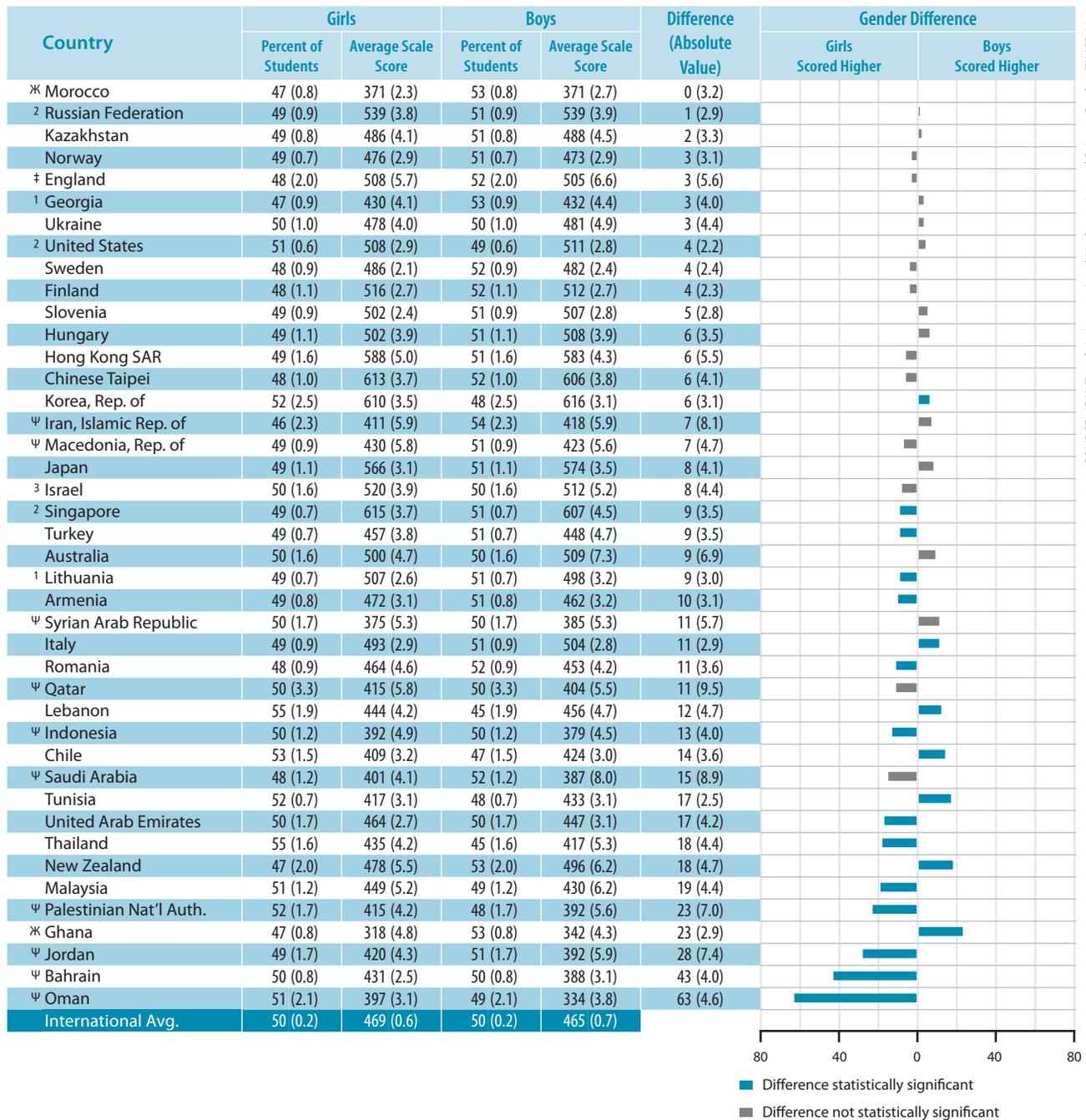
() Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 1.10: Average Mathematics Achievement by Gender (Continued)



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

Exhibit 1.11: Average Mathematics Achievement by Gender



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

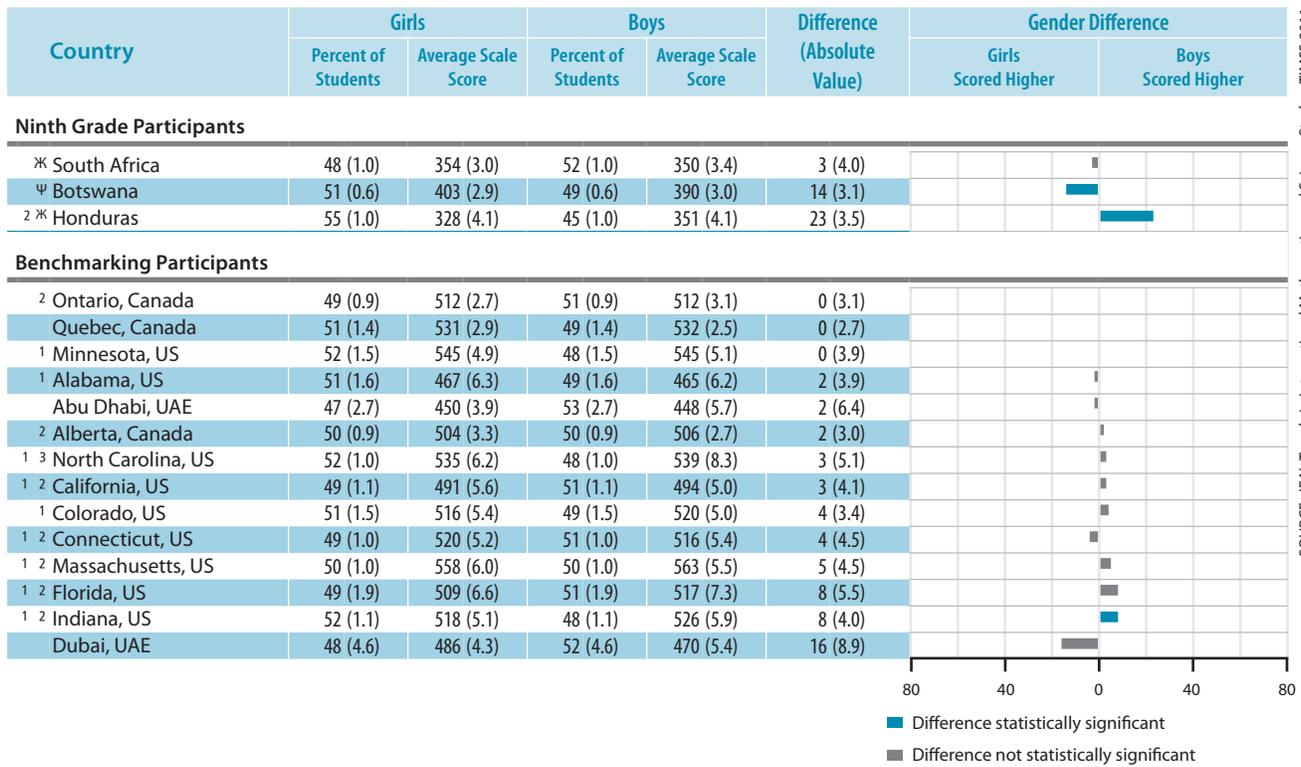
* Average achievement not reliably measured because the percentage of students with achievement too low for estimation exceeds 25%.

‡ 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.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 Mathematics Achievement by Gender (Continued)



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

fourth grade, the gender difference varied across countries, with no difference in 22 of the 42 eighth grade countries, a difference favoring boys in seven countries, and a difference favoring girls in the remaining 13 countries. For Botswana and Honduras, which assessed their ninth grade students, gender differences resembled their sixth grade results, with girls having higher mathematics achievement than boys in Botswana and boys higher than girls in Honduras. There were no gender differences among the eighth grade benchmarking participants, with the exception of the state of Indiana, where boys performed better than girls by a small margin.

At both fourth and eighth grades, and consistent with findings from TIMSS 2007, the largest achievement differences favoring girls were often in Arabic-speaking countries from the Middle East, including Qatar, Oman, Kuwait, and Abu Dhabi, UAE at fourth grade and the United Arab Emirates, Palestinian National Authority, Jordan, Bahrain, and Oman at eighth grade.

Trends in Mathematics Achievement by Gender

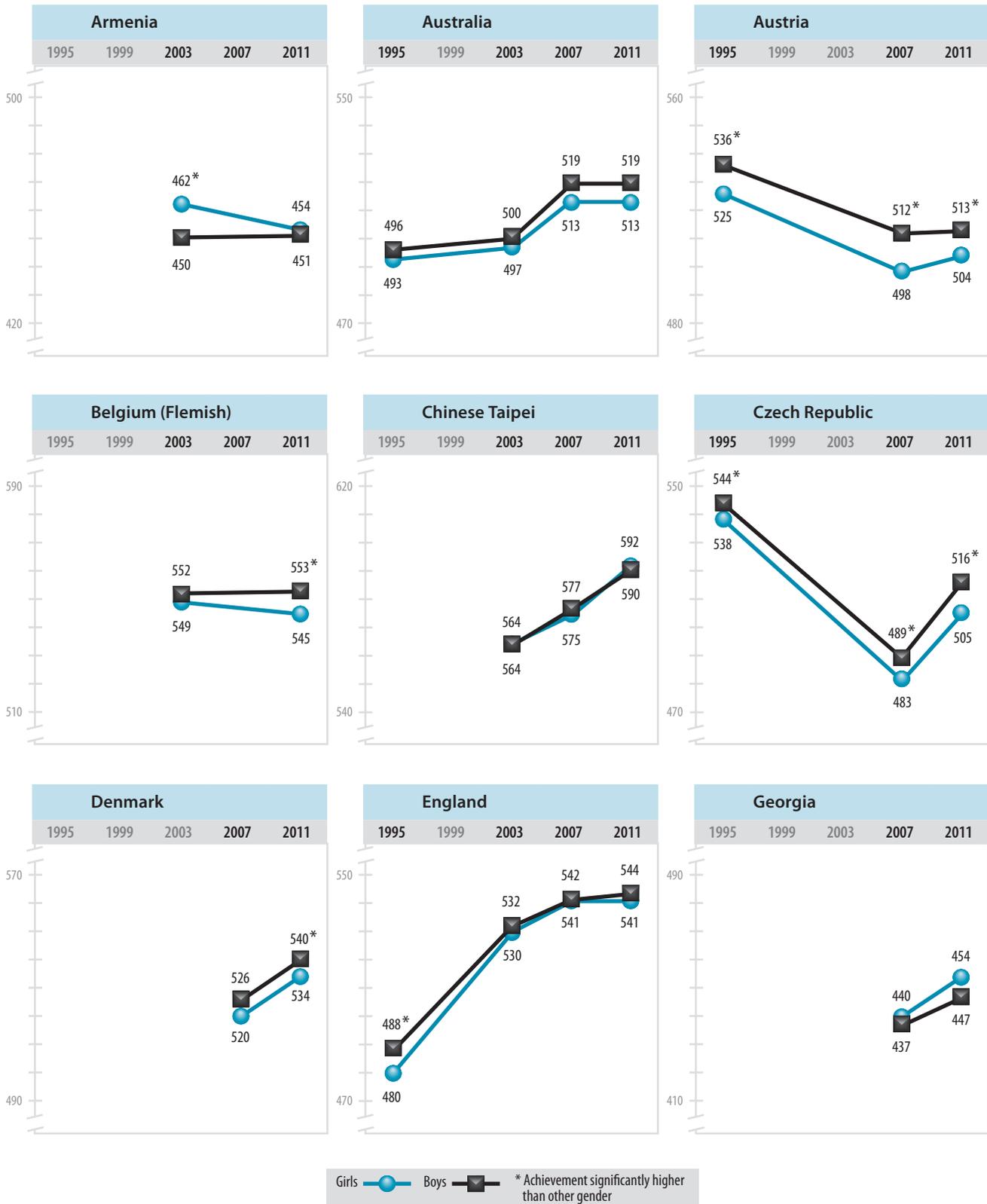
Exhibits 1.12 and 1.13 show graphic representations across the TIMSS assessments of trends in mathematics achievement of boys and girls for fourth and eighth grades, respectively. For each country that participated in one or more of the previous TIMSS assessments, these displays show how trends in mathematics 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 mathematics 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 mathematics 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 where fourth grade girls performed consistently below boys (i.e., in 2011 and on at least two other TIMSS assessments) include Austria, the Czech Republic, Italy, the Netherlands, Slovenia, the United States, and the Canadian province of Québec. In Germany, Korea, and the Slovak Republic, boys had higher average achievement than girls on each of the two TIMSS assessments in which they participated. Armenia, Sweden, and Tunisia had gender differences in earlier assessments but not in TIMSS 2011.

With greater gender differences among countries, and trends across five TIMSS assessments, trends at the eighth grade in mathematics achievement for boys and girls follow a variety of paths. A number of countries show an increasing difference across the years, including Bahrain, Indonesia, Jordan, Lithuania, Malaysia, New Zealand, Oman, and Romania. However, there were few instances of countries decreasing an existing gender gap in mathematics achievement.

Exhibit 1.12: Trends in Mathematics Achievement by Gender[◊]

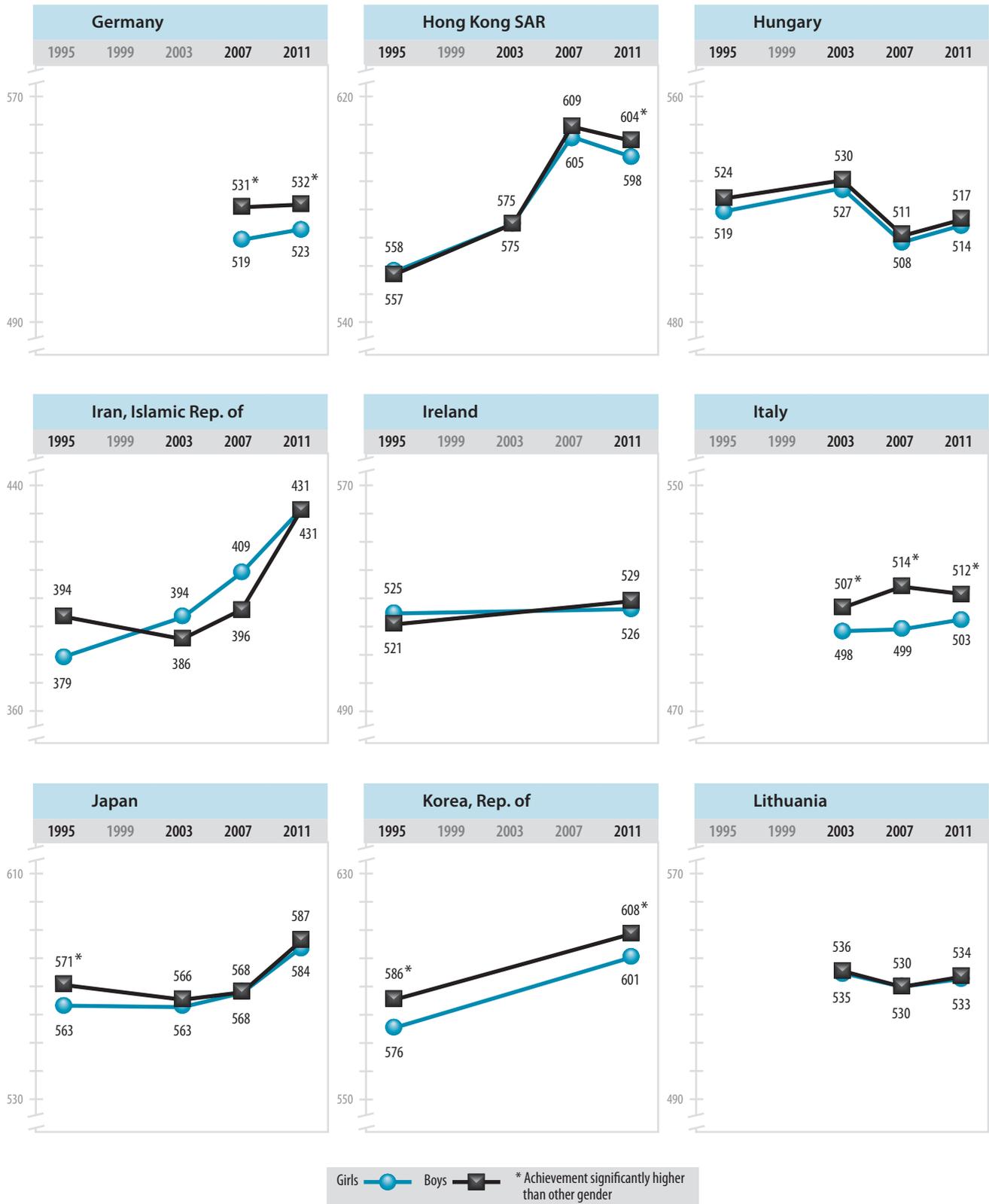
SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011



◊ 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 Mathematics Achievement by Gender^o (Continued)



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

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

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

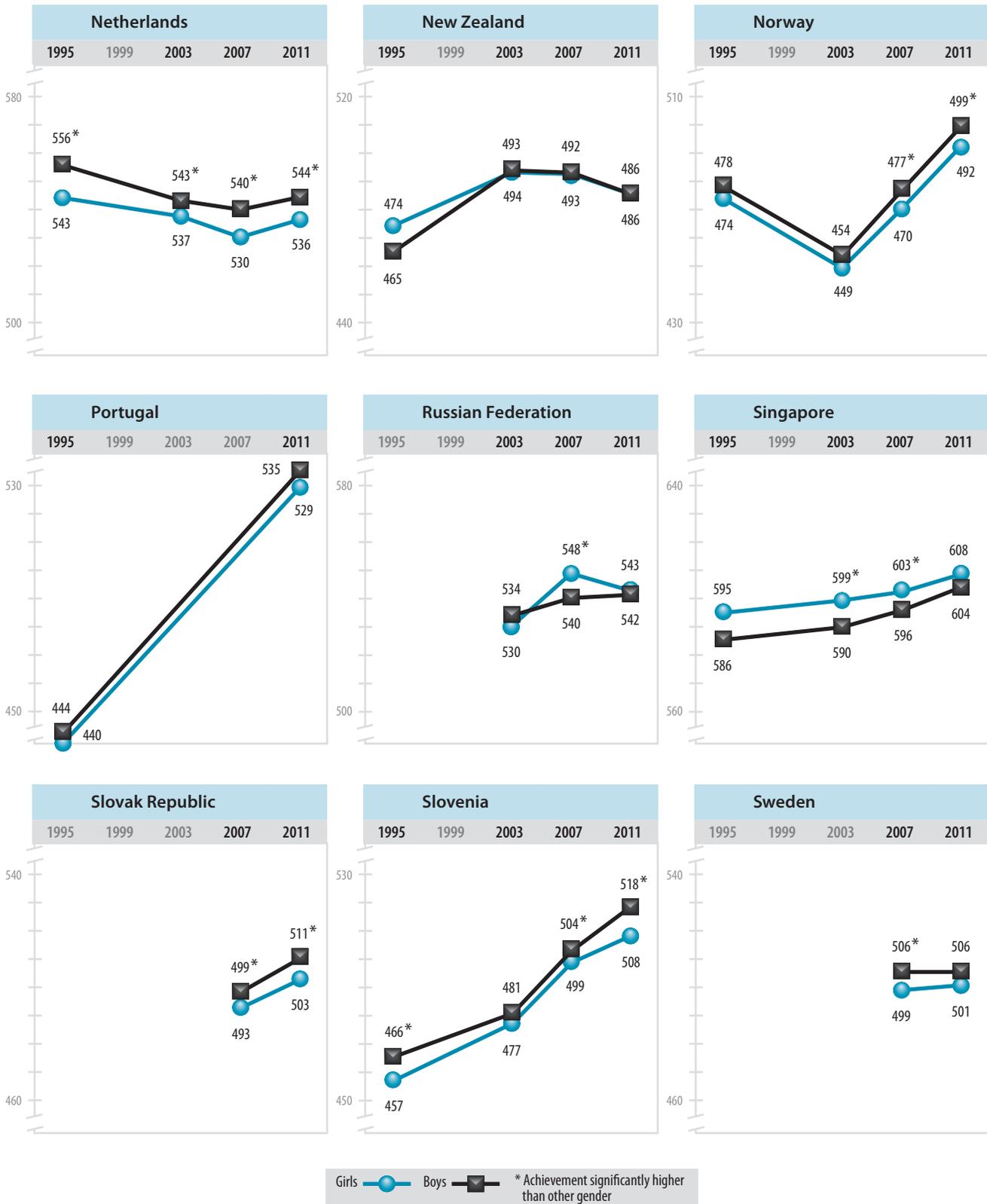
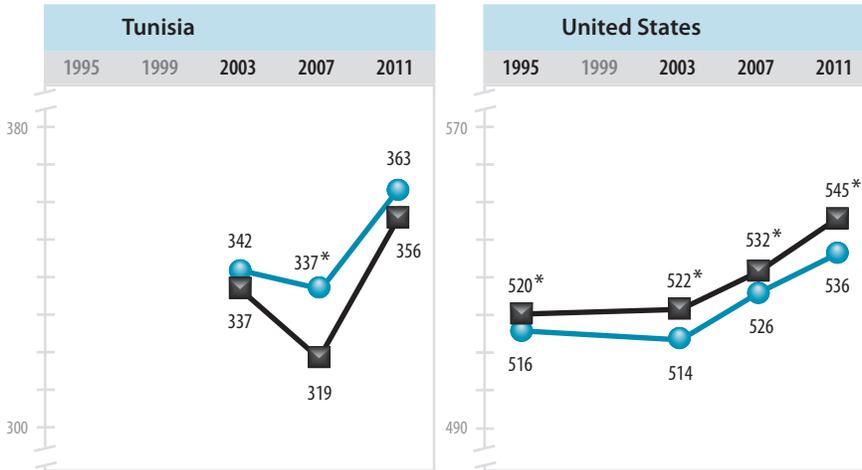
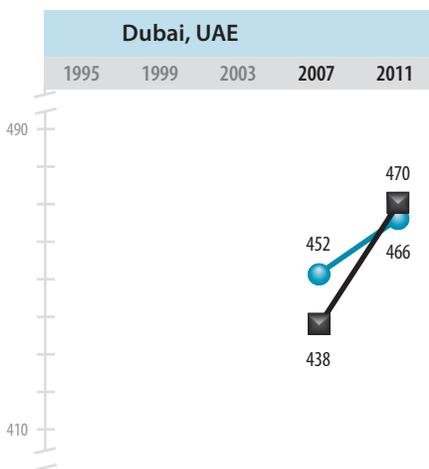
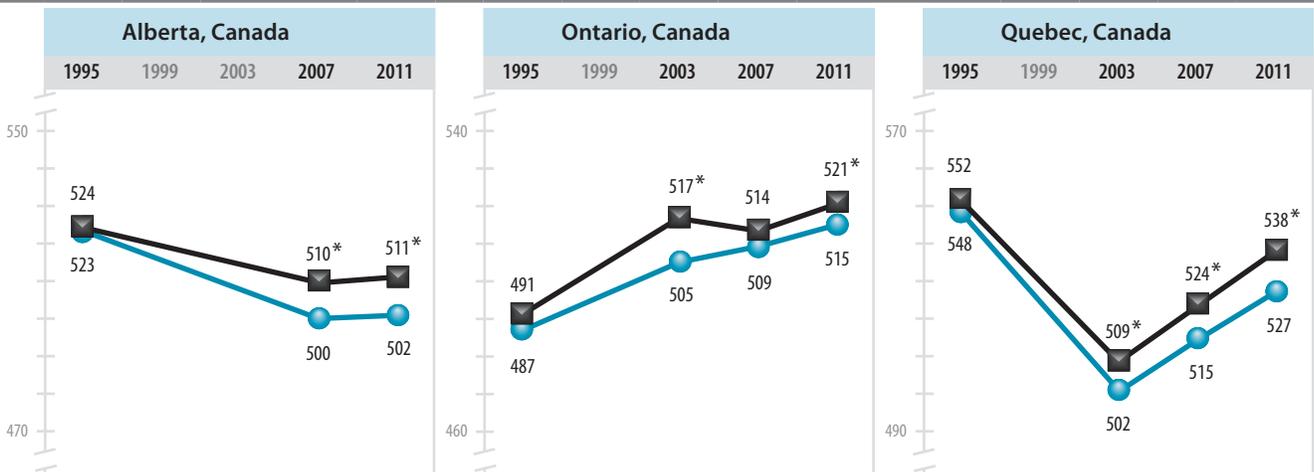


Exhibit 1.12: Trends in Mathematics Achievement by Gender^o (Continued)

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



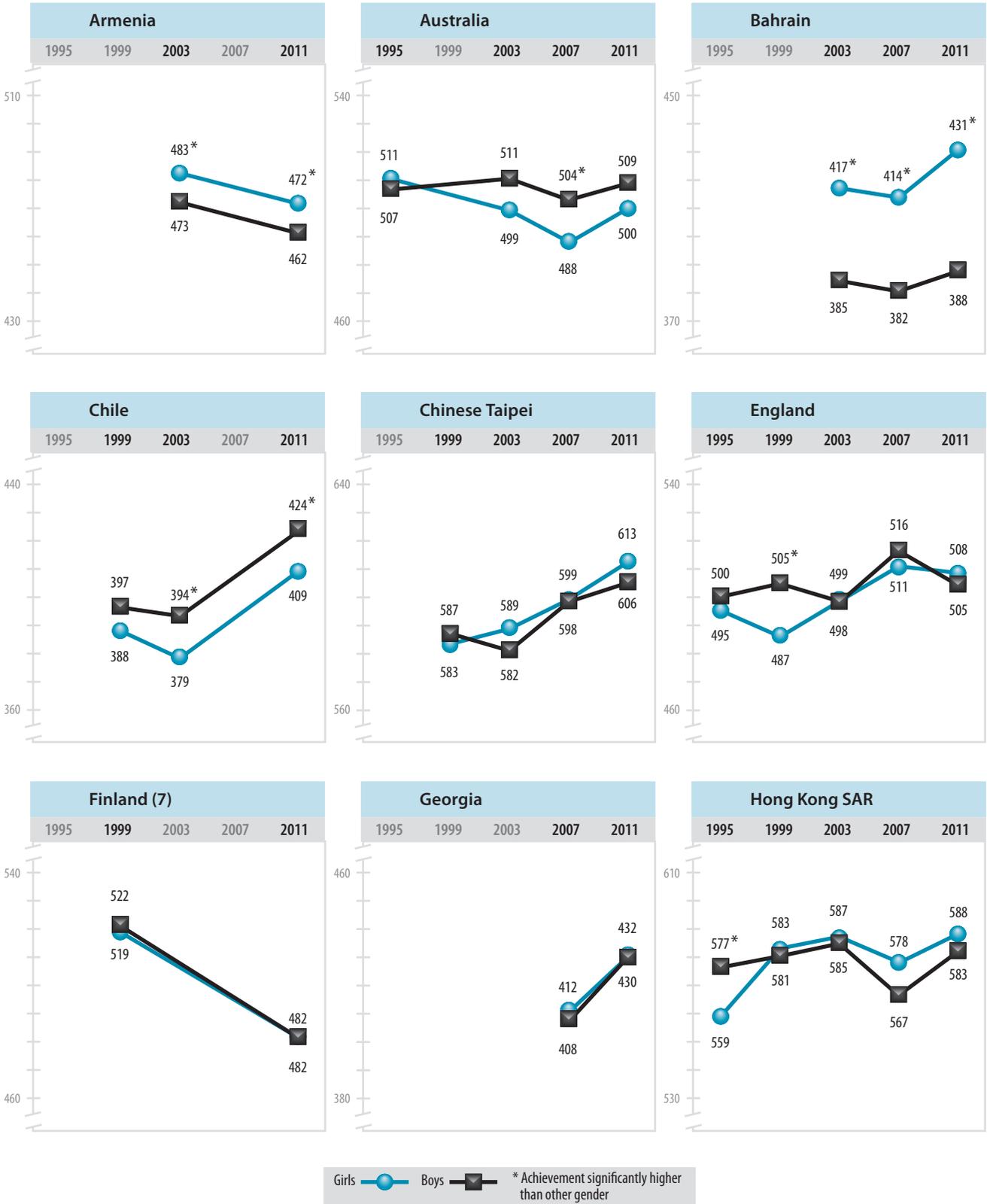
Benchmarking Participants



Girls ● Boys ■ * Achievement significantly higher than other gender

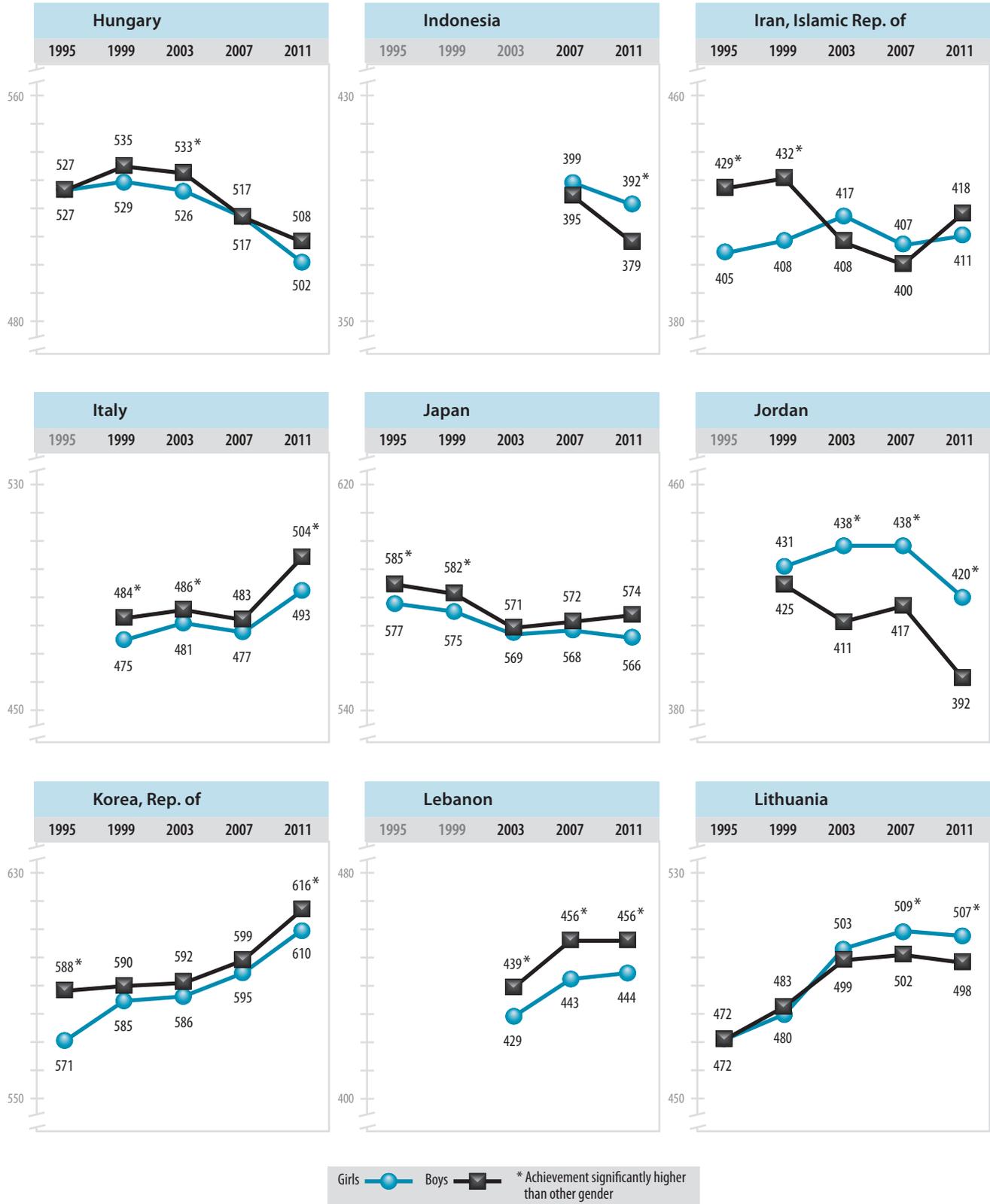
Exhibit 1.13: Trends in Mathematics Achievement by Gender

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 Mathematics Achievement by Gender (Continued)



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

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

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

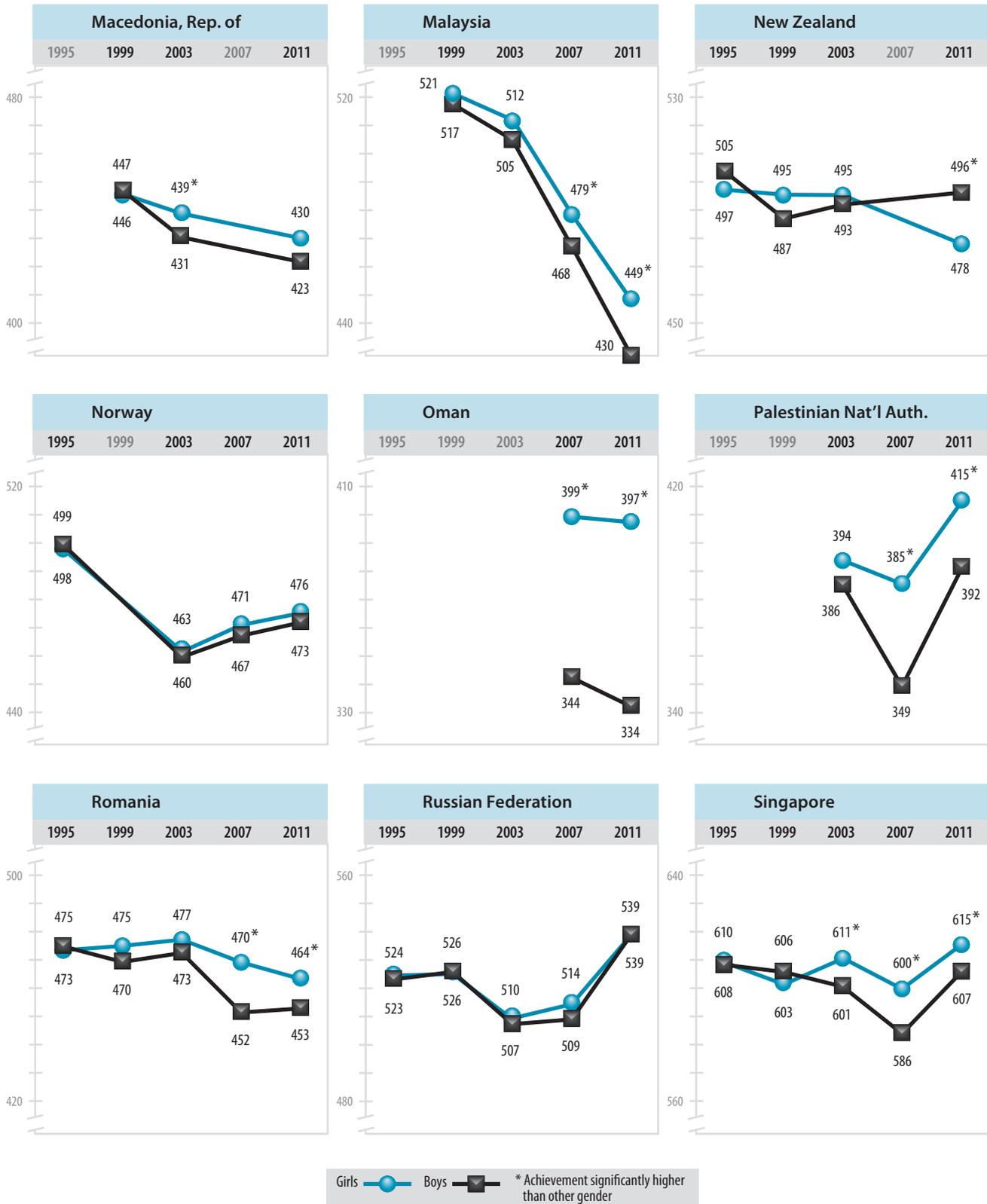
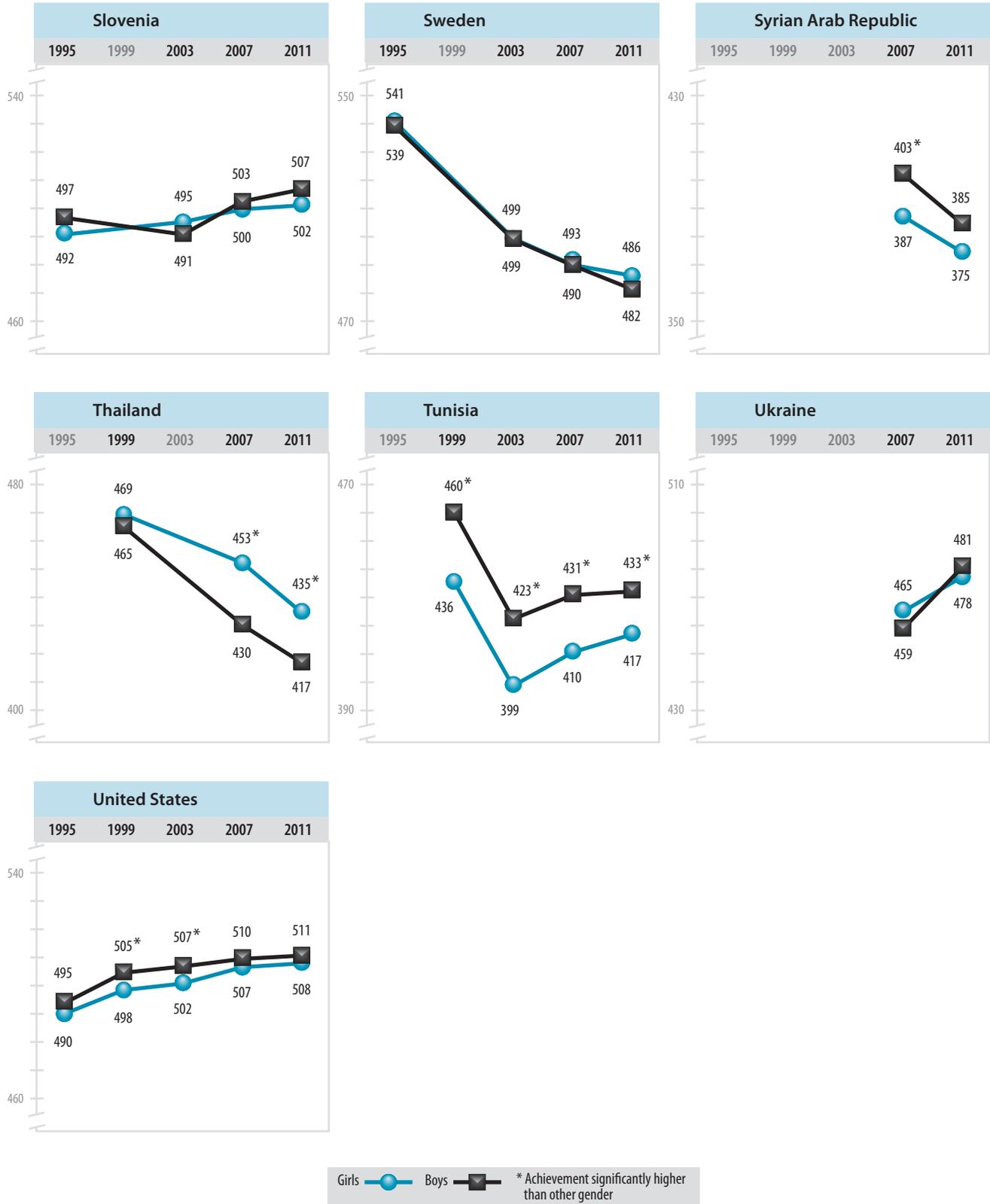


Exhibit 1.13: Trends in Mathematics Achievement by Gender (Continued)



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

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

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

Benchmarking Participants

