CHAPTER 1

Student Achievement in Mathematics

Chapter 1 summarizes eighth-grade achievement on the TIMSS 1999 mathematics assessment for each of the Benchmarking states, districts, and consortia, as well as for each participating country. Comparisons of participants' performance against international benchmarks, as well as gender differences in performance, are also provided.

How Do Participants Differ in Mathematics Achievement?

Exhibit 1.1 presents the distribution of student achievement for the 38 TIMSS 1999 countries and the 27 Benchmarking participants in a two-page display.[1] The left-hand page shows countries and Benchmarking participants together, in decreasing order of average (mean) scale score, and indicates whether the average for each participant is significantly higher or lower than the international average of 487. The international average was obtained by averaging across the mean scores for each of the 38 participating countries. On the right-hand page is a tabular display of average achievement, along with the number of years of formal schooling and the average age of students tested.

Many of the Benchmarking participants performed fairly well on the TIMSS 1999 mathematics assessment. Average performance for the 13 Benchmarking states was clustered in the middle of the international distribution of results for the 38 countries. All of the Benchmarking states performed either significantly above or similar to the international average. The United States as a whole also had average mathematics achievement just above the international average.

The Benchmarking Study underscores the extreme importance of looking beyond the averages to the range of performance found across the nation. Performance across the participating school districts and consortia reflected nearly the full range of achievement internationally. The two highest-achieving Benchmarking participants were the Naperville School District and the First in the World Consortium. These were two of the Benchmarking participants with the lowest percentages of students from low-income families (Naperville, 2 percent; First in the World, 14 percent).[2] Benchmarking participants with the lowest average mathematics achievement included four urban school districts with high percentages of students from low-income families – the Jersey City Public Schools (89 percent), the Chicago Public Schools (71 percent), the Rochester City School District (73 percent), and the Miami-Dade County Public Schools (59 percent). Although not quite as high as Singapore, Korea, and Chinese Taipei nor as low as the lowest-scoring countries in TIMSS 1999, the range of average performance across the Benchmarking districts and consortia was almost as broad as across all the TIMSS 1999 countries.

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

2 Low-income figures are percentages of students eligible to receive free or reduced-price lunch through the National School Lunch Program, as reported by participating schools.

That achievement is distributed broadly within as well as across participating entities is graphically illustrated in Exhibit 1.1 showing the distribution of student performance within each entity. Achievement for each participant is shown for the 25th and 75th percentiles as well as for the 5th and 95th percentiles.[3] Each percentile point indicates the percentages of students performing below and above that point on the scale. For example, 25 percent of the eighth-grade students in each participating entity performed below the 25th percentile for that entity, and 75 percent performed above the 25th percentile. The range between the 25th and 75th percentiles represents performance by the middle half of students. In most entities, the range of performance for the middle group was between 100 and 150 scale-score points. Performance at the 5th and 95th percentiles represents the extremes in both lower and higher achievement. The range of performance between these two score points, which includes 90 percent of the population, is between 250 and 300 points for most participants. The dark boxes at the midpoints of the distributions show the 95 percent confidence intervals around the average achievement in each entity.[4]

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

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

The data in Exhibit 1.2 reinforce the point that, when ordered by average achievement, adjacent participants usually did not significantly differ from each other, although the differences in achievement between the high-performing and low-performing participants were very large.

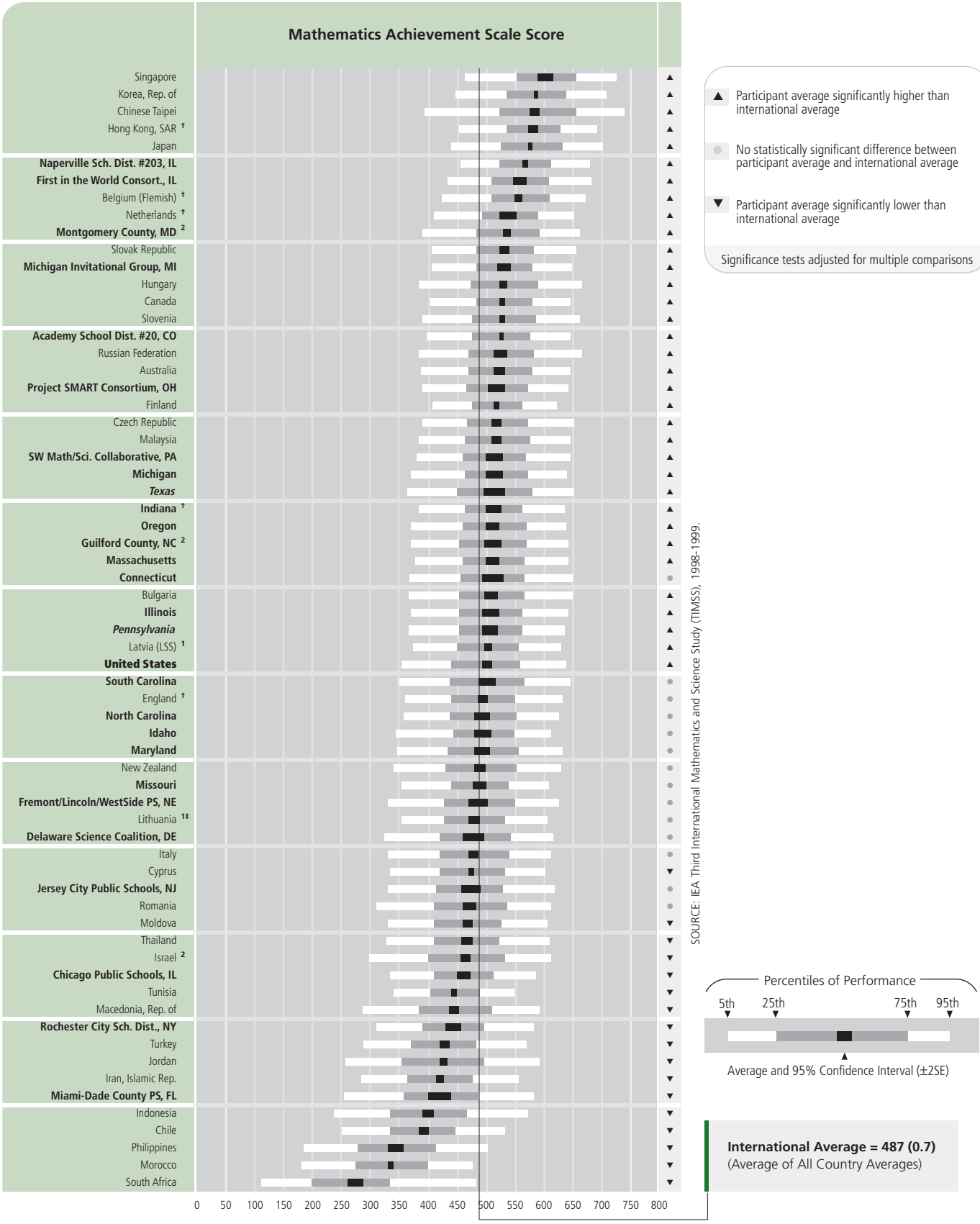
3 Tables of the percentile values and standard deviations for all participants are presented in Appendix C.

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

Singapore, Korea, Chinese Taipei, and Hong Kong had the highest performance, closely followed by Japan, the Naperville School District, the First in the World Consortium, and Belgium (Flemish).[5] Naperville and First in the World both performed similarly to Hong Kong, Japan, and Belgium (Flemish), but significantly below Singapore, Korea, and Chinese Taipei. The difference in performance from one participant to the next was often negligible. Montgomery County, the Michigan Invitational Group, the Academy School District, the Project SMART Consortium, the Southwest Pennsylvania Math and Science Collaborative, Michigan, Texas, Indiana, Oregon, Guilford County, Massachusetts, Connecticut, and Illinois were outperformed by only the top-performing eight or nine entities. These Benchmarking jurisdictions had average achievement most similar to the Netherlands, the Slovak Republic, Hungary, Canada, Slovenia, the Russian Federation, Australia, Finland, the Czech Republic, and Malaysia. Pennsylvania and South Carolina had achievement similar to that of Latvia (LSS),[6] the United States, and England, closely followed by North Carolina, Idaho, Maryland, Missouri, and the Fremont/Lincoln/Westside Public Schools. The Delaware Science Coalition and the Jersey City Public Schools had average achievement similar to that of Italy, outperforming eleven and nine of the TIMSS 1999 countries, respectively. The Chicago Public Schools had average achievement close to that in Moldova, Thailand, and Israel. The Rochester City School District and the Miami-Dade County Public Schools had average eighth-grade mathematics performance lower than most of the TIMSS 1999 countries. Rochester had performance similar to the Republic of Macedonia, but significantly higher than Indonesia and Chile. Miami-Dade had average achievement about the same as the Islamic Republic of Iran, but significantly higher than the three lowest-scoring countries (the Philippines, Morocco, and South Africa).

5 Belgium has two separate educational systems, Flemish and French. The Flemish system participated in TIMSS 1999.

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


Mathematics Achievement Scale Score
Singapore
Korea, Rep. of
Chinese Taipei
Hong Kong, SAR †
Japan
Naperville Sch. Dist. #203, IL
First in the World Consort., IL
Belgium (Flemish) †
Netherlands †
Montgomery County, MD 2
Slovak Republic
Michigan Invitational Group, MI
Hungary
Canada
Slovenia
Academy School Dist. #20, CO
Russian Federation
Australia
Project SMART Consortium, OH
Finland
Czech Republic
Malaysia
SW Math/Sci. Collaborative, PA
Michigan
Texas
Indiana †
Oregon
Guilford County, NC 2
Massachusetts
Connecticut
Bulgaria
Illinois
Pennsylvania
Latvia (LSS) 1
United States
South Carolina
England †
North Carolina
Idaho
Maryland
New Zealand
Missouri
Fremont/Lincoln/WestSide PS, NE
Lithuania 1‡
Delaware Science Coalition, DE
Italy
Cyprus
Jersey City Public Schools, NJ
Romania
Moldova
Thailand
Israel 2
Chicago Public Schools, IL
Tunisia
Macedonia, Rep. of
Rochester City Sch. Dist., NY
Turkey
Jordan
Iran, Islamic Rep.
Miami-Dade County PS, FL
Indonesia
Chile
Philippines
Morocco
South Africa
0 50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800
▲ Participant average significantly higher than international average
● No statistically significant difference between participant average and international average
▼ Participant average significantly lower than international average
Significance tests adjusted for multiple comparisons
SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.
Percentiles of Performance
5th 25th 75th 95th
Average and 95% Confidence Interval (±2SE)
International Average = 487 (0.7)
(Average of All Country Averages)

	Average Scale Score	Years of Formal Schooling	Average Age
Countries			
United States	▲ 502 (4.0)	8	14.2
Australia	▲ 525 (4.8)	8 or 9	14.3
Belgium (Flemish) †	▲ 558 (3.3)	8	14.1
Bulgaria	▲ 511 (5.8)	8	14.8
Canada	▲ 531 (2.5)	8	14.0
Chile	▼ 392 (4.4)	8	14.4
Chinese Taipei	▲ 585 (4.0)	8	14.2
Cyprus	▼ 476 (1.8)	8	13.8
Czech Republic	▲ 520 (4.2)	8	14.4
England †	● 496 (4.1)	9	14.2
Finland	▲ 520 (2.7)	7	13.8
Hong Kong, SAR †	▲ 582 (4.3)	8	14.2
Hungary	▲ 532 (3.7)	8	14.4
Indonesia	▼ 403 (4.9)	8	14.6
Iran, Islamic Rep.	▼ 422 (3.4)	8	14.6
Israel [2]	▼ 466 (3.9)	8	14.1
Italy	● 479 (3.8)	8	14.0
Japan	▲ 579 (1.7)	8	14.4
Jordan	▼ 428 (3.6)	8	14.0
Korea, Rep. of	▲ 587 (2.0)	8	14.4
Latvia (LSS) [1]	▲ 505 (3.4)	8	14.5
Lithuania [1‡]	● 482 (4.3)	8.5	15.2
Macedonia, Rep. of	▼ 447 (4.2)	8	14.6
Malaysia	▲ 519 (4.4)	8	14.4
Moldova	▼ 469 (3.9)	9	14.4
Morocco	▼ 337 (2.6)	7	14.2
Netherlands †	▲ 540 (7.1)	8	14.2
New Zealand	● 491 (5.2)	8.5 to 9.5	14.0
Philippines	▼ 345 (6.0)	7	14.1
Romania	● 472 (5.8)	8	14.8
Russian Federation	▲ 526 (5.9)	7 or 8	14.1
Singapore	▲ 604 (6.3)	8	14.4
Slovak Republic	▲ 534 (4.0)	8	14.3
Slovenia	▲ 530 (2.8)	8	14.8
South Africa	▼ 275 (6.8)	8	15.5
Thailand	▼ 467 (5.1)	8	14.5
Tunisia	▼ 448 (2.4)	8	14.8
Turkey	▼ 429 (4.3)	8	14.2
International Avg. (All Countries)	487 (0.7)		

	Average Scale Score	Years of Formal Schooling	Average Age
States			
Connecticut	● 512 (9.1)	8	14.0
Idaho	● 495 (7.4)	8	14.2
Illinois	▲ 509 (6.7)	8	14.2
Indiana †	▲ 515 (7.2)	8	14.4
Maryland	● 495 (6.2)	8	13.9
Massachusetts	▲ 513 (5.9)	8	14.1
Michigan	▲ 517 (7.5)	8	14.1
Missouri	● 490 (5.3)	8	14.3
North Carolina	● 495 (7.0)	8	14.2
Oregon	▲ 514 (6.0)	8	14.2
Pennsylvania	▲ 507 (6.3)	8	14.2
South Carolina	● 502 (7.4)	8	14.2
Texas	▲ 516 (9.1)	8	14.3
Districts and Consortia			
Academy School Dist. #20, CO	▲ 528 (1.8)	8	14.2
Chicago Public Schools, IL	▼ 462 (6.1)	8	14.2
Delaware Science Coalition, DE	● 479 (8.9)	8	14.1
First in the World Consort., IL	▲ 560 (5.8)	8	14.2
Fremont/Lincoln/WestSide PS, NE	● 488 (8.2)	8	14.2
Guilford County, NC [2]	▲ 514 (7.7)	8	14.2
Jersey City Public Schools, NJ	● 475 (8.6)	8	14.3
Miami-Dade County PS, FL	▼ 421 (9.4)	8	14.3
Michigan Invitational Group, MI	▲ 532 (5.8)	8	14.1
Montgomery County, MD [2]	▲ 537 (3.5)	8	14.0
Naperville Sch. Dist. #203, IL	▲ 569 (2.8)	8	14.1
Project SMART Consortium, OH	▲ 521 (7.5)	8	14.2
Rochester City Sch. Dist., NY	▼ 444 (6.5)	8	14.2
SW Math/Sci. Collaborative, PA	▲ 517 (7.5)	8	14.2

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

▲ Participant average significantly higher than international average

● No statistically significant difference between participant average and international average

▼ Participant average significantly lower than international average

Significance tests adjusted for multiple comparisons

States in *italics* did not fully satisfy guidelines for sample participation rates (see Appendix A for details).

† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.6).

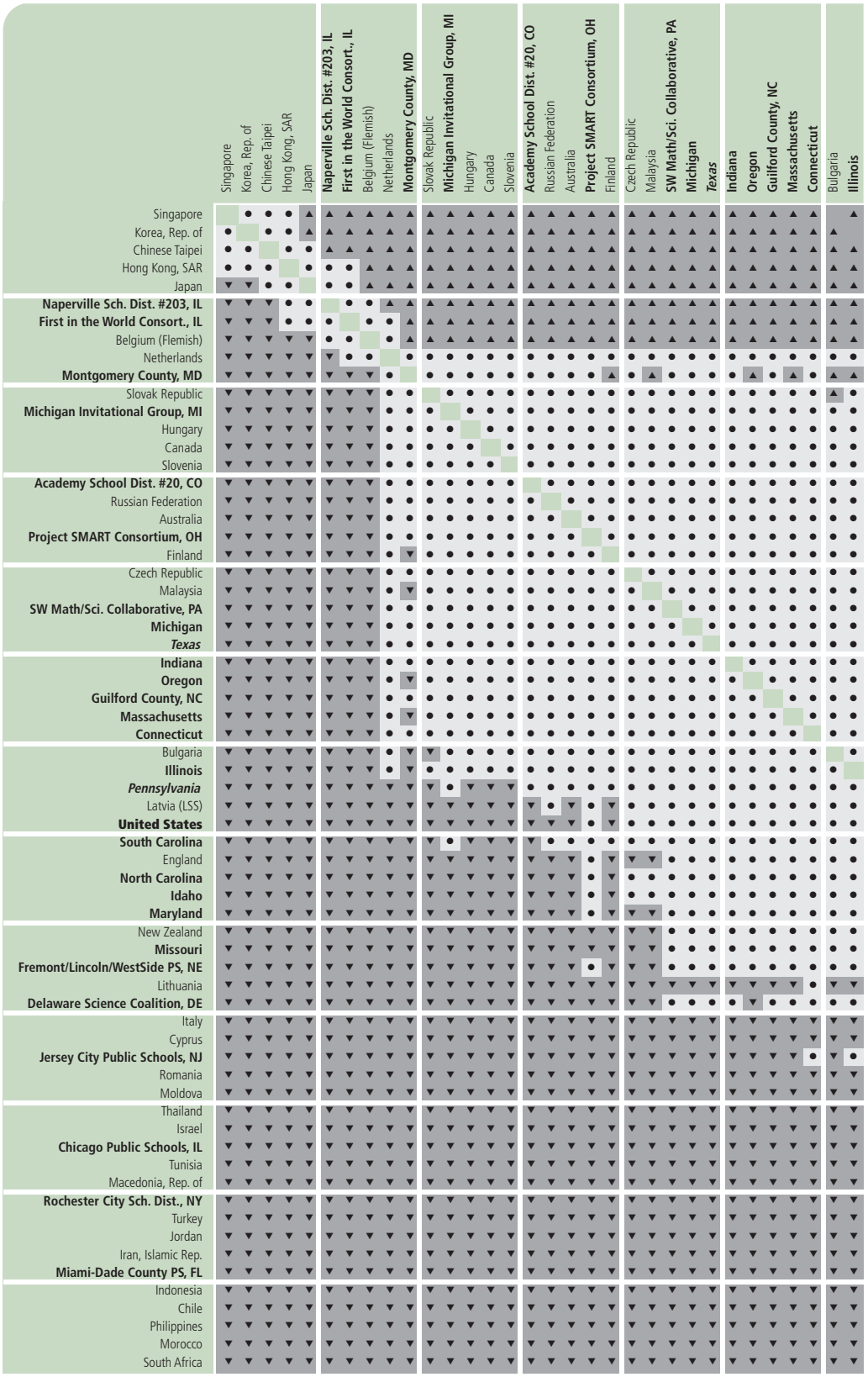
1 National Desired Population does not cover all of International Desired Population (see Exhibit A.3). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

2 National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.3).

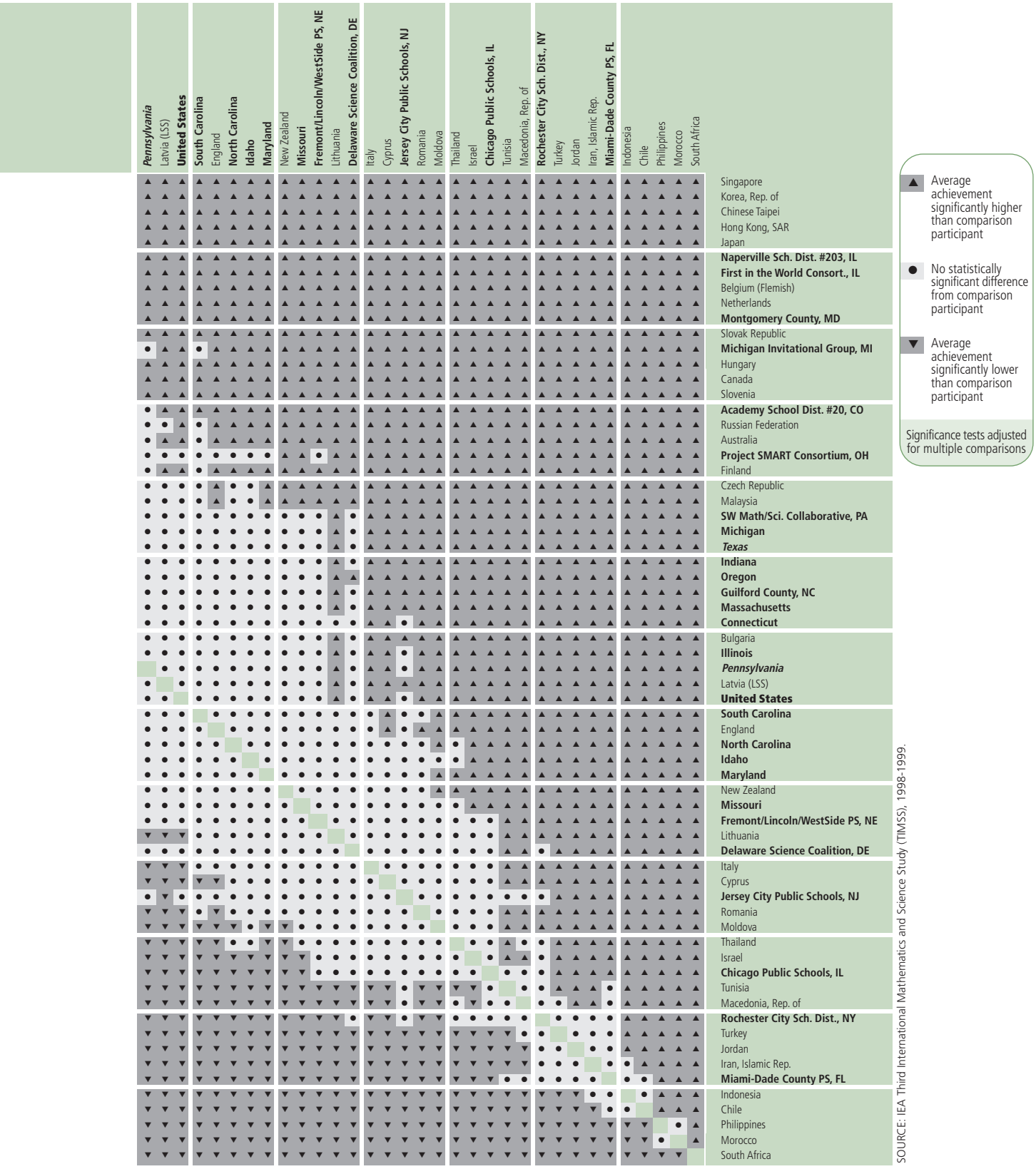
‡ Lithuania tested the same cohort of students as other countries, but later in 1999, at the beginning of the next school year.

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

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



States in *italics* did not fully satisfy guidelines for sample participation rates (see Appendix A for details).



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

How Do Benchmarking Participants Compare with International Benchmarks of Mathematics Achievement?

The TIMSS mathematics achievement scale summarizes student performance on test items designed to measure a wide range of student knowledge and proficiency. In order to provide descriptions of what performance could mean in terms of the mathematics that students know and can do, TIMSS identified four points on the scale for use as international benchmarks[7] or reference points, and conducted an ambitious scale anchoring exercise to describe students' performance at these benchmarks. Exhibit 1.3 shows the four international benchmarks of mathematics achievement and briefly describes what students scoring at these benchmarks typically know and can do. More detailed descriptions appear in Chapter 2, together with example test items illustrating performance at each benchmark.

The Top 10% Benchmark is defined at the 90th percentile on the TIMSS mathematics scale, taking into account the performance of all students in all countries participating in 1999. It corresponds to a scale score of 616 and is the point above which the top 10 percent of students in the TIMSS 1999 assessment scored. Students performing at this level demonstrated that they could organize information, make generalizations, and explain solution strategies in non-routine problem-solving situations.

The Upper Quarter Benchmark is the 75th percentile on the mathematics scale. This point, corresponding to a scale score of 555, is the point above which the top 25 percent of students scored. Students scoring at this benchmark demonstrated that they could apply their mathematical understanding and knowledge in a wide variety of relatively complex situations involving fractions, decimals, geometric properties, and algebraic expressions.

The Median Benchmark, with a score of 479, corresponds to the 50th percentile, or median. This is the point above which the top half of students scored on the TIMSS 1999 assessment. Students performing at this level showed that they could apply basic mathematical knowledge in straightforward situations, such as one-step word problems involving addition and subtraction or computational problems based on basic properties of geometric figures and simple algebraic relationships.

[7] Readers should be careful not to confuse the international benchmarks, which are points on the international mathematics achievement scale chosen to describe specific achievement levels, with the benchmarking exercise itself, which is a process by which participants compare their achievement, curriculum, and instructional practices with those of the best in the world.

The Lower Quarter Benchmark is the 25th percentile and corresponds to a scale score of 396. This score point is reached by the top 75 percent of students and may be used as a benchmark of performance for lower-achieving students. Students scoring at this level typically demonstrated computational facility with whole numbers.

Exhibit 1.4 displays the percentage of students in each participating entity that reached each international benchmark, in decreasing order by the percentage reaching the Top 10% Benchmark. If student achievement in mathematics were distributed alike in every entity, then each entity would be expected to have about 10 percent of its students reaching the Top 10% Benchmark, 25 percent the Upper Quarter Benchmark, 50 percent the Median Benchmark, and 75 percent the Lower Quarter Benchmark. Although countries such as New Zealand, and Benchmarking participants such as Maryland, North Carolina, and the Delaware Science Coalition, came fairly close, no entity followed this pattern exactly. Instead, the high-performing entities generally had greater percentages of students reaching each benchmark, and the low-performing entities had lesser percentages.

Among the high performers, for example, Singapore, Chinese Taipei, Korea, Hong Kong, and Japan had one-third or more of their students reaching the Top 10% Benchmark, about two-thirds reaching the Upper Quarter Benchmark, around 90 percent reaching the Median Benchmark, and almost all (95 to 99 percent) reaching the Lower Quarter Benchmark. In comparison, the Naperville School District and the First in the World Consortium had 24 and 22 percent of their students, respectively, reaching the Top 10% Benchmark and 59 and 56 percent, respectively, reaching the Upper Quarter Benchmark, somewhat less than in the high-performing Asian countries. More like the top-performing Asian countries, these two high-performing districts had close to 90 percent of their students reaching the Median Benchmark (91 and 87 percent, respectively) and nearly all of their students reaching the Lower Quarter Benchmark (99 and 98 percent, respectively).

In contrast, the three lowest-performing Benchmarking participants, all urban districts, had two percent of their students reaching the Top 10% Benchmark, 9 to 12 percent reaching the Upper Quarter Benchmark, and from 29 to 41 percent reaching the Median Benchmark. The lowest-performing countries of South Africa, the

Philippines, and Morocco had almost no students reaching the Top 10% Benchmark, no more than one percent reaching the Upper Quarter Benchmark, less than 10 percent reaching the Median Benchmark, and no more than 31 percent reaching the Lower Quarter Benchmark.

Although Exhibit 1.4 is organized to draw particular attention to the percentage of high-achieving students in each entity, it conveys information about the distribution of middle and low performers also. For example, Canada, Australia, and Malaysia had 12 percent of their students reaching the Top 10% Benchmark, as might be expected, but 94 to 96 percent (rather than 75 percent) reaching the Lower Quarter Benchmark. Similarly, the Academy School District, the Michigan Invitational Group, and the Project SMART Consortium had 11 to 12 percent of their students reaching the Top 10% Benchmark but 95 to 96 percent reaching the Lower Quarter Benchmark.

• Top 10% Benchmark

Students can organize information, make generalizations, and explain solution strategies in non-routine problem solving situations. They can organize information and make generalizations to solve problems; apply knowledge of numeric, geometric, and algebraic relationships to solve problems (e.g., among fractions, decimals, and percents; geometric properties; and algebraic rules); and find the equivalent forms of algebraic expressions.

90th Percentile: 616

• Upper Quarter Benchmark

Students can apply their understanding and knowledge in a wide variety of relatively complex situations. They can order, relate and compute with fractions and decimals to solve word problems; solve multi-step word problems involving proportions with whole numbers; solve probability problems; use knowledge of geometric properties to solve problems; identify and evaluate algebraic expressions and solve equations with one variable.

75th Percentile: 555

• Median Benchmark

Students can apply basic mathematical knowledge in straightforward situations. They can add or subtract to solve one-step word problems involving whole numbers and decimals; identify representations of common fractions and relative sizes of fractions; solve for missing terms in proportions; recognize basic notions of percents and probability; use basic properties of geometric figures; read and interpret graphs, tables, and scales; and understand simple algebraic relationships.

50th Percentile: 479

• Lower Quarter Benchmark

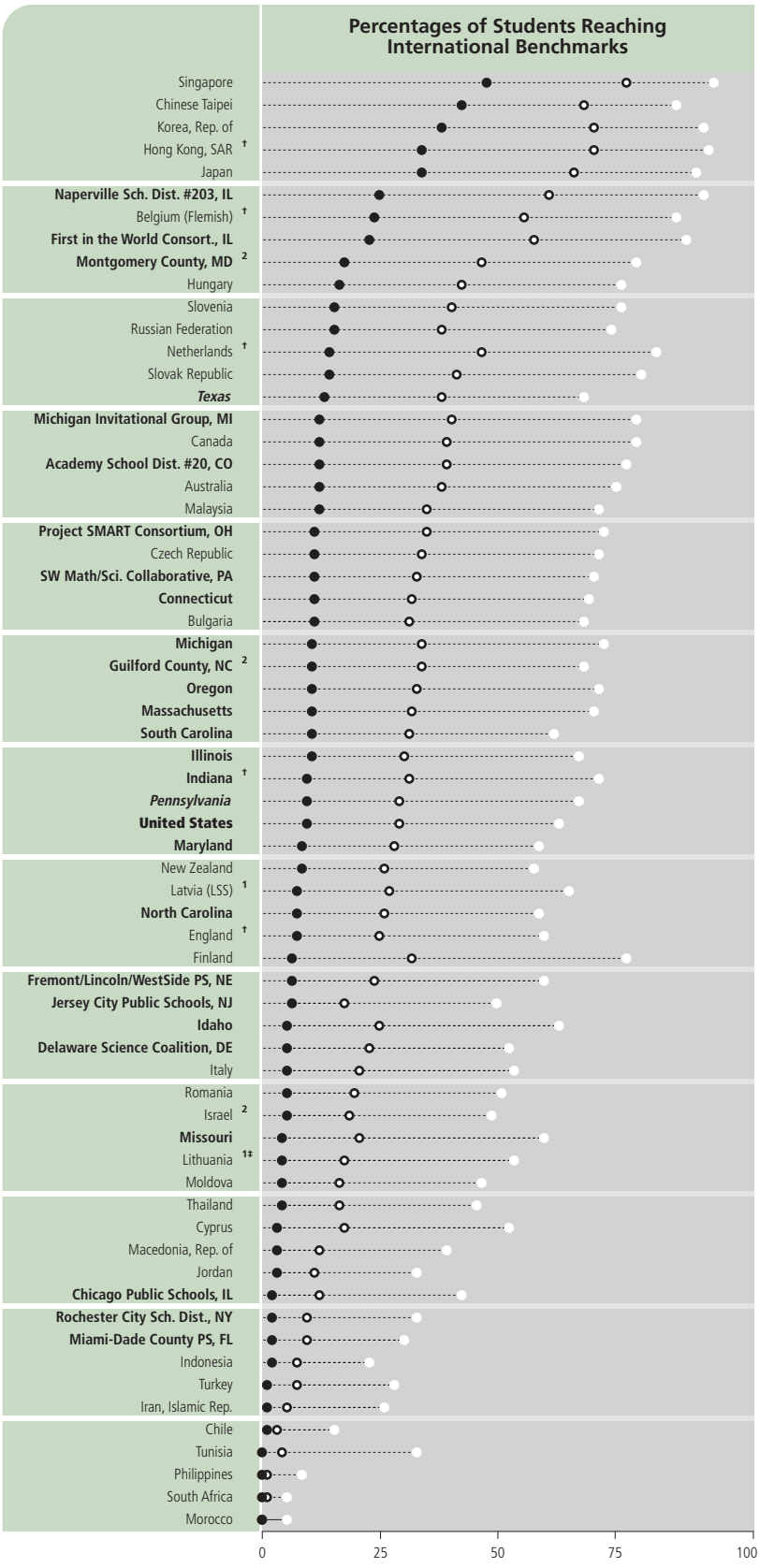
Students can do basic computations with whole numbers. The few items that anchor at this level provide some evidence that students can add, subtract, and round with whole numbers. When there are the same number of decimal places, they can subtract with multiple regrouping. Students can round whole numbers to the nearest hundred. They recognize some basic notation and terminology.

25th Percentile: 396

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

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

Percentages of Students Reaching TIMSS 1999 International Benchmarks of Mathematics Achievement



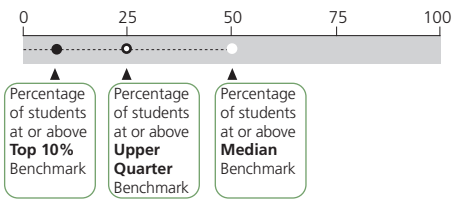
SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.

Top 10% Benchmark (90th Percentile) = 616

Upper Quarter Benchmark (75th Percentile) = 555

Median Benchmark (50th Percentile) = 479

Lower Quarter Benchmark (25th Percentile) = 396



	Top 10%	Upper Quarter	Median	Lower Quarter
Countries				
United States	9 (1.0)	28 (1.6)	61 (1.9)	88 (1.0)
Australia	12 (1.8)	37 (2.7)	73 (2.4)	94 (0.8)
Belgium (Flemish) †	23 (1.5)	54 (1.7)	85 (1.2)	98 (0.6)
Bulgaria	11 (2.3)	30 (3.0)	66 (2.6)	91 (1.3)
Canada	12 (1.1)	38 (1.5)	77 (1.3)	96 (0.6)
Chile	1 (0.5)	3 (1.1)	15 (1.8)	48 (2.0)
Chinese Taipei	41 (1.7)	66 (1.5)	85 (1.0)	95 (0.6)
Cyprus	3 (0.4)	17 (0.8)	51 (1.1)	84 (0.8)
Czech Republic	11 (1.4)	33 (2.1)	69 (2.3)	94 (1.1)
England †	7 (0.9)	24 (1.9)	58 (2.1)	89 (1.3)
Finland	6 (0.9)	31 (1.7)	75 (1.5)	96 (0.5)
Hong Kong, SAR †	33 (2.3)	68 (2.4)	92 (1.5)	99 (0.6)
Hungary	16 (1.2)	41 (1.9)	74 (1.6)	94 (1.0)
Indonesia	2 (0.4)	7 (0.9)	22 (1.4)	52 (2.2)
Iran, Islamic Rep.	1 (0.2)	5 (0.8)	25 (1.7)	63 (1.5)
Israel [2]	5 (0.6)	18 (1.3)	47 (1.8)	77 (1.9)
Italy	5 (0.7)	20 (1.4)	52 (2.1)	83 (1.4)
Japan	33 (1.1)	64 (0.9)	89 (0.5)	98 (0.3)
Jordan	3 (0.5)	11 (0.9)	32 (1.5)	62 (1.4)
Korea, Rep. of	37 (1.0)	68 (0.9)	91 (0.5)	99 (0.2)
Latvia (LSS) [1]	7 (0.9)	26 (1.8)	63 (2.0)	92 (1.0)
Lithuania [1‡]	4 (0.7)	17 (2.0)	52 (2.4)	86 (1.8)
Macedonia, Rep. of	3 (0.4)	12 (1.0)	38 (1.9)	72 (1.8)
Malaysia	12 (1.4)	34 (2.4)	69 (2.2)	94 (0.8)
Moldova	4 (0.7)	16 (1.5)	45 (2.2)	81 (1.7)
Morocco	0 (0.0)	0 (0.2)	5 (0.4)	27 (1.1)
Netherlands †	14 (2.3)	45 (4.1)	81 (3.5)	96 (1.3)
New Zealand	8 (1.2)	25 (2.4)	56 (2.5)	85 (1.5)
Philippines	0 (0.1)	1 (0.5)	8 (1.4)	31 (2.5)
Romania	5 (1.1)	19 (1.9)	49 (2.6)	80 (2.1)
Russian Federation	15 (1.8)	37 (2.8)	72 (2.7)	94 (1.2)
Singapore	46 (3.5)	75 (2.7)	93 (1.3)	99 (0.3)
Slovak Republic	14 (1.4)	40 (2.3)	78 (1.8)	96 (0.6)
Slovenia	15 (1.2)	39 (1.4)	74 (1.4)	95 (0.7)
South Africa	0 (0.2)	1 (0.4)	5 (1.0)	14 (2.0)
Thailand	4 (0.8)	16 (1.8)	44 (2.6)	81 (1.6)
Tunisia	0 (0.1)	4 (0.5)	32 (1.6)	80 (1.3)
Turkey	1 (0.3)	7 (1.0)	27 (1.9)	65 (2.0)

	Top 10%	Upper Quarter	Median	Lower Quarter
States				
Connecticut	11 (2.5)	31 (3.9)	67 (4.4)	91 (1.9)
Idaho	5 (1.1)	24 (2.9)	61 (3.5)	88 (2.2)
Illinois	10 (1.6)	29 (2.9)	65 (3.3)	92 (1.5)
Indiana †	9 (1.9)	30 (3.9)	69 (3.6)	94 (1.2)
Maryland	8 (1.4)	27 (2.5)	57 (3.2)	87 (2.0)
Massachusetts	10 (1.6)	31 (2.6)	68 (3.0)	92 (1.6)
Michigan	10 (2.0)	33 (3.7)	70 (3.3)	92 (1.7)
Missouri	4 (0.9)	20 (2.4)	58 (2.9)	89 (1.5)
North Carolina	7 (1.6)	25 (3.1)	57 (3.3)	88 (2.0)
Oregon	10 (1.8)	32 (2.8)	69 (2.8)	91 (1.4)
Pennsylvania	9 (1.3)	28 (2.6)	65 (3.0)	91 (1.8)
South Carolina	10 (2.0)	30 (3.2)	60 (3.5)	88 (1.8)
Texas	13 (2.2)	37 (3.8)	66 (4.3)	90 (2.1)
Districts and Consortia				
Academy School Dist. #20, CO	12 (0.8)	38 (1.5)	75 (1.5)	95 (0.7)
Chicago Public Schools, IL	2 (0.9)	12 (1.7)	41 (4.3)	81 (2.5)
Delaware Science Coalition, DE	5 (1.8)	22 (4.1)	51 (4.5)	83 (2.4)
First in the World Consort., IL	22 (3.2)	56 (3.3)	87 (2.1)	98 (0.6)
Fremont/Lincoln/WestSide PS, NE	6 (2.3)	23 (4.1)	58 (4.0)	84 (2.7)
Guilford County, NC [2]	10 (2.2)	33 (3.5)	66 (4.1)	91 (1.6)
Jersey City Public Schools, NJ	6 (1.9)	17 (3.4)	48 (3.9)	82 (2.9)
Miami-Dade County PS, FL	2 (0.9)	9 (2.4)	29 (3.6)	61 (3.5)
Michigan Invitational Group, MI	12 (2.4)	39 (3.4)	77 (3.0)	96 (1.3)
Montgomery County, MD [2]	17 (2.2)	45 (1.8)	77 (1.4)	95 (1.1)
Naperville Sch. Dist. #203, IL	24 (1.7)	59 (2.2)	91 (1.1)	99 (0.4)
Project SMART Consortium, OH	11 (2.9)	34 (4.7)	70 (3.1)	95 (1.0)
Rochester City Sch. Dist., NY	2 (0.9)	9 (2.5)	32 (3.2)	73 (2.9)
SW Math/Sci. Collaborative, PA	11 (2.7)	32 (3.9)	68 (3.1)	93 (1.6)

Top 10% Benchmark (90th Percentile) = 616
Upper Quarter Benchmark (75th Percentile) = 555
Median Benchmark (50th Percentile) = 479
Lower Quarter Benchmark (25th Percentile) = 396

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

States in *italics* did not fully satisfy guidelines for sample participation rates (see Appendix A for details).

† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.6).

1 National Desired Population does not cover all of International Desired Population (see Exhibit A.3). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

2 National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.3).

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

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

What Are the Gender Differences in Mathematics Achievement?

Exhibit 1.5 presents average mathematics achievement separately for girls and boys for each of the participating entities, as well as the difference between the means, in increasing order of the difference. The gender difference for each entity is shown by a bar indicating the amount of the difference, whether its direction favored girls or boys, and whether it is statistically significant (a darkened bar).

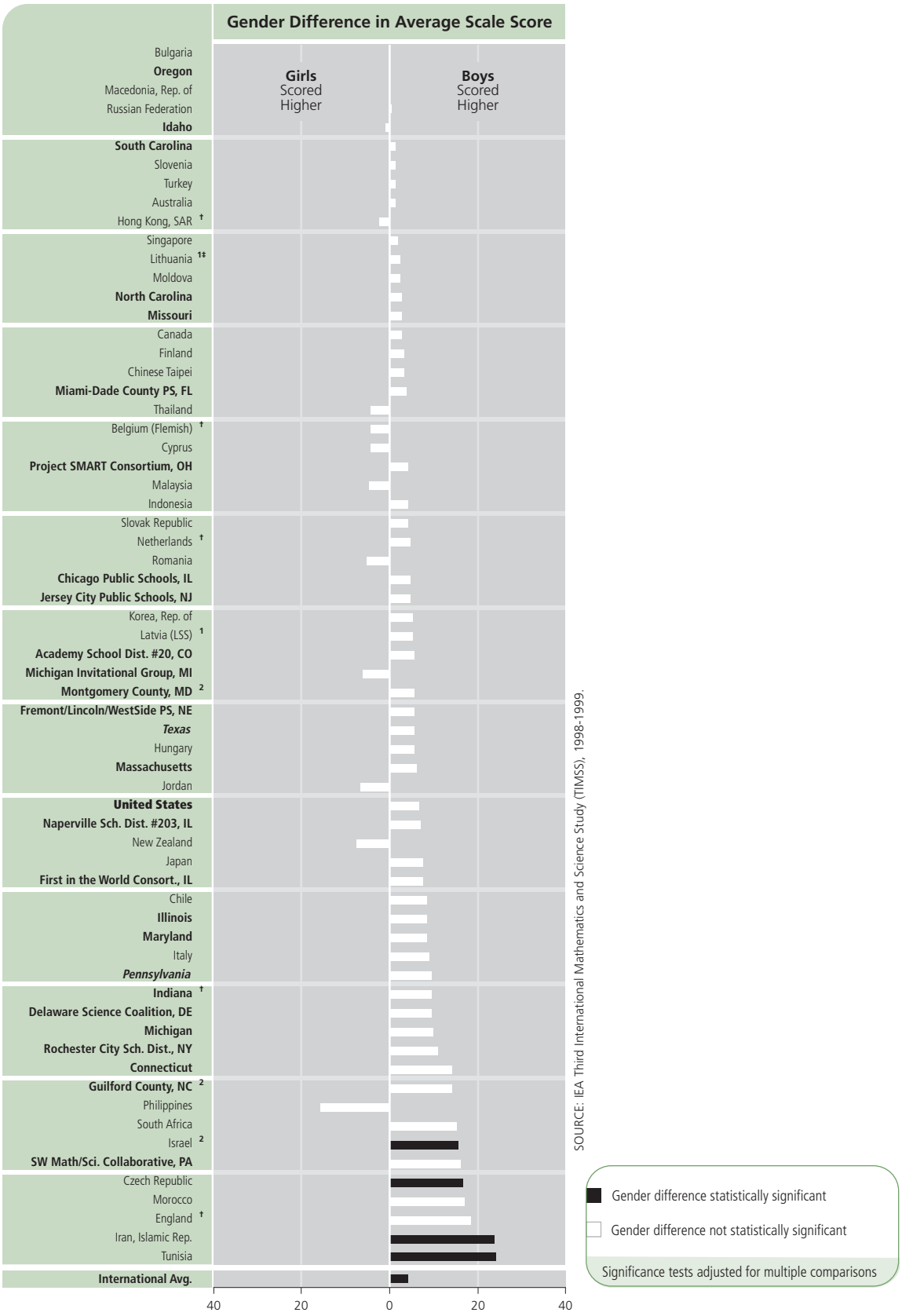
It is good news that in mathematics at the eighth grade, the TIMSS 1999 Benchmarking Study shows relatively equivalent average achievement for girls and boys in each of the Benchmarking jurisdictions. The United States as well as a number of other countries around the world appear to be making progress towards gender equity in mathematics education. On average across all TIMSS 1999 countries, there was a modest but significant difference favoring boys, although this varied considerably from country to country. The only countries with differences large enough to be statistically significant were Israel, the Czech Republic, Iran, and Tunisia.

Although achievement differences between the genders are becoming smaller in mathematics, research indicates that they still exist in those areas involving the most complex mathematical tasks, particularly as students progress to middle and secondary schools.[8] Thus, Exhibit 1.6 provides information on gender differences in mathematics achievement among students with high performance compared with those in the middle of the achievement distribution. For each entity, score levels were computed for the highest-scoring 25 percent of students, called the upper quarter level, and for the highest-scoring 50 percent, called the median level. The percentages of girls and boys in each entity reaching each of the two levels were computed. For equitable performance, 25 percent each of girls and boys should have reached the upper quarter level, and 50 percent the median level.

On average across countries, 23 percent of girls compared with 27 percent of boys reached the upper quarter level, and 49 percent of girls compared with 51 percent of boys reached the median level. These gender differences, although small, were statistically significant. In all but four countries, however, the percentages reaching the upper quarter and median levels were not significantly different, indicating

8 Fennema, E. (1996), "Mathematics, Gender, and Research" in G. Hanna (ed.), *Towards Equity in Mathematics Education*, Dordrecht, the Netherlands: Kluwer Academic Publishers.

that gender equity exists in most countries at these levels. Even though the four countries with significant differences did include the United States (as well as Israel, the Philippines, and Tunisia), this was not reflected in the results for the Benchmarking jurisdictions. Michigan was the only Benchmarking jurisdiction to show a significant gender difference favoring males among high-performing students.


Gender Difference in Average Scale Score
Girls Scored Higher
Boys Scored Higher
Bulgaria
Oregon
Macedonia, Rep. of
Russian Federation
Idaho
South Carolina
Slovenia
Turkey
Australia
Hong Kong, SAR †
Singapore
Lithuania 1‡
Moldova
North Carolina
Missouri
Canada
Finland
Chinese Taipei
Miami-Dade County PS, FL
Thailand
Belgium (Flemish) †
Cyprus
Project SMART Consortium, OH
Malaysia
Indonesia
Slovak Republic
Netherlands †
Romania
Chicago Public Schools, IL
Jersey City Public Schools, NJ
Korea, Rep. of
Latvia (LSS) 1
Academy School Dist. #20, CO
Michigan Invitational Group, MI
Montgomery County, MD 2
Fremont/Lincoln/WestSide PS, NE
Texas
Hungary
Massachusetts
Jordan
United States
Naperville Sch. Dist. #203, IL
New Zealand
Japan
First in the World Consort., IL
Chile
Illinois
Maryland
Italy
Pennsylvania
Indiana †
Delaware Science Coalition, DE
Michigan
Rochester City Sch. Dist., NY
Connecticut
Guilford County, NC 2
Philippines
South Africa
Israel 2
SW Math/Sci. Collaborative, PA
Czech Republic
Morocco
England †
Iran, Islamic Rep.
Tunisia
International Avg.
40
20
0
20
40
SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1998-1999.
Gender difference statistically significant
Gender difference not statistically significant
Significance tests adjusted for multiple comparisons

	Girls' Average Scale Score	Boys' Average Scale Score		Difference (Absolute Value)
Countries				
United States	498 (3.9)	505 (4.8)		7 (3.4)
Australia	524 (5.7)	526 (5.7)		2 (6.0)
Belgium (Flemish) †	560 (7.2)	556 (8.3)		4 (14.2)
Bulgaria	510 (5.9)	511 (6.9)		0 (5.5)
Canada	529 (2.5)	533 (3.2)		3 (2.9)
Chile	388 (4.3)	397 (5.8)		9 (5.5)
Chinese Taipei	583 (3.9)	587 (5.3)		4 (4.6)
Cyprus	479 (2.1)	474 (2.7)		4 (3.3)
Czech Republic	512 (4.0)	528 (5.8)	▲	17 (5.0)
England †	487 (5.4)	505 (5.0)		19 (6.5)
Finland	519 (3.0)	522 (3.5)		3 (3.6)
Hong Kong, SAR †	583 (4.7)	581 (5.9)		2 (6.5)
Hungary	529 (4.0)	535 (4.3)		6 (3.7)
Indonesia	401 (5.4)	405 (5.0)		5 (3.3)
Iran, Islamic Rep.	408 (4.2)	432 (4.8)	▲	24 (6.5)
Israel 2	459 (4.2)	474 (4.8)	▲	16 (4.6)
Italy	475 (4.5)	484 (4.3)		9 (4.2)
Japan	575 (2.4)	582 (2.3)		8 (3.3)
Jordan	431 (4.7)	425 (5.9)		7 (8.1)
Korea, Rep. of	585 (3.1)	590 (2.2)		5 (3.7)
Latvia (LSS) 1	502 (3.8)	508 (4.4)		5 (4.5)
Lithuania 1‡	480 (4.7)	483 (4.8)		3 (4.0)
Macedonia, Rep. of	446 (5.3)	447 (4.3)		0 (4.5)
Malaysia	521 (4.7)	517 (6.0)		5 (6.1)
Moldova	468 (4.1)	471 (4.7)		3 (4.1)
Morocco	326 (5.3)	344 (4.1)		17 (7.7)
Netherlands †	538 (7.6)	542 (7.0)		5 (3.0)
New Zealand	495 (5.5)	487 (7.6)		7 (8.3)
Philippines	352 (6.9)	337 (6.5)		15 (6.1)
Romania	475 (6.3)	470 (6.2)		5 (4.7)
Russian Federation	526 (6.0)	526 (6.4)		1 (3.3)
Singapore	603 (6.1)	606 (7.5)		2 (5.7)
Slovak Republic	532 (4.2)	536 (4.5)		5 (3.6)
Slovenia	529 (3.0)	531 (3.6)		1 (3.6)
South Africa	267 (7.5)	283 (7.3)		16 (5.9)
Thailand	469 (5.7)	465 (5.5)		4 (4.9)
Tunisia	436 (2.4)	460 (2.9)	▲	25 (2.2)
Turkey	428 (4.7)	429 (4.4)		2 (2.8)
International Avg. (All Countries)	485 (0.8)	489 (0.9)	▲	4 (1.1)

	Girls' Average Scale Score	Boys' Average Scale Score	Difference (Absolute Value)
States			
Connecticut	506 (8.9)	520 (9.8)	14 (5.3)
Idaho	495 (7.1)	495 (8.2)	1 (4.3)
Illinois	505 (8.0)	514 (6.1)	9 (4.5)
Indiana †	510 (6.8)	519 (8.0)	10 (3.9)
Maryland	490 (6.4)	499 (6.8)	9 (4.2)
Massachusetts	510 (6.4)	517 (6.0)	6 (3.5)
Michigan	512 (7.2)	522 (8.1)	10 (3.9)
Missouri	488 (5.9)	491 (5.6)	3 (4.5)
North Carolina	494 (7.9)	497 (6.9)	3 (4.9)
Oregon	514 (6.6)	514 (6.9)	0 (6.0)
Pennsylvania	503 (6.2)	512 (7.2)	10 (4.2)
South Carolina	501 (8.0)	502 (7.6)	1 (5.0)
Texas	513 (8.2)	519 (10.7)	6 (5.7)
Districts and Consortia			
Academy School Dist. #20, CO	526 (2.9)	531 (3.4)	6 (5.2)
Chicago Public Schools, IL	460 (6.3)	465 (6.7)	5 (4.5)
Delaware Science Coalition, DE	475 (8.9)	485 (11.1)	10 (9.2)
First in the World Consort., IL	556 (6.7)	564 (6.8)	8 (7.1)
Fremont/Lincoln/WestSide PS, NE	485 (8.3)	491 (10.2)	6 (8.7)
Guilford County, NC 2	507 (8.3)	521 (8.2)	14 (5.8)
Jersey City Public Schools, NJ	472 (8.8)	478 (9.2)	5 (4.6)
Miami-Dade County PS, FL	419 (9.3)	423 (12.1)	4 (10.3)
Michigan Invitational Group, MI	535 (5.4)	529 (7.4)	6 (5.8)
Montgomery County, MD 2	534 (5.5)	540 (4.4)	6 (7.0)
Naperville Sch. Dist. #203, IL	566 (3.3)	573 (3.3)	7 (3.3)
Project SMART Consortium, OH	518 (7.8)	523 (8.1)	4 (5.0)
Rochester City Sch. Dist., NY	439 (7.8)	450 (6.6)	11 (6.2)
SW Math/Sci. Collaborative, PA	509 (7.5)	525 (8.5)	16 (5.3)

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

▲ Significantly higher than other gender

Significance tests adjusted for multiple comparisons

States in *italics* did not fully satisfy guidelines for sample participation rates (see Appendix A for details).

† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.6).

1 National Desired Population does not cover all of International Desired Population (see Exhibit A.3). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

2 National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.3)

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

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

Countries	Upper Quarter		Median	
	Percent of Girls	Percent of Boys	Percent of Girls	Percent of Boys
United States	23 (1.3)	27 (1.9) ▲	49 (2.0)	51 (2.3)
Australia	24 (2.8)	26 (2.6)	49 (3.2)	51 (3.0)
Belgium (Flemish) †	25 (2.5)	25 (2.5)	50 (3.1)	50 (3.5)
Bulgaria	24 (3.1)	26 (3.5)	51 (3.0)	49 (3.2)
Canada	24 (1.2)	26 (1.4)	49 (1.3)	51 (1.9)
Chile	23 (1.9)	27 (2.6)	48 (2.2)	52 (2.4)
Chinese Taipei	22 (1.5)	28 (1.9)	49 (1.9)	51 (2.1)
Cyprus	24 (1.4)	26 (1.4)	50 (1.4)	50 (1.5)
Czech Republic	22 (1.6)	28 (2.5)	46 (2.4)	54 (2.9)
England †	20 (2.7)	30 (2.4)	46 (3.0)	54 (2.7)
Finland	23 (1.8)	27 (2.2)	49 (1.9)	51 (2.2)
Hong Kong, SAR †	24 (2.5)	26 (2.4)	50 (2.9)	50 (3.1)
Hungary	24 (1.9)	26 (1.8)	48 (2.2)	52 (2.1)
Indonesia	25 (1.6)	25 (1.7)	49 (2.1)	52 (2.1)
Iran, Islamic Rep.	19 (2.0)	29 (2.2)	43 (2.5)	55 (2.5)
Israel 2	21 (1.5)	29 (1.7) ▲	47 (2.0)	53 (2.2)
Italy	23 (1.8)	28 (1.7)	47 (2.2)	53 (2.2)
Japan	23 (1.3)	27 (1.1)	47 (1.5)	53 (1.3)
Jordan	24 (1.7)	26 (2.1)	51 (2.0)	49 (2.2)
Korea, Rep. of	24 (1.1)	26 (1.0)	48 (1.5)	52 (1.3)
Latvia (LSS) 1	24 (1.9)	27 (2.1)	49 (2.2)	52 (2.2)
Lithuania 1‡	24 (2.5)	26 (2.3)	50 (2.5)	50 (2.5)
Macedonia, Rep. of	26 (1.8)	24 (1.6)	51 (2.4)	49 (2.0)
Malaysia	26 (2.3)	24 (2.9)	52 (2.6)	48 (3.4)
Moldova	24 (1.6)	27 (2.1)	50 (2.1)	51 (2.2)
Morocco	21 (1.7)	28 (1.5)	45 (2.2)	54 (1.7)
Netherlands †	24 (3.6)	26 (3.2)	48 (4.2)	52 (4.4)
New Zealand	26 (2.6)	24 (3.5)	52 (3.0)	48 (3.5)
Philippines	27 (2.7)	23 (2.5)	53 (2.7) ▲	46 (2.5)
Romania	25 (2.3)	25 (2.4)	51 (2.8)	49 (2.8)
Russian Federation	24 (2.4)	26 (2.5)	49 (2.9)	51 (3.2)
Singapore	23 (3.1)	26 (3.4)	49 (3.6)	51 (4.2)
Slovak Republic	23 (2.0)	27 (2.2)	48 (2.6)	52 (2.7)
Slovenia	24 (1.6)	26 (1.5)	49 (1.7)	51 (2.0)
South Africa	23 (2.7)	27 (2.3)	47 (2.5)	53 (2.1)
Thailand	25 (2.6)	24 (2.4)	50 (2.9)	50 (2.7)
Tunisia	19 (1.4)	31 (1.6) ▲	42 (1.7)	59 (1.6) ▲
Turkey	25 (1.8)	25 (1.9)	50 (2.2)	50 (1.8)
International Avg. (All Countries)	23 (0.4)	27 (0.4) ▲	49 (0.4)	51 (0.4) ▲

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

▲ Significantly greater percentage than other gender

Significance tests adjusted for multiple comparisons

States in *italics* did not fully satisfy guidelines for sample participation rates (see Appendix A for details).

† Met guidelines for sample participation rates only after replacement schools were included (see Exhibit A.6).

1 National Desired Population does not cover all of International Desired Population (see Exhibit A.3). Because coverage falls below 65%, Latvia is annotated LSS for Latvian-Speaking Schools only.

2 National Defined Population covers less than 90 percent of National Desired Population (see Exhibit A.3).

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

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

	Upper Quarter		Median	
	Percent of Girls	Percent of Boys	Percent of Girls	Percent of Boys
States				
Connecticut	21 (3.1)	29 (3.9)	47 (4.7)	53 (4.4)
Idaho	24 (3.0)	26 (3.0)	49 (3.5)	51 (4.1)
Illinois	23 (3.1)	27 (2.9)	48 (3.7)	52 (3.1)
Indiana †	22 (3.6)	28 (3.7)	47 (4.1)	53 (5.1)
Maryland	22 (2.6)	28 (2.6)	48 (3.4)	52 (3.2)
Massachusetts	23 (2.7)	27 (2.7)	48 (3.4)	52 (3.0)
Michigan	22 (3.3)	29 (3.6) ▲	48 (4.3)	52 (3.6)
Missouri	23 (2.7)	27 (2.7)	49 (3.3)	51 (2.5)
North Carolina	24 (3.5)	26 (2.8)	49 (3.6)	51 (3.5)
Oregon	24 (2.7)	27 (2.8)	49 (3.2)	51 (3.5)
Pennsylvania	22 (3.0)	28 (2.9)	48 (3.2)	52 (3.6)
South Carolina	24 (3.2)	27 (3.2)	49 (3.8)	51 (3.3)
Texas	22 (3.1)	28 (3.7)	48 (4.4)	52 (4.7)
Districts and Consortia				
Academy School Dist. #20, CO	22 (1.6)	28 (1.9)	48 (2.3)	52 (2.1)
Chicago Public Schools,	23 (2.9)	27 (3.6)	50 (4.3)	51 (3.5)
Delaware Science Coalition, DE	22 (4.3)	29 (5.2)	47 (4.9)	53 (5.1)
First in the World Consort., IL	22 (3.8)	28 (3.7)	49 (3.6)	51 (3.9)
Fremont/Lincoln/WestSide PS, NE	24 (3.7)	26 (4.7)	50 (4.0)	50 (4.1)
Guilford County, NC [2]	22 (3.0)	28 (4.2)	47 (4.6)	54 (4.3)
Jersey City Public Schools, NJ	24 (3.8)	26 (4.7)	49 (4.6)	51 (3.5)
Miami-Dade County PS, FL	23 (4.1)	27 (3.5)	50 (3.9)	50 (5.0)
Michigan Invitational Group, MI	25 (3.6)	25 (3.6)	51 (4.2)	49 (4.5)
Montgomery County, MD [2]	24 (2.3)	26 (2.2)	48 (2.8)	52 (2.0)
Naperville Sch. Dist. #203, IL	23 (1.9)	27 (2.1)	49 (2.6)	51 (2.7)
Project SMART Consortium, OH	24 (4.5)	26 (4.4)	49 (4.8)	51 (5.0)
Rochester City Sch. Dist., NY	22 (3.9)	29 (3.0)	48 (4.4)	52 (3.7)
SW Math/Sci. Collaborative, PA	22 (3.1)	29 (4.2)	47 (4.3)	54 (4.3)

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

▲ Significantly greater percentage than other gender

Significance tests adjusted for multiple comparisons