

INDICATORS PART II

Academic Performance—Reading

FOURTH-GRADE READING LITERACY

G-8 Countries Included: England, France, Germany, Italy, Russian Federation, Scotland, United States

On average, U.S. fourth-graders scored higher in reading literacy than their peers in Scotland and France, but lower than their peers in Italy and the Russian Federation. Twelve percent of U.S. fourth-graders reached the advanced international benchmark.

The Progress in International Reading Literacy Study (PIRLS) measures the reading comprehension of students in their 4th year of formal schooling. Begun in 2001, PIRLS is implemented every 5 years. Using data from PIRLS 2006, this indicator presents average scores on the combined reading literacy scale and the percentage of fourth-grade students reaching four established international benchmarks (low, intermediate, high, and advanced) in participating G-8 countries.⁷

In 2006, average scores on the combined reading literacy scale ranged from 522 in France to 565 in the Russian Federation (figure 4a). Every participating G-8 country had an average score above the PIRLS scale average of 500.⁸ Fourth-graders in the Russian Federation outperformed their peers in all other participating G-8 countries in terms of average scores, with students in Scotland and France scoring the lowest among the G-8 countries. Fourth-graders in the United States scored 540 on the combined reading literacy scale, higher than their peers in Scotland and France but lower than their peers in Italy and the Russian Federation.

A greater percentage of students reached the advanced benchmark in the Russian Federation than in all other participating G-8

countries except England (England's score was not measurably different from the Russian Federation's), and France had the smallest percentage of students reach the advanced benchmark. The advanced benchmark is the highest of four benchmarks PIRLS uses to describe performance at various skill levels. In the United States, 12 percent of fourth-graders reached the advanced benchmark (figure 4b). The percentage in the United States was higher than in France (5 percent) but lower than in England (15 percent) and the Russian Federation (19 percent).

Sixty-one percent of fourth-graders scored at or above the high benchmark (the second highest PIRLS benchmark) in the Russian Federation. This is higher than in all other participating G-8 countries, where the percentages ranged from 35 percent in France to 52 percent in Germany and Italy. In the United States, 47 percent of fourth-graders scored at or above the high benchmark. The percentage in the United States was higher than in Scotland and France but lower than in Italy and the Russian Federation.

A greater percentage of students scored at or above the intermediate benchmark in the Russian Federation than in all other participating G-8 countries except Italy (Italy's score was not measurably different from the Russian Federation's). In the United States, 82 percent of fourth-graders scored at or above the intermediate benchmark (figure 4b). The percentage in the United States was higher than in Scotland (77 percent) and France (76 percent), but lower than in Germany and Italy (both at 87 percent) and the Russian Federation (90 percent).

Definitions and Methodology

In PIRLS 2006, countries were required to sample students in the grade that corresponded to the end of 4 years of formal schooling, providing that the mean age at the time of testing was at least 9.5 years. As defined by PIRLS, the 1st year of formal schooling begins with the 1st year of primary school (ISCED97 level 1), which should mark the beginning of formal instruction in reading, writing, and mathematics. Note that kindergarten is not counted. For most countries, the target grade was fourth grade, or its national equivalent.

PIRLS scores are reported on a scale from 0 to 1,000 with the scale average fixed at 500 and the standard deviation fixed at 100. Since the PIRLS reading achievement scales were designed to reliably measure student achievement over time, the metric of the scales was established originally with the 2001 assessment. In order to describe concretely the knowledge and skills attained along the performance scale, PIRLS 2006 established four international achievement benchmarks in reading literacy (low, intermediate, high, and advanced). These benchmarks are identical to the cutpoints used for the Trends in International Mathematics and Science Study (TIMSS). Information about the rationale underlying the benchmarks and the procedures used to set the cutpoints is available in Martin, Mullis, and Kennedy (2007). Four points on the

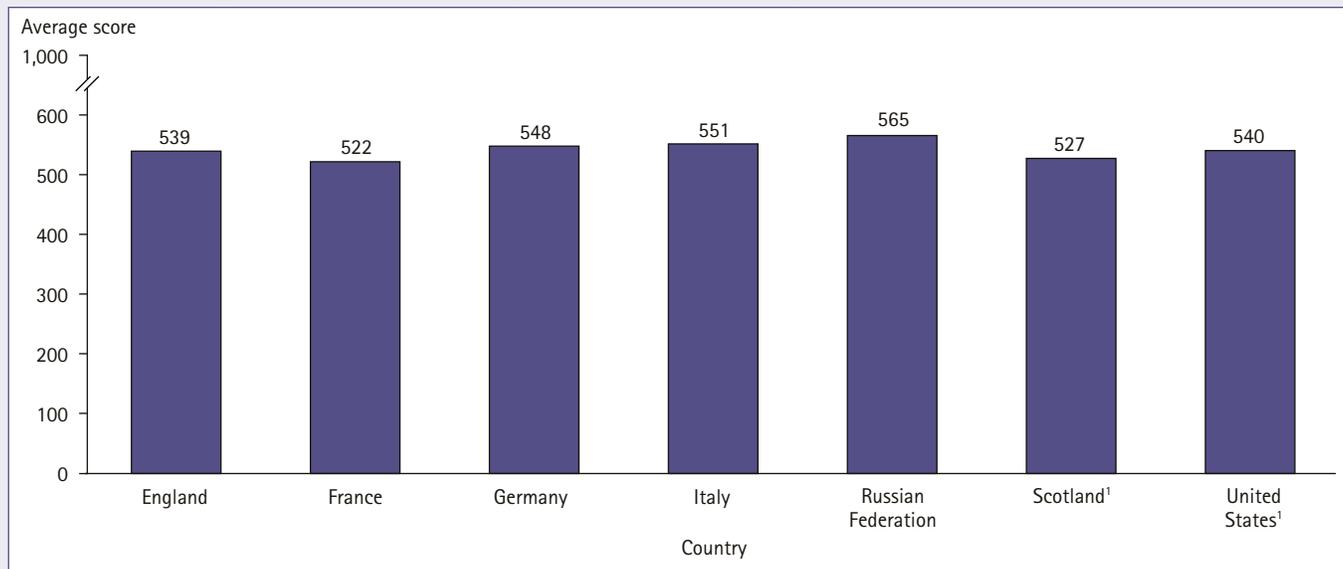
scales were identified for use as international benchmarks: 400 for the low benchmark, 475 for the intermediate benchmark, 550 for the high benchmark, and 625 for the advanced benchmark. These were selected to represent the range of performance shown by students internationally.

Students at the low benchmark display basic reading skills, such as retrieving explicitly stated details from literary and informational texts. Students at the intermediate benchmark demonstrate some reading proficiency. They can identify central events, plot sequences, and relevant story details; make some inferences and connections across parts of the text; and use text organizers (e.g., headings and illustrations) to find information. At the high benchmark, students are competent readers who can recognize some textual features, such as figurative language and abstract messages. They can make inferences on the basis of abstract or embedded information and integrate information to recognize main ideas and provide explanations. Students at the advanced benchmark demonstrate the highest level of reading proficiency. They can interpret figurative language, distinguish and interpret complex information from different parts of text, and integrate ideas across text to provide interpretations about characters' intentions and feelings.

⁷ Canada also participated in PIRLS 2006 but not at the national level. Canada participated as separate provinces (Alberta, British Columbia, Nova Scotia, Ontario, and Quebec). Data corresponding with this indicator for the participating Canadian provinces can be found in appendix tables B1 and B2.

⁸ PIRLS scores are reported on a scale from 0 to 1,000. The metric of the scale was established originally with the 2001 assessment, with the scale average fixed at 500 and the standard deviation fixed at 100.

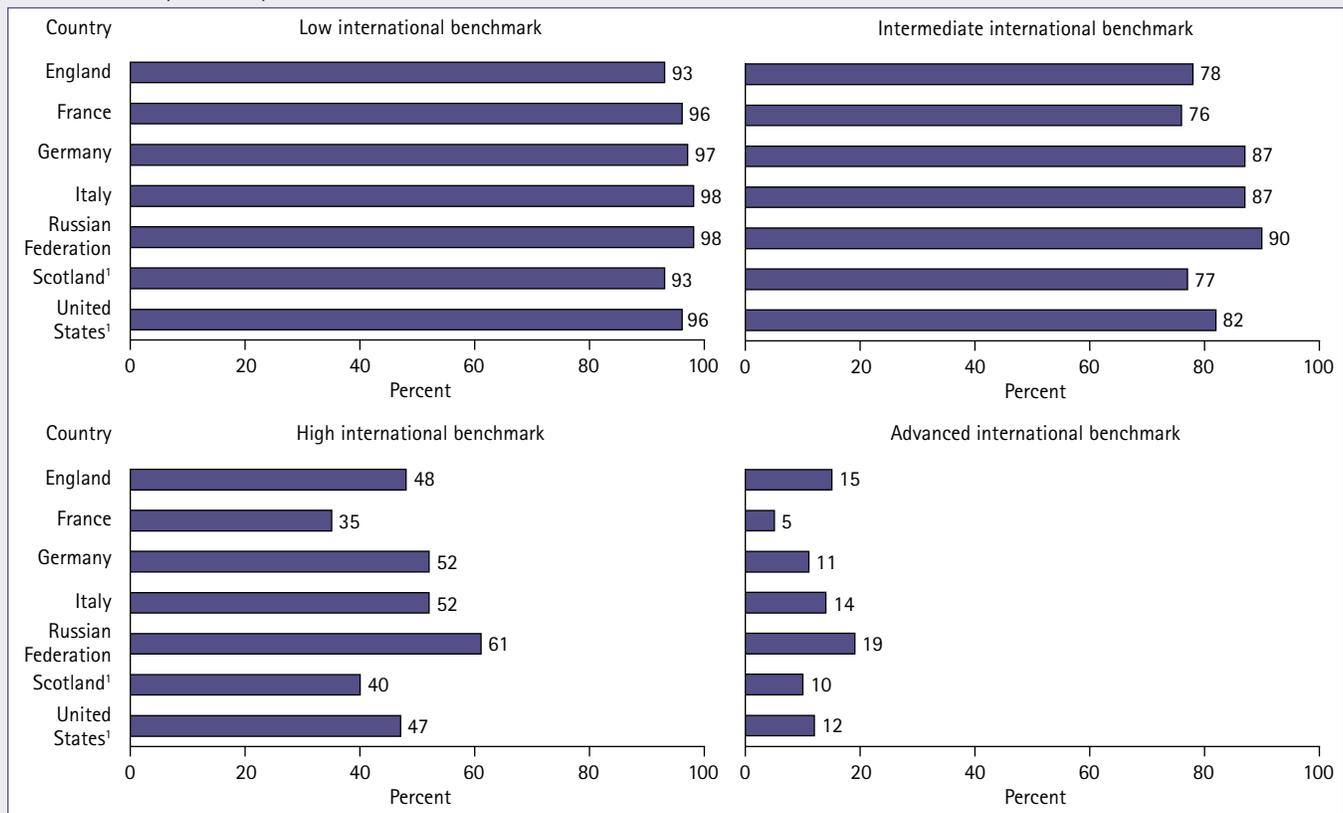
Figure 4a. Average scale scores of fourth-grade students in reading literacy, by country: 2006



¹Met international guidelines for participation rates only after substitute schools were included. That is, to avoid sample size losses resulting from sampled schools not participating, a mechanism was instituted to identify, a priori, substitute schools that have similar characteristics to the sampled schools that they may replace.

SOURCE: Mullis, I.V.S., Martin, M.O., Kennedy, A.M., and Foy, P. (2007). *PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries*, exhibit 1.1. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Figure 4b. Percentage of fourth-grade students reaching PIRLS international benchmarks in reading literacy, by country: 2006



¹Met international guidelines for participation rates only after substitute schools were included. That is, to avoid sample size losses resulting from sampled schools not participating, a mechanism was instituted to identify, a priori, substitute schools that have similar characteristics to the sampled schools that they may replace.

SOURCE: Mullis, I.V.S., Martin, M.O., Kennedy, A.M., and Foy, P. (2007). *PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries*, exhibit 2.1. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

DIFFERENCES IN FOURTH-GRADE READING LITERACY BY SEX

G-8 Countries Included: England, France, Germany, Italy, Russian Federation, Scotland, United States

In 2006, average scores of fourth-grade females in reading literacy were higher than the average scores of fourth-grade males in all participating G-8 countries. In the United States, females outperformed males by 10 points.

The Progress in International Reading Literacy Study (PIRLS) measures the reading comprehension of students in their 4th year of formal schooling. Using data from PIRLS 2006, this indicator compares differences by sex in average scores on the combined reading literacy scale among fourth-graders in participating G-8 countries.⁹

In 2006, average scores of fourth-grade females in reading literacy were higher than the average scores of fourth-grade males in all participating G-8 countries. Among countries with the largest score differences between males and females were Scotland (538 for females vs. 516 for males), with a difference of 22 points; England (549 for females vs. 530 for males), with a difference of 19 points; and the Russian Federation (572 for females vs. 557 for males), with a difference of 15 points (figure 5a and figure 5b). Among countries with the smallest differences related to sex were Germany (551 for females vs. 544 for males) and Italy (555 for females vs. 548 for males), both with a difference of 7 points. In the United States, females outperformed males by 10 points (545 for females vs. 535 for males).

Definitions and Methodology

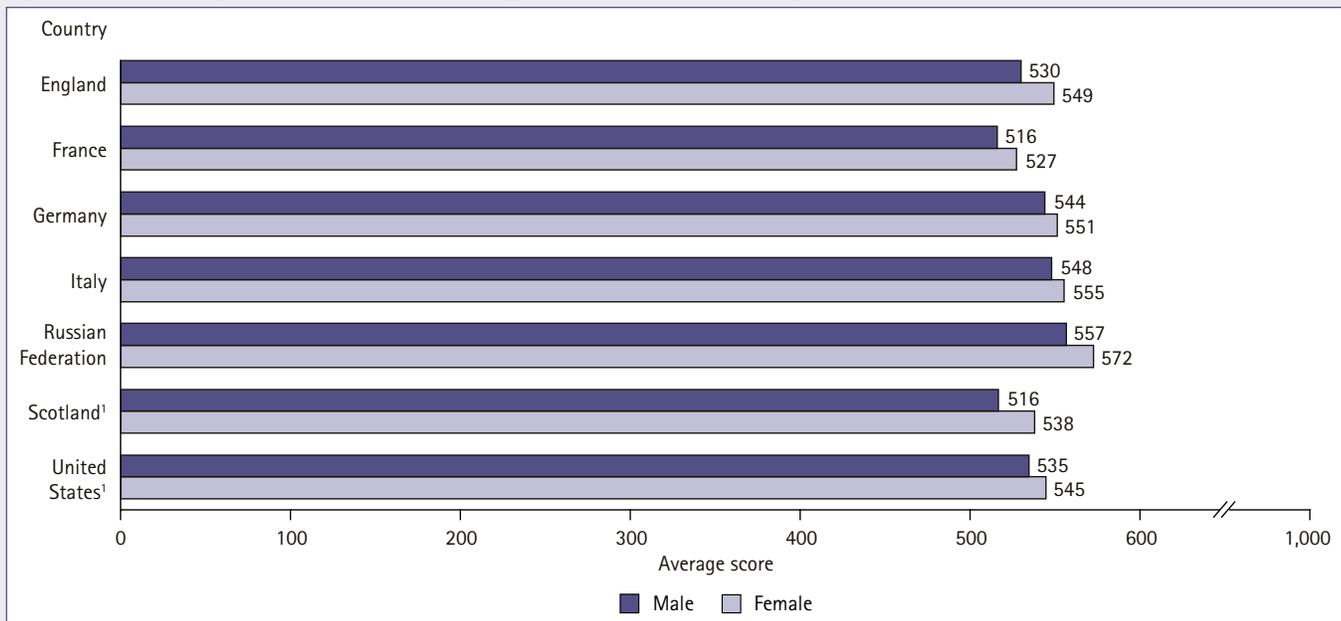
In PIRLS 2006, countries were required to sample students in the grade that corresponded to the end of 4 years of formal schooling, providing that the mean age at the time of testing was at least 9.5 years. As defined by PIRLS, the 1st year of formal schooling begins with the 1st year of primary school (ISCED97 level 1), which should mark the beginning of formal instruction in reading, writing, and mathematics. Note that kindergarten is not counted. For most countries, the target grade was fourth grade, or its national equivalent.

PIRLS scores are reported on a scale from 0 to 1,000 with the scale average fixed at 500 and the standard deviation fixed at 100. Since the PIRLS reading achievement scales were designed to reliably measure student achievement over time, the metric of the scales was established originally with the 2001 assessment.

Male-female score-point differences in reading literacy presented in the text and in figure 5b were computed from unrounded numbers; therefore, they may differ from computations made using the rounded whole numbers that appear in figure 5a.

⁹ Canada also participated in PIRLS 2006 but not at the national level. Canada participated as separate provinces (Alberta, British Columbia, Nova Scotia, Ontario, and Quebec). Data corresponding with this indicator for the participating Canadian provinces can be found in appendix table B3.

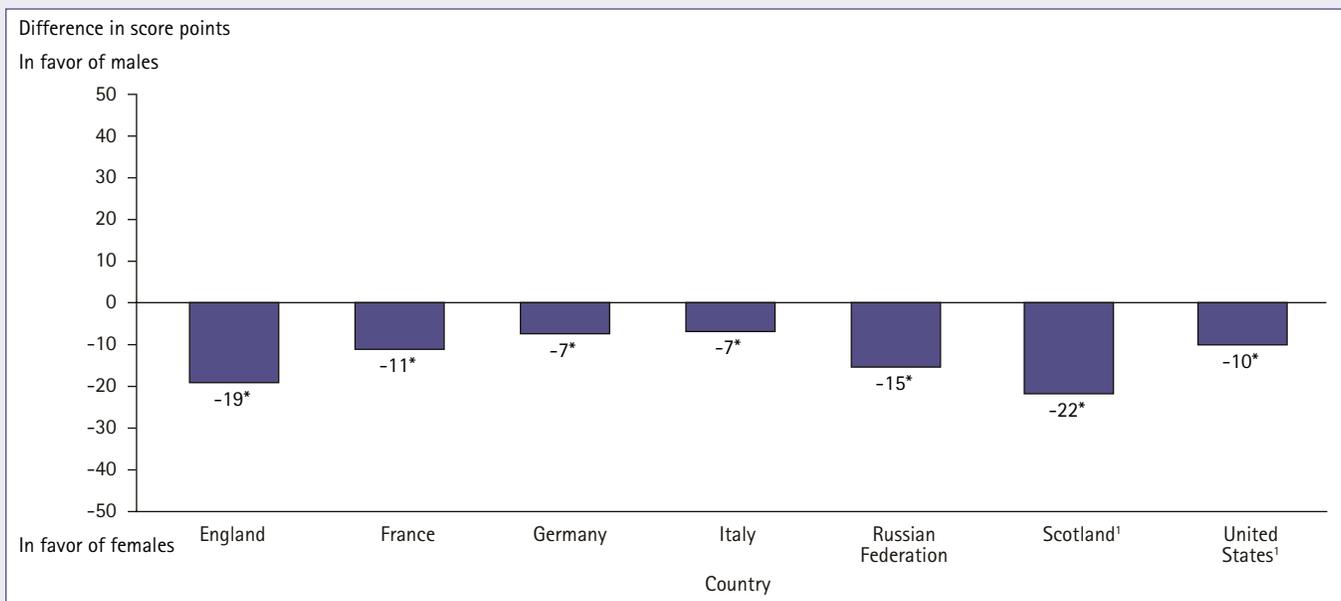
Figure 5a. Average scale scores of fourth-grade students in reading literacy, by sex and country: 2006



¹ Met international guidelines for participation rates only after substitute schools were included. That is, to avoid sample size losses resulting from sampled schools not participating, a mechanism was instituted to identify, a priori, substitute schools that have similar characteristics to the sampled schools that they may replace.

SOURCE: Mullis, I.V.S., Martin, M.O., Kennedy, A.M., and Foy, P. (2007). *PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries*, exhibit 1.4. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Figure 5b. Difference in average scale scores between fourth-grade males and females in reading literacy, by country: 2006



* $p < .05$ (difference in score points is statistically significant).

¹ Met international guidelines for participation rates only after substitute schools were included. That is, to avoid sample size losses resulting from sampled schools not participating, a mechanism was instituted to identify, a priori, substitute schools that have similar characteristics to the sampled schools that they may replace.

SOURCE: Mullis, I.V.S., Martin, M.O., Kennedy, A.M., and Foy, P. (2007). *PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries*, exhibit 1.4. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.



INDICATORS PART II

Academic Performance—Mathematics

PERFORMANCE OF FOURTH- AND EIGHTH-GRADERS IN MATHEMATICS

G-8 Countries Included: England, Germany,¹⁰ Italy, Japan, Russian Federation, Scotland, United States

On TIMSS 2007, students in Japan outperformed students in the other participating G-8 countries in mathematics, with higher percentages of Japanese fourth- and eighth-graders reaching each of the four international benchmarks.

The Trends in International Mathematics and Science Study (TIMSS) assessed students in fourth and eighth grade in mathematics and science in 2007. This indicator presents the percentages of fourth- and eighth-graders reaching the four established international benchmarks in mathematics (low, intermediate, high, and advanced).

On the TIMSS 2007 fourth-grade mathematics assessment, students in Japan outperformed students in the other participating G-8 countries, with higher percentages of Japanese fourth-graders reaching each of the four international benchmarks. The highest international benchmark, advanced, was reached by 23 percent of Japan's fourth-graders in mathematics, compared with percentages ranging from 4 percent in Scotland to 16 percent in the Russian Federation and England (figure 6). In the United States, 10 percent of fourth-graders reached the advanced benchmark. The percentage in the United States was higher than in Germany, Italy, and Scotland, but lower than in the Russian Federation, England, and Japan.

In Japan, 61 percent of fourth-graders reached the high benchmark in mathematics; the percentages in the other G-8 countries ranged from 25 percent in Scotland to 48 percent in England and the Russian Federation. In the United States, 40 percent of fourth-graders reached the high benchmark, a larger percentage than in Italy and Scotland but smaller than in England, the Russian Federation, and Japan.

As at fourth grade, eighth-graders in Japan outperformed their peers in the other participating G-8 countries, with higher percentages of Japanese eighth-graders reaching each of the four benchmarks. The advanced benchmark was reached by 26 percent of Japan's eighth-graders in mathematics, compared with percentages ranging from 3 percent in Italy to 8 percent in the Russian Federation and England. In the United States, 6 percent of eighth-graders reached the advanced benchmark. The percentage in the United States was higher than in Scotland and Italy, but lower than in the Russian Federation and Japan.

In Japan, 61 percent of eighth-graders reached the high benchmark in mathematics; the percentages in the other G-8 countries ranged from 17 percent in Italy to 35 percent in England. In the United States, 31 percent of eighth-graders reached the high benchmark, a larger percentage than in Scotland and Italy.

Definitions and Methodology

In TIMSS 2007 at fourth grade, countries were required to sample students in the grade that corresponded to the end of 4 years of formal schooling (the end of primary school), providing that the mean age at the time of testing was at least 9.5 years. At eighth grade, countries were required to sample students in the grade that corresponded to the end of 8 years of formal schooling (the end of lower secondary education), providing that the mean age at the time of testing was at least 13.5 years.

TIMSS scores are reported on a scale from 0 to 1,000 with the scale average fixed at 500 and the standard deviation fixed at 100. Since the TIMSS mathematics achievement scales were designed to reliably measure student achievement over time, the metric of the scales was established originally with the 1995 assessment. In order to describe concretely the knowledge and skills attained along the performance scales, TIMSS established four international achievement benchmarks in mathematics and science (low, intermediate, high, and advanced). Four points on the scales were identified for use as international benchmarks: 400 for the low benchmark, 475 for the intermediate benchmark, 550 for the high benchmark, and 625 for the advanced benchmark. These were selected to represent the range of performance shown by students internationally.

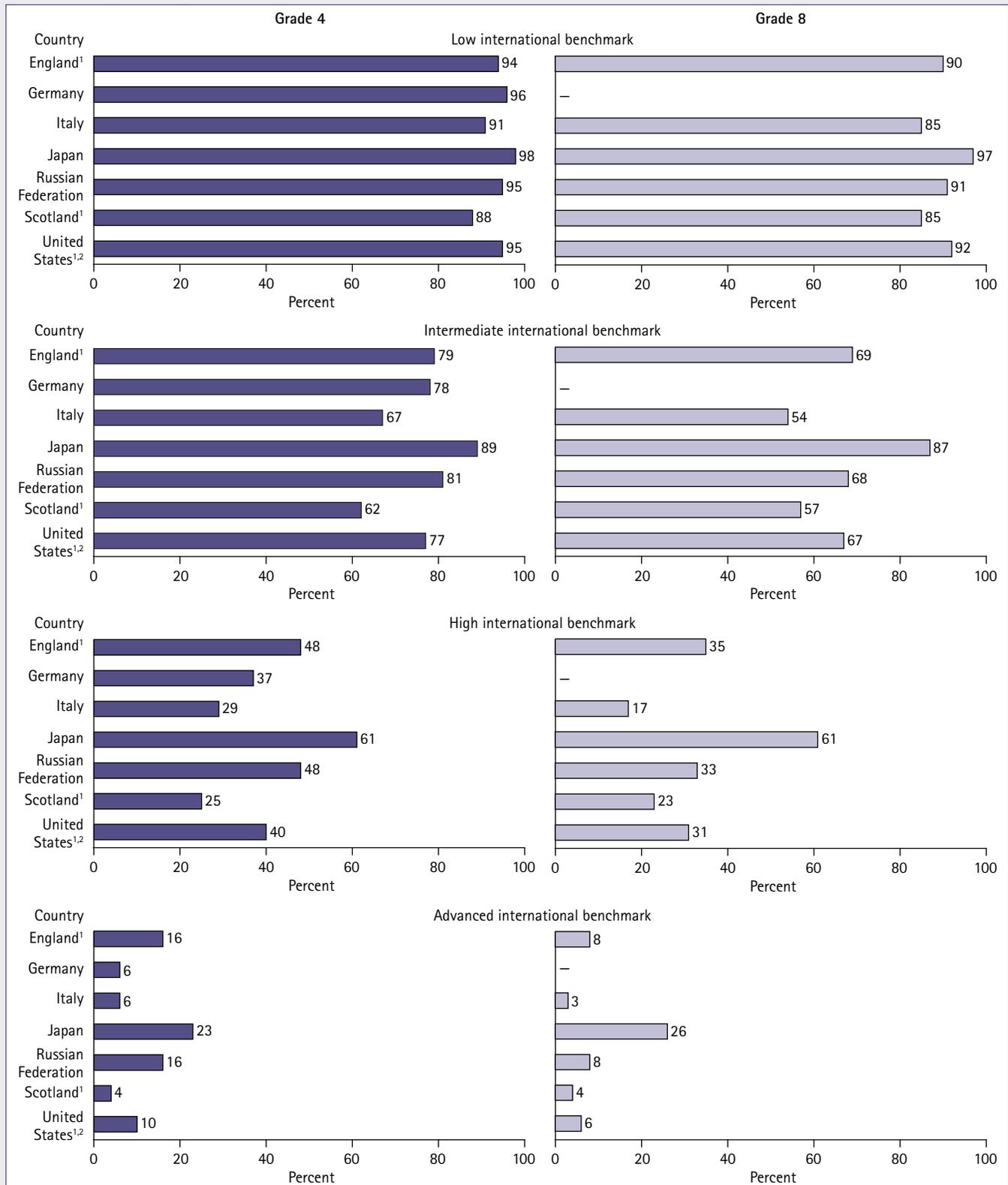
At the fourth-grade level in mathematics, students at the low benchmark have some basic mathematical knowledge, such as an understanding of whole numbers and the properties of basic geometric shapes. At the intermediate benchmark, students can apply basic mathematical knowledge in straightforward situations,

such as performing operations with 3- and 4-digit numbers and decimals and extending simple patterns. At the high benchmark, students can apply their knowledge and understanding to solve multistep word problems involving addition, multiplication, and division and problems requiring the use of data in tables and graphs. Students at the advanced benchmark can apply their understanding and knowledge in a wide variety of relatively complex situations to solve problems involving fractions, decimals, proportions, area, and rotation.

At the eighth-grade level in mathematics, students at the low benchmark have some basic mathematical knowledge, such as the ability to do basic computations with whole numbers and reading information from a line on a graph. At the intermediate benchmark, students can apply basic mathematical knowledge in straightforward situations. For example, they can understand simple algebraic relationships, interpret graphs and tables, and perform basic computations to solve one-step word problems involving whole numbers and decimals. Students at the high benchmark can apply their understanding and knowledge in a wide variety of relatively complex situations to solve problems involving fractions, decimals, negative integers, proportions, area, volume, and probability. Students at the advanced benchmark can organize information, make generalizations, solve nonroutine problems, and draw and justify conclusions from data. For example, they can compute percent change, solve simultaneous linear equations, and model simple situations algebraically.

¹⁰ Data for Germany are only available at the fourth grade, as Germany did not participate in TIMSS 2007 at the eighth grade.

Figure 6. Percentage of fourth- and eighth-grade students reaching TIMSS international benchmarks in mathematics, by country: 2007



- Not available. Data for Germany are only available at the fourth grade, as Germany did not participate in TIMSS 2007 at the eighth grade.
¹Met international guidelines for participation rates only after substitute schools were included. That is, to avoid sample size losses resulting from sampled schools not participating, a mechanism was instituted to identify, a priori, substitute schools that have similar characteristics to the sampled schools that they may replace. For England, this applies to eighth-grade only.
²National Defined Population covers 90 percent to 95 percent of National Target Population.
 SOURCE: Mullis, I.V.S., Martin, M.O., and Foy, P. (2008). *TIMSS 2007 International Mathematics Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*, exhibit 2.2. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

DIFFERENCES IN FOURTH- AND EIGHTH-GRADE MATHEMATICS ACHIEVEMENT BY SEX

G-8 Countries Included: England, Germany,¹¹ Italy, Japan, Russian Federation, Scotland, United States

In 2007, fourth-grade males in Italy, Germany, Scotland, and the United States scored higher, on average, than fourth-grade females in mathematics; however, no measurable differences related to sex were detected among eighth-graders.

This indicator addresses differences by sex in mathematics achievement among fourth- and eighth-grade students in the G-8 countries that participated in the Trends in International Mathematics and Science Study (TIMSS) in 2007.

On the TIMSS 2007 fourth-grade mathematics assessment, males in Italy, Germany, Scotland, and the United States outperformed females. In Italy, the difference in performance was 15 points,

with males scoring an average of 514 compared with 499 among females (figures 7a and 7b). In Germany, the difference by sex was 12 points (531 for males vs. 519 for females); in Scotland, the difference by sex was 9 points (499 for males vs. 490 for females); and in the United States, the difference by sex was 6 points (532 for males vs. 526 for females). In the Russian Federation, however, females outperformed males by 7 points (540 for males vs. 548 for females). In England and Japan, no measurable differences were detected between the average scale scores of fourth-grade males and females.

On the TIMSS 2007 eighth-grade mathematics assessment, no measurable differences related to sex were detected for any of the participating G-8 countries.

Definitions and Methodology

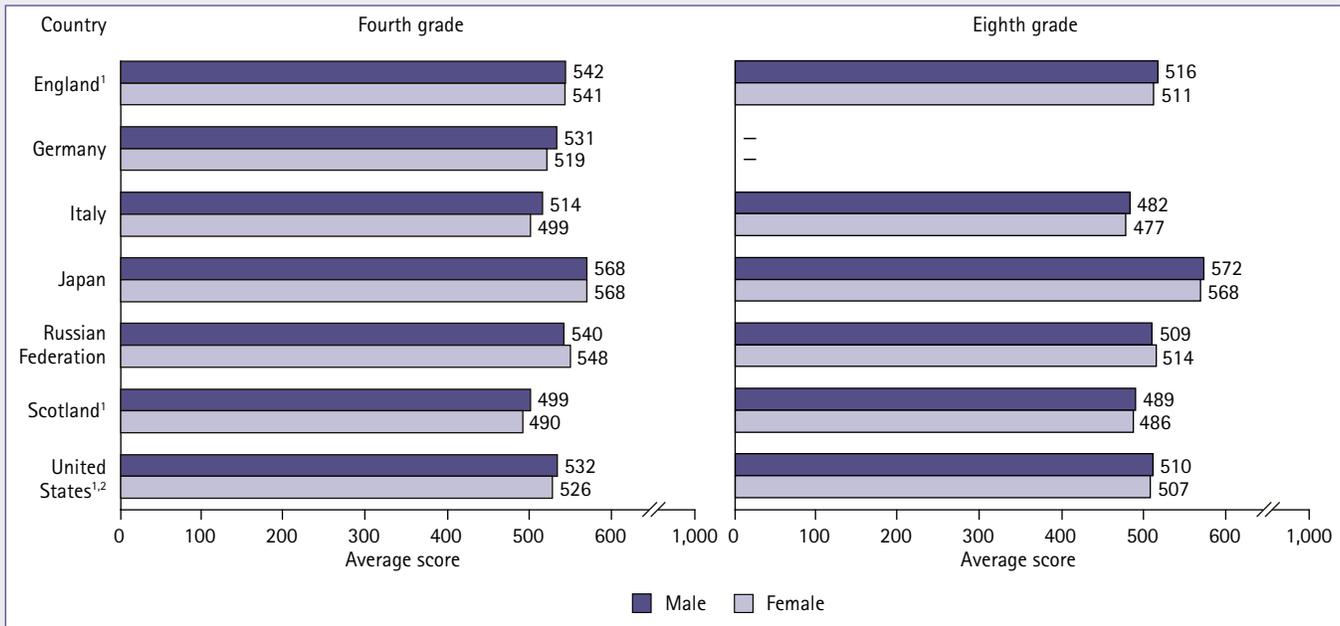
In TIMSS 2007 at fourth grade, countries were required to sample students in the grade that corresponded to the end of 4 years of formal schooling (the end of primary school), providing that the mean age at the time of testing was at least 9.5 years. At eighth grade, countries were required to sample students in the grade that corresponded to the end of 8 years of formal schooling (the end of lower secondary education), providing that the mean age at the time of testing was at least 13.5 years.

TIMSS scores are reported on a scale from 0 to 1,000 with the scale average fixed at 500 and the standard deviation fixed at 100. Since the TIMSS mathematics achievement scales were designed to reliably measure student achievement over time, the metric of the scales was established originally with the 1995 assessment.

Male-female score-point differences in mathematics achievement presented in the text and in figure 7b were computed from unrounded numbers; therefore, they may differ from computations made using the rounded whole numbers that appear in figure 7a.

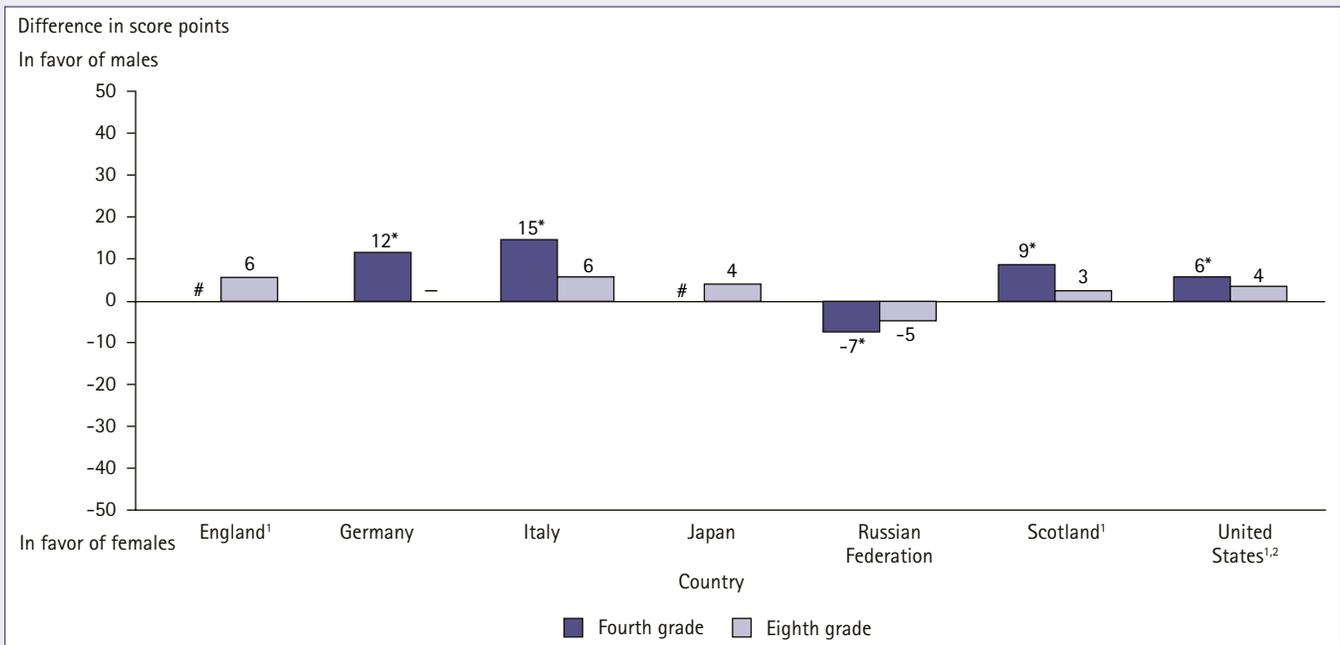
¹¹ Data for Germany are only available at the fourth grade, as Germany did not participate in TIMSS 2007 at the eighth grade.

Figure 7a. Average scale scores of fourth- and eighth-grade students in mathematics literacy, by sex and country: 2007



– Not available. Data for Germany are only available at the fourth grade, as Germany did not participate in TIMSS 2007 at the eighth grade.
¹Met international guidelines for participation rates only after substitute schools were included. That is, to avoid sample size losses resulting from sampled schools not participating, a mechanism was instituted to identify, a priori, substitute schools that have similar characteristics to the sampled schools that they may replace. For England, this applies to eighth grade only.
²National Defined Population covers 90 percent to 95 percent of National Target Population.
 SOURCE: Mullis, I.V.S., Martin, M.O., and Foy, P. (2008). *TIMSS 2007 International Mathematics Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*, exhibit 1.5. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Figure 7b. Difference in average scale scores between fourth- and eighth-grade males and females in mathematics, by country: 2007



– Not available. Data for Germany are only available at the fourth grade, as Germany did not participate in TIMSS 2007 at the eighth grade.
 # Rounds to zero.
 * $p < .05$ (difference in score points is statistically significant).
¹Met international guidelines for participation rates only after substitute schools were included. That is, to avoid sample size losses resulting from sampled schools not participating, a mechanism was instituted to identify, a priori, substitute schools that have similar characteristics to the sampled schools that they may replace. For England, this applies to eighth grade only.
²National Defined Population covers 90 percent to 95 percent of National Target Population.
 NOTE: Differences shown are computed by subtracting the average unrounded score for females from the average unrounded score for males. Thus, positive values indicate higher average scores for males.
 SOURCE: Mullis, I.V.S., Martin, M.O., and Foy, P. (2008). *TIMSS 2007 International Mathematics Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*, exhibit 1.5. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.