NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education
NCES 2006-073

# U.S. Student and Adult Performance on International Assessments of Educational Achievement

Findings from
The Condition of Education 2006

This page intentionally left blank.

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education
NCES 2006-073

# U.S. Student and Adult Performance on International Assessments of Educational Achievement

## Findings from The Condition of Education 2006

**June 2006**

Mariann Lemke
Patrick Gonzales
**National Center for Education Statistics**

**U.S. Department of Education**
Margaret Spellings
*Secretary*

**Institute of Education Sciences**
Grover J. Whitehurst
*Director*

**National Center for Education Statistics**
Mark Schneider
*Commissioner*

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in other countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other U.S. Department of Education, National Center for Education Statistics product or report, we would like to hear from you. Please direct your comments to

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street, NW
Washington, DC 20006–5651

June 2006

The NCES World Wide Web Home Page is http://nces.ed.gov.
The NCES World Wide Web Electronic Catalog is http://nces.ed.gov/pubsearch.

**Suggested Citation**

Lemke, M., and Gonzales, P. (2006). *U.S. Student and Adult Performance on International Assessments of Educational Achievement: Findings from The Condition of Education 2006* (NCES 2006-073). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

For ordering information on this report, write to

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

call toll free 1–877–4ED–PUBS; or order online at http://www.edpubs.org.

**Content Contact**
Michael Planty
(202) 502–7312
michael.planty@ed.gov

## Preface

*The Condition of Education* summarizes important developments and trends in education using the latest available data. The report, which is required by law, is an indicator report intended for a general audience of readers who are interested in education. The indicators represent a consensus of professional judgment on the most significant national measures of the condition and progress of education for which accurate data are available. The 2006 print edition includes 50 indicators in five main areas: (1) enrollment trends and student characteristics at all levels of the education system from early childhood education to graduate and first-professional programs; (2) student achievement and the longer term, enduring effects of education; (3) student effort and rates of progress through the educational system among different population groups; (4) the contexts of elementary and secondary education in terms of courses taken, teacher characteristics, and other factors; and (5) the contexts of postsecondary education.

The 2006 edition also includes a special analysis that presents key findings of recent international assessments that examine the performance of U.S. students in reading, mathematics, and science and the literacy of adults compared with the performance of their peers in other countries. To make the special analysis available to audiences interested in U.S. performance on international assessments of educational achievement, the special analysis is reprinted here as a separate volume. Technical notes about the data sources, methodology, and standard errors are included at the end of this booklet.

Special analyses included in the 2000–06 editions of *The Condition of Education* are available both as booklets and in the full print volumes. They are also available on the NCES *Condition of Education* website (http://nces.ed.gov/programs/coe).

This page intentionally left blank.

# U.S. Student and Adult Performance on International Assessments of Educational Achievement

## Introduction

As part of its congressional mandate, the National Center for Education Statistics (NCES) is required to report on the state of education in the United States and other countries (Education Sciences Reform Act of 2002). To carry out this mission, NCES engages in a number of activities designed to gather information and produce indicators on how the performance of U.S. students, teachers, and schools compares with that of their counterparts in other countries. NCES and other offices within the U.S. Department of Education work with foreign ministries of education and international organizations, such as the Organization for Economic Cooperation and Development (OECD), the International Association for the Evaluation of Educational Achievement (IEA), and the United Nations Educational, Scientific and Cultural Organization (UNESCO) to plan, develop, and implement reliable and meaningful measures across countries.

The United States participates in several international assessments designed to provide comparable information about achievement in various subject areas. These assessments offer an opportunity to compare the performance of U.S. students and adults with that of their peers in other countries. They also provide an opportunity to observe characteristics associated with high and low achievement across countries and to posit questions about policies and practices that could be applied in U.S. schools to improve student learning.

The United States has participated in developing and conducting cross-national assessments since the 1960s. Since the first comparative assessments were given, the number and scope of international assessments have grown. The implementation of technical standards and increased monitoring, along with the expertise that the international community has contributed to assessment design, has improved the quality of data over time. For complete details on the methods instituted to ensure

data quality and comparability, see Adams (2005); Martin, Mullis, and Chrostowski (2004); Martin, Mullis, and Kennedy (2003); and Statistics Canada (2005).

Currently, the United States participates in four international assessments: the Progress in International Reading Literacy Study (PIRLS), which assesses reading performance in grade 4; the Program for International Student Assessment (PISA), which assesses the reading, mathematics, and science literacy of 15-year-olds;[1] the Trends in International Mathematics and Science Study (TIMSS), which assesses mathematics and science performance in grades 4 and 8; and the Adult Literacy and Lifeskills Survey (ALL), which assesses the adult literacy and numeracy skills of 16- to 65-year-olds (table 1). Each international assessment measures one or more dimensions of the performance or ability of U.S. students or adults. Combined with data from national assessments,[2] these international assessment data provide educators and policymakers with a more complete picture of educational achievement in the United States.

This special analysis will present major findings from each of these assessments. The purpose of this special analysis is three-fold: (1) to discuss the similarities and differences in the countries participating in the assessments; (2) to report the most recent findings of these assessments; and (3) to compare the overall performance of students and adults in the United States with their peers in other countries.

## Which Countries Participate?

Countries around the world are invited to participate in each assessment by the sponsoring international organization. Because they volunteer to participate, the number and range of countries (e.g., developed vs. developing) vary from assessment to assessment. Though TIMSS, PIRLS, and PISA include developed and developing countries, a larger proportion of developing countries have participated in TIMSS and PIRLS than in PISA and ALL (table 1). PISA is primarily administered in the member countries of the OECD—an intergovernmental organization of 30 industrialized countries seeking to promote trade and economic growth. ALL was conducted only among 6 countries in 2003, but additional countries collected data in 2005, and more countries plan to participate in future years.

Differences in the combinations of countries that participate in the assessments can affect how various measures, such as the international average, are calculated and interpreted. For example, because national average scores in developing countries tend to be lower than those in developed countries, the international averages can vary from administration to administration, depending on which countries participate. In TIMSS and PIRLS, the international averages are calculated using results from both developing and developed countries while in PISA, they are calculated using results only from the OECD-member countries.

## Table 1. Recent international assessments

| Study | Age/grade assessed | Subjects assessed | Year administered | Number of participating countries[1] | Average GDP per capita of participating countries (in U.S. dollars using PPP)[2] | Average HDI of participating countries[3] |
|---|---|---|---|---|---|---|
| Progress in International Reading Literacy Study (PIRLS) | 4th grade | Reading | 2001<br>2006 | 35 | $13,229 | 0.865 |
| Trends in International Mathematics and Science Study (TIMSS) | 4th grade<br>8th grade[4] | Mathematics<br>Science | 1995<br>1999<br>2003<br>2007 | 25 at grade 4;<br>45 at grade 8 | $15,911 (grade 4);<br>$10,808 (grade 8) | 0.863 (grade 4);<br>0.820 (grade 8) |
| Program for International Student Assessment (PISA) | 15-year-olds | Reading literacy<br>Mathematics literacy<br>Science literacy | 2000<br>2003<br>2006 | 39 | $26,172 | 0.917 |
| Adult Literacy and Lifeskills Survey (ALL) | 16- to 65-year-olds | Literacy<br>Numeracy | 2003 | 6 | $33,598 | 0.947 |

[1] Number of participating countries based on the most recently completed year of the assessment.

[2] Average gross domestic product (GDP) per capita is based on the averages of the participating countries in 2003 that completed all necessary steps to appear in the international reports. GDP per capita is taken from the United Nations Development Program (UNDP) *Human Development Report 2005*. Figures are converted using purchasing power parity (PPP) conversion factors that take into account differences in the relative prices of goods and services—particularly non-tradables—and therefore provide a better overall measure of the real value of output produced by an economy compared to other economies. PPP GDP is measured in current international dollars which, in principle, have the same purchasing power as a dollar spent on gross national index in the U.S. economy. Average GDP per capita for PISA includes Organization for Economic Cooperation and Development (OECD)-member nations only. Average GDP per capita for TIMSS, PIRLS, and ALL includes all nations for which data were available. GDP per capita data were unavailable for Bermuda, Chinese Taipei, Lichtenstein, Macao-China, Palestinian National Authority, and Serbia and are thus not included in the averages.

[3] Average Human Development Index (HDI) is based on the HDI of particpating countries in 2003 and includes only those countries that completed all necessary steps to appear in the international reports. The HDI is a composite index that takes into account three dimensions of human development: life expectancy; knowledge; and standard of living. HDI figures are taken from the UNDP *Human Development Report 2005*. HDI scores range from 0 (lowest) to 1 (highest). Average HDI for PISA includes OECD-member nations only. Average HDI for TIMSS, PIRLS, and ALL includes all nations for which data were available. HDI figures were unavailable for Bermuda, Chinese Taipei, Lichtenstein, Macao-China, and Serbia and are thus not included in the averages.

[4] Fourth-graders were only assessed in 1995 and 2003.

SOURCE: International Association for the Evaluation of Educational Achievement (IEA), Progress in International Reading Literacy Study (PIRLS), 2001; Trends in International Mathematics and Science Study (TIMSS), 2003; Statistics Canada and Organization for Economic Cooperation and Development (OECD), Adult Literacy and Lifeskills Survey (ALL), 2003; OECD, Program for International Student Assessment (PISA), 2003; and United Nations Development Program (UNDP), *Human Development Report 2005*, previously unpublished tabulation (October 2005).

## How Comparable Are the Schools and Students That Participate?

One challenge in comparing assessment data from countries around the world is determining the extent that variations in the characteristics of student and adult populations relate to achievement scores. For example, restrictions in attrition rates as students move through the educational system, the economic and social status of students and their families, and parental levels of education may each affect the comparability of findings both within and across assessments. In developing international assessments, the challenge of making student populations comparable is generally dealt with in two ways.

First, countries that participate in international assessments such as TIMSS, PIRLS, ALL, and PISA are required to select national probability samples from *all* students or adults in a particular grade or of a particular age. Exclusions are strictly limited, must be clearly documented, and are reported along with participation rates at each level of sampling. Countries with exclusion rates that are above established levels or with samples that are not representative of the population being assessed run the risk of being eliminated from reports.

Second, in the school-based assessments, the grades or ages selected for assessment are chosen to maximize the likelihood of youth being enrolled in school; for example, PISA samples are drawn from the population of 15-year-old students enrolled in school. In 2003, the most recent year for which data are available, the percentage of the population ages 5–14 enrolled in school was 90 percent or higher in most developed countries, including the United States, and 80 percent or higher in most developing countries that participated in international assessments (OECD 2004a, table C1.2). The percentage of the U.S. population ages 15–19 enrolled in public or private school was 75 percent, which is comparable to or below that of most other industrialized countries. Comparisons of graduation rates from upper secondary school (high school in the United States) paint a similar picture: the U.S. graduation rate (73 percent) is comparable to or below that of most industrialized countries, where 80 percent or more of students finish upper secondary school (OECD 2004a, table A2.1).

Further differences among countries in terms of their student population characteristics, especially those found to be significantly related to achievement, can also be evaluated and explained in comparative analyses. Research has established that students' economic and social characteristics, such as their immigrant status and family income, are associated with academic achievement (Coleman et al. 1966; Entwisle and Alexander 1993; Shavit and Blossfield 1993). Moreover, research has shown that these factors are often interrelated, further complicating the picture (McLanahan and Sandefur 1994; Schmid 2001). For example, minority status, family income, language ability, and family structure are associated with students' achievement in the United States (Coleman et al. 1966; Jencks et al. 1979; McLanahan and Sandefur 1994; Schmid 2001), and such relationships are also found in many other countries (Buchmann 2002). The uneven distribution

of students' economic and social factors across countries, as well as the potential cross-national variation in the relationship between student achievement and these factors, may affect the outcomes of cross-national comparisons.

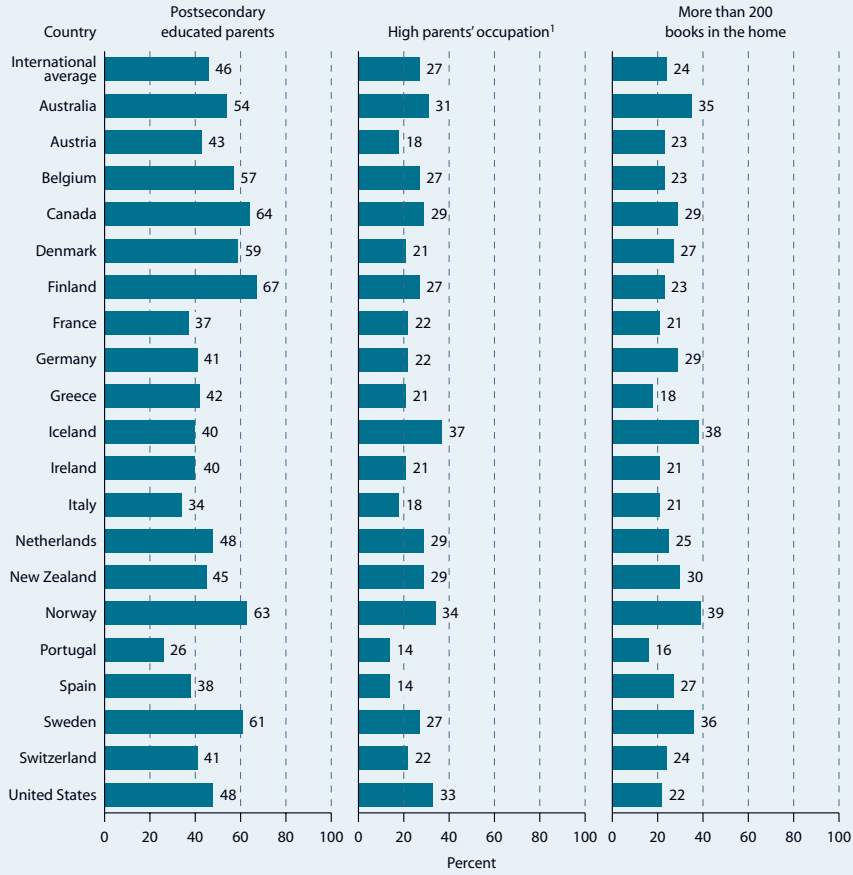
Recent comparisons of PISA 2003 data have explored how variations in student population characteristics across countries may affect the reported outcomes of international studies. For example, it is true that some characteristics of the U.S. student population are different from those of student populations in countries like Japan and Korea, where there are few foreign-born students; however, student populations in other countries are often not measurably different from the U.S. student population in terms of the distribution of salient social and economic factors (figures 1 and 2; Hampden-Thompson and Johnston 2006). For example, 48 percent of 15-year-old students in the United States reported having at least one parent who had a college degree or a postsecondary vocational qualification (figure 1). When the United States was compared to the other 19 countries in this study, 11 countries were found to have a smaller percentage of students with postsecondary-educated parents when compared with the United States. Seven countries had a higher percentage of 15-year-old students who reported that at least one of their parents was educated to the postsecondary level (Australia, Belgium, Canada, Denmark, Finland, Norway, and Sweden). Also, the data show that 9 percent of U.S. 15-year-olds did not speak the language of the test at home (i.e., English; figure 2). Of the 19 other countries, 6 had a greater percentage of 15-year-olds who did not speak the language of the test at home, and 8 countries had a lower percentage.

Cross-national comparisons of student populations and their social and economic contexts show that the United States shares many of the same educational challenges as other countries. For example, while the strength of the association may vary, many studies report a fairly consistent relationship between lower socioeconomic status and lower student achievement (Buchmann 2002). The cross-national comparisons of achievement displayed in the sections that follow have not been adjusted for socioeconomic or other factors.

## How Do U.S. Students and Adults Compare With Their Peers in Other Countries?

Results for U.S. students and adults on international assessments vary by subject, grade or age, and assessment. Although it would be desirable, it is not possible to directly compare the international assessment scores from the various studies because of differences in the countries participating, the purpose of the assessments, the items used, and the target populations. Without making direct comparisons between studies, the following section presents highlights of the key findings of several recent international studies that looked at students' and adults' achievement in reading, mathematics, and science.

## Figure 1. Percentage of 15-year-olds whose parents had a postsecondary education, had high occupational status, and had more than 200 books in the home, by country: 2003

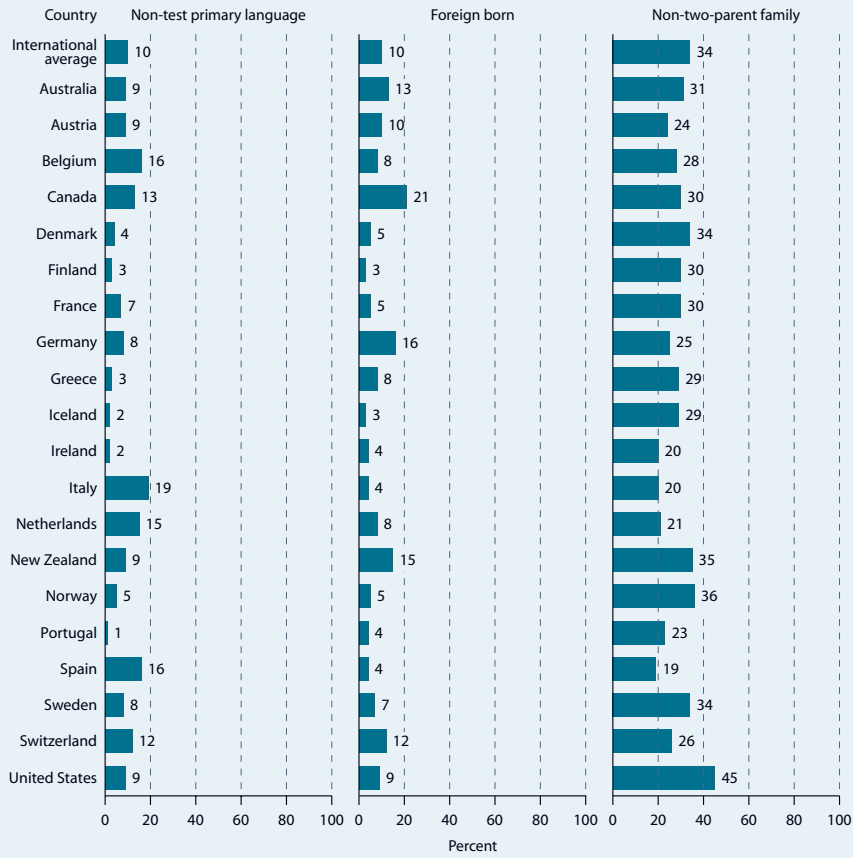



[1] Parents' occupation is reported by the student and coded to the International Standard Classification of Occupations (ISCO-88) and then grouped into major occupational groups. For further information, see Ganzeboom et al. (1992).

NOTE: The international average is the weighted mean of the data values for the 20 countries included in the analysis. Parent education, parent occupational status, and number of books in the home are based on students' reports. If either of a student's parents completed a bachelor's, master's, or postgraduate degree (corresponding to the International Standard Classification of Education (ISCED) levels 5A, 5B, or 6), the student was considered as having postsecondary-educated parents. Parent occupational status is based on either of the student's parents' occupation (whichever is higher), and the variable was transformed into quarters with "high" occupational status representing the upper quarter. The response rate in New Zealand for parent occupational status was below 85 percent.

SOURCE: Hampden-Thompson, G., and Johnston, J.S. (2006). *Variation in the Relationship Between Nonschool Factors and Student Achievement on International Assessments* (NCES 2006-014), table 1. Data from Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2003.

## Figure 2. Percentage of 15-year-olds who spoke a non-test language, were foreign born, and were from non-two-parent families, by country: 2003




NOTE: The international average is the weighted mean of the data values for the 20 countries included in the analysis. Language spoken at home, immigrant status, and family structure are based on students' reports. "Test-language" students reported speaking the language in which the test was administered always or most of the time at home while "non-test-language" students reported using another language always or most of the time at home. Students from a "two-parent family" reported living with both their mother and father. The category "non-two-parent family" encompasses all other responses.

SOURCE: Hampden-Thompson, G., and Johnston, J.S. (2006). *Variation in the Relationship Between Nonschool Factors and Student Achievement on International Assessments* (NCES 2006-014), table 1. Data from Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2003.

## Reading

Three international assessments measure aspects of reading skills. The Progress in International Reading Literacy Study (PIRLS) assesses 4th-grade reading skills; the Program for International Student Assessment (PISA) focuses on the ability of 15-year-olds to apply their reading skills to a wide variety of materials within a real-life context; and the Adult Literacy and Lifeskills Survey (ALL) assesses the literacy skills of adults ages 16–65.

### PIRLS

Administered in 35 countries in 2001, PIRLS defines reading literacy as

> The ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment. (Mullis et al. 2004a, p. 3).

To measure the reading literacy skills and abilities of 4th-graders, PIRLS used a combination of literary texts—passages drawn from children's books—and informational texts—passages providing information on people, places, and things. Students were asked to demonstrate skills and abilities such as retrieving specific information, making inferences, interpreting and integrating ideas and information, and examining and evaluating content and language.

- *U.S. 4th-graders had higher average reading literacy scores than the international average and higher scores than students in 23 of the 34 other participating countries in 2001.*

The results from PIRLS indicate that U.S. 4th-graders performed as well as or better than most of their international peers in the other 34 participating countries (table 2). Specifically, U.S. 4th-graders performed above the international average, and, on average, they outperformed students in two-thirds of the other participating countries. The performance of students in about one-quarter of the participating countries was not measurably different from that of U.S. students. Students in three countries (Sweden, the Netherlands, and England) outperformed U.S. students, on average. The average score of U.S. 4th-graders was not measurably different from the average student scores in other industrialized countries such as Canada (Ontario and Quebec), Italy, and Germany. U.S. 4th-graders outscored their peers in some industrialized countries, such as New Zealand, Scotland, France, and Norway, as well as in a number of developing countries.

In addition to overall reading scores, PIRLS provides subscale scores for specific reading skills: reading for literary experience and reading to acquire and use information. On average, U.S. 4th-graders performed as well as or better than their peers

**Table 2. Average PIRLS reading literacy scores of 4th-graders, by country: 2001**

| Country | Average score |
|---|---|
| International average | 500 |
| Sweden | 561 |
| Netherlands[1] | 554 |
| England[1, 2] | 553 |
| Bulgaria | 550 |
| Latvia | 545 |
| Canada (O,Q)[3, 4] | 544 |
| Lithuania[3] | 543 |
| Hungary | 543 |
| **United States[1]** | **542** |
| Italy | 541 |
| Germany | 539 |
| Czech Republic | 537 |
| New Zealand | 529 |
| Scotland[1] | 528 |
| Singapore | 528 |
| Russian Federation[2] | 528 |
| Hong Kong SAR[5] | 528 |
| France | 525 |
| Greece[2] | 524 |
| Slovak Republic | 518 |
| Iceland | 512 |
| Romania | 512 |
| Israel[2] | 509 |
| Slovenia | 502 |
| Norway | 499 |
| Cyprus | 494 |
| Moldova | 492 |
| Turkey | 449 |
| Macedonia | 442 |
| Colombia | 422 |
| Argentina | 420 |
| Iran | 414 |
| Kuwait | 396 |
| Morocco[6] | 350 |
| Belize | 327 |

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
■ Average is lower than the U.S. average.

[1] Met international guidelines for sample participation rates only after replacement schools were included.
[2] National defined population covers less than 95 percent of national desired population.
[3] National desired population does not cover all of international desired population.
[4] Canada is represented by the provinces of Ontario and Quebec (O, Q) only.
[5] Hong Kong SAR is a Special Administrative Region (SAR) of the People's Republic of China.
[6] Nearly satisfied guidelines for sample participation rates after replacement schools were included.
NOTE: Participants were scored on a 1,000-point scale. The international standard deviation is 100 points. The test for significance between the U.S. average and the international average was adjusted to account for the contribution of the U.S. average to the international average.
SOURCE: Ogle, L.T., Sen, A., Pahlke, E., Jocelyn, L., Kastberg, D., Roey, S., and Williams, T. (2003). *International Comparisons in Fourth-Grade Reading Literacy: Findings from the Progress in International Reading Literacy Study (PIRLS) of 2001* (NCES 2003-073), figure 3. Data from International Association for the Evaluation of Educational Achievement, Progress in International Reading Literacy Study (PIRLS), 2001.

in most countries in both reading subscales (Ogle et al. 2003). Students in only one country, Sweden, outperformed U.S. students in reading for literacy experience; students in five countries (Sweden, the Netherlands, Bulgaria, Latvia, and England) outperformed U.S. students in reading to acquire and use information.

As with all international assessments in which the United States participates, PIRLS data can be analyzed to provide information on the achievement of student

subpopulations. For example, 19 percent of U.S. students performed among the top 10 percent of all 4th-graders across the 35 countries that participated in PIRLS in 2001, a percentage exceeded only in England (Ogle et al. 2003). Among U.S. 4th-graders, a larger percentage of White students performed in the top 10 percent of all students than their Black or Hispanic peers. In all 35 countries, including the United States, girls outperformed boys in reading. Girls in Sweden, England, the Netherlands, and Bulgaria outperformed U.S. girls in reading, on average, while boys in the Netherlands and Sweden outperformed U.S. boys.

PIRLS will be repeated in 2006, providing more information about the progress of U.S. students in reading relative to other countries. Results of the PIRLS 2001 assessment can be found in Ogle et al. (2003; available at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2003073) and Mullis et al. (2003; available at http://isc.bc.edu/pirls2001i/PIRLS2001_Pubs_IR.html).

### PISA

PISA measured the reading literacy of 15-year-olds in 2000. In this study, reading literacy was defined as "understanding, using, and reflecting on written texts in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society" (OECD 1999, p. 20). PISA measured the extent to which students could apply different reading processes (retrieving information, interpreting text, and reflecting on text) to a range of reading materials they were likely to encounter as young adults, such as government forms, newspaper articles, manuals, books, and magazines.

- *U.S. 15-year-olds scored at the OECD average in reading literacy in 2000.*

PISA 2000 results showed that U.S. 15-year-olds performed as well as or better than most of their peers in the 30 other participating countries (table 3). On average, students in Finland, Canada, and New Zealand outperformed U.S. students, but the U.S. average scores were not significantly different from those in most other industrialized countries as well as the OECD average.[3] PISA also provided subscale scores based on processes used when reading a text: retrieving information from text; interpreting texts; and reflecting on texts to relate to other experiences, knowledge, or ideas. U.S. 15-year-olds scored at the OECD average on all three reading processes measured. However, students in five countries outperformed U.S. students on a measure of retrieving information, and students in four countries outperformed U.S. students on a measure of reflecting on texts. On a measure of interpreting texts, students in two countries—Finland and Canada—outperformed U.S. 15-year-olds (Lemke et al. 2001).

Thirteen percent of U.S. students performed among the top 10 percent of all 15-year-olds in the OECD-member countries that participated in PISA 2000 (Lemke

**Table 3. Average PISA reading literacy scores of 15-year-olds, by country: 2000**

| Country | Average score |
|---|---|
| OECD average | 500 |
| **OECD countries** | |
| Finland | 546 |
| Canada | 534 |
| New Zealand | 529 |
| Australia | 528 |
| Ireland | 527 |
| Korea, Republic of | 525 |
| United Kingdom | 523 |
| Japan | 522 |
| Sweden | 516 |
| Austria | 507 |
| Belgium | 507 |
| Iceland | 507 |
| Norway | 505 |
| France | 505 |
| **United States** | **504** |
| Denmark | 497 |
| Switzerland | 494 |
| Spain | 493 |
| Czech Republic | 492 |
| Italy | 487 |
| Germany | 484 |
| Hungary | 480 |
| Poland | 479 |
| Greece | 474 |
| Portugal | 470 |
| Luxembourg | 441 |
| Mexico | 422 |
| **Non-OECD countries** | |
| Liechtenstein | 483 |
| Russian Federation | 462 |
| Latvia | 458 |
| Brazil | 396 |

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
■ Average is lower than the U.S. average.

NOTE: The test for significance between the United States and the Organization for Economic Cooperation and Development (OECD) average was adjusted to account for the contribution of the U.S. average to the OECD average. Because PISA is principally an OECD study, the results for non-OECD countries are displayed separately from those for the OECD countries and are not included in the OECD average. Due to low response rates, data for the Netherlands are not included. Participants were scored on a 1,000-point scale. The international standard deviation is 100 points.

SOURCE: Lemke, M., Calsyn, C., Lippman, L., Jocelyn, L., Kastberg, D., Liu, Y.Y., Roey, S., Williams, T., Kruger, T., and Bairu, G. (2001). *Outcomes of Learning: Results From the 2000 Program for International Student Assessment of 15-Year-Olds in Reading, Mathematics, and Science Literacy* (NCES 2002-115), figure 3. Data from Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2000.

et al. 2001), and about one-third of U.S. students were found to read at the two highest levels of performance. Similar to the results in the PIRLS 2001 study, girls outperformed boys in reading literacy in the United States and all other participating PISA countries (Lemke et al. 2001). More information on the performance of other student population groups can be found in Lemke et al. (2001; available at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002115) and OECD (2001; available at https://www.pisa.oecd.org/dataoecd/44/53/33691596.pdf).

## ALL

In 2003, the United States participated in ALL along with five other countries. The study assessed the literacy and numeracy skills of adults ages 16–65 through a written test administered in respondents' homes. In this study, literacy was

defined as the knowledge and skills needed by adults, in life and at work, to use information from various texts (e.g., news stories, editorials, manuals, brochures) in various formats (e.g., texts, maps, tables, charts, forms, time tables) (Statistics Canada and OECD 2005). The ALL test questions were developed to assess the respondent's ability to retrieve, compare, integrate, and synthesize information from texts and to make inferences, among other skills.

### ■ *U.S. adults had lower literacy skills, on average, than adults in Norway, Bermuda, Switzerland, and Canada in 2003 and had higher literacy skills than adults in Italy.*

Results from ALL showed that U.S. adults outperformed adults in Italy in 2003, but were outperformed by adults in Norway, Bermuda, Switzerland, and Canada (table 4). Adults in Bermuda, Norway, and Canada had higher literacy scores than U.S. adults at both the high and low ends of the score distribution (Lemke et al. 2005). The highest performers (the top 10 percent of adults) had literacy scores of 353 or higher in Bermuda, 348 or higher in Norway, and 344 or higher in Canada, compared with 333 or higher in the United States. The lowest performers (those in the bottom 10 percent) in Bermuda had literacy scores of 213 or lower, 233 or lower in Norway, and 209 or lower in Canada, compared with 201 or lower in the United States. The lowest performers in Switzerland also outperformed their U.S. counterparts in literacy, scoring 216 or lower.

In contrast to the results in PIRLS and PISA, there was no measurable difference in the literacy performance of men and women in the United States and in Bermuda, Canada, and Norway (Lemke et al. 2005). In Italy and Switzerland, men outperformed women. In the United States, White adults outscored Black and Hispanic adults, on average, on literacy tasks.

More countries will have collected data by 2005, allowing for additional comparisons of adult skills and knowledge. Detailed information on the results from ALL 2003 can be found in Statistics Canada and OECD (2005; available at http://www.statcan.ca/english/freepub/89-603-XIE/2005001/pdf.htm).

**Table 4. Average ALL literacy scores of adults ages 16–65, by country: 2003**

| Country | Average score |
|---|---|
| Norway | 293 |
| Bermuda | 285 |
| Switzerland | 274 |
| Canada | 281 |
| **United States** | **269** |
| Italy | 228 |

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
Average is lower than the U.S. average.

NOTE: Participants were scored on a 500-point scale.
SOURCE: Lemke, M., Miller, D., Johnston, J., Krenzke, T., Alvarez-Rojas, L., Kastberg, D., and Jocelyn, L. (2005). *Highlights From the 2003 International Adult Literacy and Lifeskills Survey (ALL)—(Revised)* (NCES 2005-117rev), table 1. Data from Statistics Canada and Organization for Economic Cooperation and Development (OECD), Adult Literacy and Lifeskills Survey (ALL), 2003.

## Mathematics

Three international assessments measure aspects of mathematical skills. The Trends in International Mathematics and Science Study (TIMSS), which assesses 4th- and 8th-grade mathematics knowledge and skills; the Program for International Student Assessment (PISA), which focuses on mathematics literacy, or the ability of 15-year-olds to apply mathematics to a wide variety of materials within a real-life context; and the Adult Literacy and Lifeskills Survey (ALL), which measures the numeracy skills of adults ages 16–65.

### TIMSS

TIMSS, which was administered in grades 4 and 8 in 1995 and 2003 and in grade 8 in 1999, is designed to measure the achievement of 4th- and 8th-graders in mathematics and science. The study is closely linked to the curricula of the participating countries, providing an indication of the degree to which students have learned the concepts of mathematics that they have studied in school. Some 46 countries participated in TIMSS in 2003, at either the 4th- or 8th-grade level, or both.

- *From 1995 to 2003, U.S. 4th-graders showed no measurable change in their mathematics performance, while the performance of 8th-graders improved.*

In mathematics, students in some countries (notably several Asian countries, such as Japan and Hong Kong, but also including the Netherlands and Belgium) consistently outperformed U.S. students, on average, regardless of the year of assessment, measure, grade, or age tested (Gonzales et al. 2004). Overall, however, the current picture of U.S. performance, as measured by TIMSS, is mixed at the 4th- and 8th-grade levels.

When comparing the United States with the other 24 countries participating at grade 4 in 2003, U.S. 4th-graders performed better, on average, than their peers in 13 countries but worse than their peers in 11 countries (table 5). TIMSS also provided scores for five mathematics content areas at grade 4: number, patterns and relationships, measurement, geometry, and data. U.S. 4th-graders performed above the international average in four of the five content areas in 2003 (all but measurement); they performed best in data and least well in measurement (Mullis et al. 2004b).

Comparing results from 1995 and 2003 suggests that while the performance of U.S. students was stable during this period, it did not keep pace with improved scores among students in several other countries (Gonzales et al. 2004). That is, of the other 14 countries participating in both 1995 and 2003, 4th-graders in more countries outperformed their U.S. peers in 2003 than in 1995, on average.

## Table 5. Average TIMSS mathematics scores of 4th- and 8th-graders, by country: 2003

| Grade 4 | |
|---|---|
| **Country** | **Average score** |
| International average | 495 |
| Singapore | 594 |
| Hong Kong SAR[1,2] | 575 |
| Japan | 565 |
| Chinese Taipei | 564 |
| Belgium-Flemish[3] | 551 |
| Netherlands[2] | 540 |
| Latvia | 536 |
| Lithuania[4] | 534 |
| Russian Federation[3] | 532 |
| England[2] | 531 |
| Hungary[3] | 529 |
| **United States[2]** | **518** |
| Cyprus | 510 |
| Moldova, Republic of | 504 |
| Italy | 503 |
| Australia[2] | 499 |
| New Zealand | 493 |
| Scotland[2] | 490 |
| Slovenia | 479 |
| Armenia | 456 |
| Norway | 451 |
| Iran, Islamic Republic of[3] | 389 |
| Philippines | 358 |
| Morocco | 347 |
| Tunisia | 339 |

| Grade 8 | |
|---|---|
| **Country** | **Average score** |
| International average[5] | 466 |
| Singapore | 605 |
| Korea, Republic of | 589 |
| Hong Kong SAR[1,2] | 586 |
| Chinese Taipei | 585 |
| Japan | 570 |
| Belgium-Flemish | 537 |
| Netherlands[2] | 536 |
| Estonia | 531 |
| Hungary[3] | 529 |
| Malaysia | 508 |
| Latvia | 508 |
| Russian Federation[3] | 508 |
| Slovak Republic | 508 |
| Australia | 505 |
| **United States[6]** | **504** |
| Lithuania[4] | 502 |
| Sweden | 499 |
| Scotland[2] | 498 |
| Israel[3] | 496 |
| New Zealand | 494 |
| Slovenia | 493 |
| Italy | 484 |
| Armenia | 478 |
| Serbia[4] | 477 |
| Bulgaria | 476 |
| Romania | 475 |
| Norway | 461 |
| Moldova, Republic of | 460 |
| Cyprus | 459 |
| Macedonia, Republic of[3] | 435 |
| Lebanon | 433 |
| Jordan | 424 |
| Iran, Islamic Republic of[3] | 411 |
| Indonesia[4] | 411 |
| Tunisia | 410 |
| Egypt | 406 |
| Bahrain | 401 |
| Palestinian National Authority | 390 |
| Chile | 387 |
| Morocco[4,6] | 387 |
| Philippines | 378 |
| Botswana | 366 |
| Saudi Arabia | 332 |
| Ghana | 276 |
| South Africa | 264 |

[1] Hong Kong is a Special Administrative Region (SAR) of the People's Republic of China.

[2] Met international guidelines for participation rates only after replacement schools were included.

[3] National defined population covers less than 95 percent of national desired population.

[4] National desired population does not cover all of the international desired population.

[5] The international average reported here differs from that reported in Mullis et al. (2004) due to the deletion of England. In Mullis et al., the reported international average is 467.

[6] Nearly satisfied guidelines for sample participation rates after replacement schools were included.

NOTE: The test for significance between the United States and the international average was adjusted to account for the U.S. contribution to the international average. Countries were required to sample students in the upper of the two grades that contained the largest number of 9-year-olds and 13-year-olds. In the United States and most countries, this corresponds to grades 4 and 8, respectively. Participants were scored on a 1,000-point scale. The international standard deviation is 100 points.

SOURCE: Gonzales, P., Guzman, J.C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., and Williams, T. (2004). *Highlights From the Trends in International Mathematics and Science Study (TIMSS) 2003* (NCES 2005-005), tables 2 and 3. Data from International Association for the Evaluation of Educational Achievement (IEA), Trends in International Mathematics and Science Study (TIMSS), 2003.

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
■ Average is lower than the U.S. average.

Students in seven countries (Singapore, Japan, Hong Kong, the Netherlands, Latvia, England, and Hungary) outscored U.S. students in 2003, while students in four countries (Singapore, Japan, Hong Kong, and the Netherlands) outscored U.S. students in 1995.

In grade 8, U.S. students showed gains in their mathematics skills and abilities. As mentioned above, TIMSS assessed 8th-graders in mathematics in 1995, 1999, and 2003. In comparison to the other 44 countries that assessed 8th-graders in 2003, U.S. 8th-graders outperformed their peers in 25 countries, on average, and were outperformed by students in 9 countries (table 5; Gonzales et al. 2004). U.S. 8th-graders had higher average scores in 2003 than in 1995, with the increase occurring primarily between 1995 and 1999. Moreover, the relative standing of U.S. 8th-graders was higher in 2003 than in 1995 in relation to students in the 21 other countries participating in TIMSS in both years. That is, of the 21 other countries participating in both 1995 and 2003, U.S. 8th-graders were outscored by their international peers, on average, in fewer countries in 2003 than in 1995 (12 countries in 1995 vs. 7 countries in 2003). In addition, TIMSS provided achievement results in five mathematics content areas: number, algebra, measurement, geometry, and data. U.S. 8th-graders improved their performance in two of these content areas (algebra and data) between 1999 and 2003.

TIMSS 2003 also examined the mathematics performance of 4th- and 8th-graders by achievement level, sex, and race/ethnicity. At both grades, 7 percent of U.S. students performed at the highest international benchmark (called "advanced") in 2003, percentages that were not measurably different from the international averages (Mullis et al. 2004b). In the United States, boys outperformed girls in mathematics at both grades 4 and 8. The gap in mathematics achievement scores between White and Black 4th- and 8th-graders narrowed between 1995 and 2003 (Gonzales et al. 2004). More detailed results for TIMSS 2003 can be found in Gonzales et al. (2004; available at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005005) and Mullis et al. (2004b; available at http://isc.bc.edu/timss2003i/mathD.html).

## PISA

While the primary emphasis of PISA in 2000 was on reading literacy, in 2003, the assessment turned its focus to mathematics literacy of 15-year-olds, with 39 countries participating. PISA uses the term *mathematics literacy* to indicate its broader focus on students' ability to apply their mathematical knowledge and skills to a range of situations they are likely to encounter in their everyday lives. Thus, unlike TIMSS, PISA does not focus exclusively on outcomes that can be directly linked to curricula, but instead emphasizes larger ideas such as space and shape or uncertainty in mathematics. PISA complements information obtained from studies such as TIMSS because it addresses whether students can apply what they have learned, both in and out of school.

- *U.S. 15-year-olds had lower average mathematics literacy scores than the OECD average and lower scores than their peers in 20 of the other 28 OECD countries participating in 2003.*

The PISA 2003 results suggest that when applying mathematical skills, U.S. 15-year-olds performed worse, on average, than many of their international peers (table 6). For this age group, the mathematics literacy performance of U.S. students was lower than the average student performance for the majority of the 28 other OECD-member countries, and below the OECD average.[4] In addition to overall mathematics literacy scores, PISA reports on performance by four broad content areas connected to overarching ideas in mathematics: space and shape,

**Table 6. Average PISA mathematics literacy scores of 15-year-olds, by country: 2003**

| Country | Average score |
|---|---|
| OECD average | 500 |
| **OECD countries** | |
| Finland | 544 |
| Korea, Republic of | 542 |
| Netherlands | 538 |
| Japan | 534 |
| Canada | 532 |
| Belgium | 529 |
| Switzerland | 527 |
| Australia | 524 |
| New Zealand | 523 |
| Czech Republic | 516 |
| Iceland | 515 |
| Denmark | 514 |
| France | 511 |
| Sweden | 509 |
| Austria | 506 |
| Germany | 503 |
| Ireland | 503 |
| Slovak Republic | 498 |
| Norway | 495 |
| Luxembourg | 493 |
| Poland | 490 |
| Hungary | 490 |
| Spain | 485 |
| **United States** | **483** |
| Portugal | 466 |
| Italy | 466 |
| Greece | 445 |
| Turkey | 423 |
| Mexico | 385 |
| **Non-OECD countries** | |
| Hong Kong-China | 550 |
| Liechtenstein | 536 |
| Macao-China | 527 |
| Latvia | 483 |
| Russian Federation | 468 |
| Serbia and Montenegro | 437 |
| Uruguay | 422 |
| Thailand | 417 |
| Indonesia | 360 |
| Tunisia | 359 |

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
■ Average is lower than the U.S. average.

NOTE: The test for significance between the United States and the Organization for Economic Cooperation and Development (OECD) average was adjusted to account for the contribution of the U.S. average to the OECD average. Because PISA is principally an OECD study, the results for non-OECD countries are displayed separately from those for the OECD countries and are not included in the OECD average. Due to low response rates, data for the United Kingdom are not included. Participants were scored on a 1,000 point scale. The international standard deviation is 100 points.

SOURCE: Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., and Jocelyn, L. (2004). *International Outcomes of Learning in Mathematics Literacy and Problem Solving: PISA 2003 Results from the U.S. Perspective* (NCES 2005-003), table 2. Data from Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2003.

change and relationships, quantity, and uncertainty. In each content area, U.S. 15-year-olds were outperformed, on average, by students in a majority of OECD countries and performed below the OECD average (Lemke et al. 2004). Fifteen-year-olds in 23 OECD countries outperformed their U.S. counterparts on the quantity measure (which focuses on quantitative reasoning and understanding of numerical patterns and measures and includes number sense, estimating, and computations) than on the other content areas measured. For the other content areas, the number of OECD countries in which students outperformed their U.S. counterparts was 16 countries on the uncertainty measure (which focuses on data and chance), 18 countries on the change and relationships measure (which focuses on the representation of change, including mathematics functions such as linear or exponential), and 20 countries on the space and shape measure (which focuses on recognizing shapes and patterns, describing and decoding visual information, and the relationship between visual representations and real shapes and images).

Further analysis of these data shows that, in 2003, the United States had a greater percentage of students than the OECD average at the lowest levels of performance in mathematics literacy and the four broad content areas (Lemke et al. 2004).

Differences in mathematics literacy performance within the United States were apparent by sex and race/ethnicity. U.S. 15-year-old females scored lower in mathematics literacy than their male counterparts, a pattern evidenced in 25 other countries (20 OECD and 5 non-OECD countries) as well (Lemke et al. 2004). Among U.S. 15-year-olds, Black and Hispanic students scored lower in mathematics literacy, on average, than their White and Asian counterparts, but Hispanic students outperformed their Black peers. More detailed information on the PISA 2003 results can be found in Lemke et al. (2004; available at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005003) and OECD (2004b; available at http://www.pisa.oecd.org/dataoecd/1/60/34002216.pdf).

### ALL

The ALL 2003 study included measures of adult numeracy skills, defined as knowledge and skills required to manage mathematical demands in diverse situations. Unlike mathematics literacy skills, numeracy skills go beyond the ability to apply arithmetic skills to include number sense, estimation, measurement, and statistics. Adults were asked to complete items that required understanding of arithmetic, proportionality, data reading and interpretation, estimation, measurement, recognition of patterns and relationships, and the ability to solve simple and multi-step problems. The goal of ALL was to ascertain the degree to which the adult population could perform tasks that they would likely encounter in everyday life and workplace situations.

- *U.S. adults outperformed adults in Italy in numeracy skills in 2003, but were outperformed by adults in Switzerland, Norway, Bermuda, and Canada.*

Mirroring the ALL 2003 results on literacy skills and knowledge, U.S. adults outperformed Italian adults in numeracy, on average. Adults in Switzerland, Norway, Bermuda, and Canada scored better, on average, than their U.S. peers (table 7).

Besides outperforming U.S. adults on average, adults in the four higher performing countries had higher numeracy scores than U.S. adults at both the high and low ends of the score distribution (Lemke et al. 2005). The highest performers (the top 10 percent of adults) had numeracy scores of 352 or higher in Switzerland, 343 or higher in Norway, 342 or higher in Bermuda, and 341 or higher in Canada, compared with 333 or higher in the United States. The lowest performers (those in the bottom 10 percent) in Bermuda and Canada had average scores of 198 or lower, 224 or lower in Norway, and 230 or lower in Switzerland, compared with 185 or lower in the United States.

Further analysis also revealed that among U.S. adults, males outperformed females in numeracy skills, and White adults outscored Black and Hispanic adults, on average (Lemke et al. 2005).

As additional countries collect ALL data, international comparisons of adults' numeracy and mathematics literacy skills should reveal more information. Details on the results from the first round of ALL can be found in Statistics Canada and OECD (2005; available at http://www.statcan.ca/english/freepub/89-603-XIE/2005001/pdf.htm).

**Table 7. Average ALL numeracy scores of adults ages 16–65, by country: 2003**

| Country | Average score |
|---|---|
| Switzerland | 290 |
| Norway | 285 |
| Bermuda | 270 |
| Canada | 272 |
| **United States** | **261** |
| Italy | 233 |

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
■ Average is lower than the U.S. average.

NOTE: Participants were scored on a 500-point scale.
SOURCE: Lemke, M., Miller, D., Johnston, J., Krenzke, T., Alvarez-Rojas, L., Kastberg, D., and Jocelyn, L. (2005). *Highlights From the 2003 International Adult Literacy and Lifeskills Survey (ALL)—(Revised)* (NCES 2005-117rev), table 1. Data from Statistics Canada and Organization for Economic Cooperation and Development (OECD), Adult Literacy and Lifeskills Survey (ALL), 2003.

## Science

Two international assessments measure aspects of science skills. The Trends in International Mathematics and Science Study (TIMSS) focuses on students' performance on science that they are likely to have encountered in school by grades 4 and 8; and the Program for International Student Assessment (PISA) focuses on the ability of 15-year-olds to apply science knowledge and skills to a variety of materials with a real-life context.

### TIMSS

As noted earlier, TIMSS was administered three times (in grades 4 and 8 in 1995 and 2003 and in grade 8 in 1999) across a range of countries. Closely linked with the curricula of the participating countries, TIMSS provides a measure of the degree to which students have learned concepts that they have encountered in school.

In every science administration, regardless of the measure, grade, or age tested, Japanese students, on average, outperformed U.S. students in science (Lemke et al. 2004; Gonzales et al. 2004). Otherwise, U.S. students' performance in science is mixed: U.S. students performed better than their international peers in some countries and worse than their peers in other countries.

- *From 1995 to 2003, U.S. 4th-graders showed no measurable change in science performance on average, while 8th-graders showed some improvement.*

According to TIMSS, over time U.S. 4th-graders are being outpaced by their international peers in science, while U.S. 8th-graders are making progress (Gonzales et al. 2004).

TIMSS 2003 science results at the 4th grade show that, on average, U.S. students performed above the international average, and had higher average scores than their peers in 16 of the 24 other participating countries (table 8). Students in three countries—Singapore, Chinese Taipei, and Japan—outperformed U.S. 4th-graders, on average. Nonetheless, U.S. 4th-graders made no significant progress between 1995 and 2003, and they did not keep pace with improved scores among students in several other countries (Gonzales et al. 2004). Fourth-graders in nine countries demonstrated improvement in their average science scores over this period. Consequently, among the 14 other countries that participated at 4th grade in both years, students in the United States outperformed students in fewer countries in 2003 than in 1995 (8 compared with 13). Taken together, these data suggest that U.S. 4th-graders are not keeping pace with their international peers in science.

## Table 8. Average TIMSS science scores of 4th- and 8th-graders, by country: 2003

| Grade 4 | |
|---|---|
| **Country** | **Average score** |
| International average | 489 |
| Singapore | 565 |
| Chinese Taipei | 551 |
| Japan | 543 |
| Hong Kong SAR[1,2] | 542 |
| England[2] | 540 |
| **United States[2]** | **536** |
| Latvia | 532 |
| Hungary[3] | 530 |
| Russian Federation[3] | 526 |
| Netherlands[2] | 525 |
| Australia[2] | 521 |
| New Zealand | 520 |
| Belgium-Flemish[3] | 518 |
| Italy | 516 |
| Lithuania[4] | 512 |
| Scotland[2] | 502 |
| Moldova, Republic of | 496 |
| Slovenia | 490 |
| Cyprus | 480 |
| Norway | 466 |
| Armenia | 437 |
| Iran, Islamic Republic of[3] | 414 |
| Philippines | 332 |
| Tunisia | 314 |
| Morocco | 304 |

| Grade 8 | |
|---|---|
| **Country** | **Average score** |
| International average[5] | 473 |
| Singapore | 578 |
| Chinese Taipei | 571 |
| Korea, Republic of | 558 |
| Hong Kong SAR[1,2] | 556 |
| Estonia | 552 |
| Japan | 552 |
| Hungary[3] | 543 |
| Netherlands[2] | 536 |
| **United States[6]** | **527** |
| Australia | 527 |
| Sweden | 524 |
| Slovenia | 520 |
| New Zealand | 520 |
| Lithuania[4] | 519 |
| Slovak Republic | 517 |
| Belgium-Flemish | 516 |
| Russian Federation[3] | 514 |
| Latvia | 512 |
| Scotland[2] | 512 |
| Malaysia | 510 |
| Norway | 494 |
| Italy | 491 |
| Israel[3] | 488 |
| Bulgaria | 479 |
| Jordan | 475 |
| Moldova, Republic of | 472 |
| Romania | 470 |
| Serbia[4] | 468 |
| Armenia | 461 |
| Iran, Islamic Republic of[3] | 453 |
| Macedonia, Republic of[3] | 449 |
| Cyprus | 441 |
| Bahrain | 438 |
| Palestinian National Authority | 435 |
| Egypt | 421 |
| Indonesia[4] | 420 |
| Chile | 413 |
| Tunisia | 404 |
| Saudi Arabia | 398 |
| Morocco[4,6] | 396 |
| Lebanon | 393 |
| Philippines | 377 |
| Botswana | 365 |
| Ghana | 255 |
| South Africa | 244 |

[1] Hong Kong is a Special Administrative Region (SAR) of the People's Republic of China.
[2] Met international guidelines for participation rates only after replacement schools were included.
[3] National defined population covers less than 95 percent of national desired population.
[4] National desired population does not cover all of the international desired population.
[5] The international average reported here differs from that reported in Martin et al. (2004) due to the deletion of England. In Martin et al., the reported international average is 474.
[6] Nearly satisfied guidelines for sample participation rates after replacement schools were included.

NOTE: The test for significance between the United States and the international average was adjusted to account for the U.S. contribution to the international average. Countries were required to sample students in the upper of the two grades that contained the largest number of 9- and 13-year-olds. In the United States and most countries, this corresponds to grades 4 and 8, respectively. Participants were scored on a 1,000-point scale. The international standard deviation is 100 points.

SOURCE: Gonzales, P., Guzman, J.C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., and Williams, T. (2004). *Highlights From the Trends in International Mathematics and Science Study (TIMSS) 2003* (NCES 2005-005), tables 8 and 9. Data from International Association for the Evaluation of Educational Achievement (IEA), Trends in International Mathematics and Science Study (TIMSS), 2003.

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
■ Average is lower than the U.S. average.

U.S. 4th-graders performed above the international average in all three science content areas (life science, physical science, and earth science) in 2003 (Martin et al. 2004). In addition, a greater percentage of U.S. students performed at the advanced TIMSS international benchmark compared with the international average (13 vs. 7 percent), but even so, the percentage of U.S. 4th-graders performing at this level declined from 1995 (when it was 19 percent).

Turning to 8th grade, U.S. students, on average, performed above the international average and had higher science scores than their peers in 32 of the 44 other participating countries in 2003 (table 8). U.S. 8th-graders improved their average science performance between 1995 and 2003, with the gain occurring primarily between 1999 and 2003 (Gonzales et al. 2004). Moreover, the relative standing of U.S. 8th-graders was higher in 2003 than in 1995 in relation to students in the 21 other countries participating in TIMSS in both years. That is, of the countries participating in both 1995 and 2003, U.S. 8th-graders outscored their international peers, on average, in 11 countries in 2003 compared with 5 countries in 1995.

Based on five science content areas measured in TIMSS (life science, chemistry, physics, earth science, and environmental science), U.S. 8th-graders showed improvement in earth science and physics between 1999 and 2003 (Gonzales et al. 2004). In 2003, a greater percentage of U.S. 8th-graders performed at the advanced TIMSS international benchmark compared with the international average (11 vs. 5 percent), though there had been no measurable change in the percentage of U.S. 8th-graders performing at this level in science since 1995.

Differences exist in science achievement within subgroups in the United States. At both 4th and 8th grade, boys outperformed girls in 2003 (Gonzales et al. 2004). Fourth-grade boys' scores declined from 1995 to 2003 while at 8th grade, both boys and girls showed improvement. White 4th- and 8th-graders had higher average science scores than their Black and Hispanic peers in 2003. At 4th grade, White student scores declined and Black student scores increased from 1995 to 2003. At 8th grade, the average scores of Black and Hispanic students increased between 1995 and 2003, while the average score of their White peers was not measurably different. Thus, the gap between White and Black students decreased at both grades. Further details on the TIMSS science results can be found in Gonzales et al. (2004; available at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005005) and Martin et al. (2004; available at http://isc.bc.edu/timss2003i/scienceD.html).

### PISA

While the primary emphases of PISA have been reading literacy in 2000 and mathematics literacy in 2003, each assessment contained a small section on the other two domains (science and mathematics or reading, respectively). PISA uses

the term *science literacy* to indicate its broader focus on students' ability to apply their science knowledge and skills to a range of situations they are likely to encounter in their everyday lives.

- *U.S. 15-year-olds scored below the OECD average in science literacy and below the average scores of students in 15 of the 28 other participating OECD countries in 2003.*

Based on PISA, U.S. 15-year-olds scored below the science literacy average of the 29 participating OECD countries (table 9). Students in 15 OECD countries had higher average scores than students in the United States, and 6 OECD countries had lower average scores. No information about U.S. performance on specific sci-

**Table 9. Average PISA science literacy scores of 15-year-olds, by country: 2003**

| Country | Average score |
|---|---|
| OECD average | 500 |
| **OECD countries** | |
| Finland | 548 |
| Japan | 548 |
| Korea, Republic of | 538 |
| Australia | 525 |
| Netherlands | 524 |
| Czech Republic | 523 |
| New Zealand | 521 |
| Canada | 519 |
| Switzerland | 513 |
| France | 511 |
| Belgium | 509 |
| Sweden | 506 |
| Ireland | 505 |
| Hungary | 503 |
| Germany | 502 |
| Poland | 498 |
| Slovak Republic | 495 |
| Iceland | 495 |
| **United States** | **491** |
| Austria | 491 |
| Spain | 487 |
| Italy | 487 |
| Norway | 484 |
| Luxembourg | 483 |
| Greece | 481 |
| Denmark | 475 |
| Portugal | 468 |
| Turkey | 434 |
| Mexico | 405 |
| **Non-OECD countries** | |
| Hong Kong-China | 540 |
| Liechtenstein | 525 |
| Macao-China | 525 |
| Russian Federation | 489 |
| Latvia | 489 |
| Uruguay | 438 |
| Serbia and Montenegro | 436 |
| Thailand | 429 |
| Indonesia | 395 |
| Tunisia | 385 |

■ Average is higher than the U.S. average.
□ Average is not measurably different from the U.S. average.
■ Average is lower than the U.S. average.

NOTE: The test for significance between the United States and the Organization for Economic Cooperation and Development (OECD) average was adjusted to account for the contribution of the U.S. average to the OECD average. Because PISA is principally an OECD study, the results for non-OECD countries are displayed separately from those for the OECD countries and are not included in the OECD average. Due to low response rates, data for the United Kingdom are not included. Participants were scored on a 1,000-point scale. The international standard deviation is 100 points.

SOURCE: Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., and Jocelyn, L. (2004). *International Outcomes of Learning in Mathematics Literacy and Problem Solving: PISA 2003 Results from the U.S. Perspective* (NCES 2005-003), table B-17. Data from Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2003.

ence topics was available in PISA, but science literacy will be the primary domain covered in 2006, after which detailed information about U.S. performance will be available. Further details on the PISA science literacy results can be found in Lemke et al. (2004; available at http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005003) and OECD (2004b; available at http://www.pisa.oecd.org/dataoecd/1/60/34002216.pdf).

## Conclusion

Based on the results of recent international assessments, measures of students' and adults' skills and abilities in reading, mathematics, and science present a mixed picture (table 10). U.S. students perform relatively well in reading literacy compared with their peers around the world, including those in highly industrialized countries (based on PIRLS and PISA data). In addition, U.S. students perform relatively well in mathematics at the lower grades compared to their

**Table 10. U.S. performance on international assessments of mathematics, science, and reading relative to other countries**

| Subject and grade or age | Number of countries[1] | Number of countries with average score relative to the United States | | |
|---|---|---|---|---|
| | | Significantly higher | Not significantly different | Significantly lower |
| Reading | | | | |
| 4th-graders (2001) | 34 | 3 | 8 | 23 |
| 15-year-olds (2000) | 30 | 3 | 20 | 7 |
| Mathematics | | | | |
| 4th-graders (2003) | 24 | 11 | 0 | 13 |
| 8th-graders (2003) | 44 | 9 | 10 | 25 |
| 15-year-olds (2003) | 38 | 23 | 4 | 11 |
| Science | | | | |
| 4th-graders (2003) | 24 | 3 | 5 | 16 |
| 8th-graders (2003) | 44 | 7 | 5 | 32 |
| 15-year-olds (2003) | 38 | 18 | 9 | 11 |
| Adult literacy | | | | |
| Ages 16–65 (2003) | 5 | 4 | 0 | 1 |
| Adult numeracy | | | | |
| Ages 16–65 (2003) | 5 | 4 | 0 | 1 |

[1] Includes those countries with approved data appearing in reports. Total excludes the United States.

SOURCE: International Association for the Evaluation of Educational Achievement (IEA), Progress in International Reading Literacy Study (PIRLS), 2001; Trends in International Mathematics and Science Study (TIMSS), 2003; Statistics Canada and Organization for Economic Cooperation and Development (OECD), Adult Literacy and Lifeskills Survey (ALL), 2003; OECD, Program for International Student Assessment (PISA), 2003, previously unpublished tabulation (October 2005).

peers in other countries—though the data suggest that their performance may not be keeping pace with that of their peers—and are showing improvement in the middle school years (based on TIMSS data). However, when older U.S. students are asked to apply what they have learned in mathematics, they demonstrate less ability than most of their peers in other highly industrialized countries (based on PISA data). In science, U.S. students also perform relatively well at the lower grades compared with their peers in other countries—though, again, the data suggest that their performance may not be keeping pace with their peers—and are showing improvement in the middle school years (based on TIMSS data). This progress, though, may not carry over to tasks that are embedded in a real-life context: when asked to apply scientific skills, U.S. 15-year-olds performed worse than about half of their international peers (based on PISA data). Data on the literacy and numeracy skills of U.S. adults in comparison with their peers from other countries are fairly limited, but suggest that the skills of U.S. adults do not compare favorably (based on ALL data).

Future data collections for TIMSS, PIRLS, and PISA will provide additional opportunities to compare the performance of U.S. students in mathematics, science, and reading to international benchmarks.

## Notes

[1] PISA assesses each subject every 3 years. However, each assessment cycle focuses on one particular subject. In 2000, the focus was on reading literacy; in 2003, the focus was on mathematics literacy; in 2006, PISA will focus on science literacy.

[2] The international results may differ from trends reported in the National Assessment of Educational Progress (NAEP) and other national assessments. For further discussion of the differences between NAEP and the international student assessments, see http://nces.ed.gov/TIMSS/pdf/naep_timss_pisa_comp.pdf.

[3] The international average reported for PISA is based on results only from the OECD-member countries. Because PISA is primarily an OECD study, results for non-OECD-member countries are displayed separately from those of OECD countries and are not included in the OECD average.

[4] The international average reported for PISA is based on results only from the OECD-member countries. Because PISA is primarily an OECD study, results for non-OECD-member countries are displayed separately from those of OECD countries and are not included in the OECD average.

## References


Adams, R. (Ed.). (2005). *PISA 2003 Technical Report*. Paris: Organization for Economic Cooperation and Development.

Buchmann, C. (2002). Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges. In A.C. Porter and A. Gamoran (Eds.), *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150–197). Washington, DC: National Academy Press.

Coleman, J.S., Kelly, D.L., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., and York, R.L. (1966). *Equality of Educational Opportunity*. U.S. Department of Health, Education, and Welfare. Washington, DC: U.S. Government Printing Office.

Education Sciences Reform Act of 2002, P.L. 107-279, 116 Stat. 1940 (2002).

Entwisle, D.R., and Alexander, K. (1993). Entry Into School: The Beginning School Transition and Educational Stratification in the United States. *Annual Review of Sociology, 19*: 401–423.

Ganzeboom, H.B.G., DeGraaf, P.M., and Treiman, D.J. (1992). A Standard International Socio-Economic Index of Occupational Status. *Social Science Research, 21*: 1–56.

Gonzales, P., Guzman, J., Partelow, L., Pahlke, E., Jocelyn, L., Mak, K., Kastberg, D., and Williams, T. (2004). *Highlights From the Trends in International Mathematics and Science Study (TIMSS) 2003* (NCES 2005-005). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Hampden-Thompson, G., and Johnston, J. (2006). *Variation in the Relationship Between Nonschool Factors and Student Achievement on International Assessments* (NCES 2006-014). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Jencks, C., Bartlett, S., Corcoran, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, K., Mueser, P., Olneck, M., Schwarz, J., Ward, S., and Williams, J. (1979). *Who Gets Ahead? The Determinants of Economic Success in America*. New York: Basic Books.

Lemke, M., Calsyn, C., Lippman, L., Jocelyn, L., Kastberg, D., Liu, Y.Y., Roey, S., Williams, T., Kruger, T., and Bairu, G. (2001). *Outcomes of Learning: Results From the 2000 Program for International Student Assessment of 15-Year-Olds in Reading, Mathematics, and Science Literacy* (NCES 2002-115). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Lemke, M., Miller, D., Johnston, J., Krenzke, T., Alvarez-Rojas, L., Kastberg, D., and Jocelyn, L. (2005). *Highlights From the 2003 International Adult Literacy and Lifeskills Survey (ALL)-(Revised)* (NCES 2005-117rev). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., and Jocelyn, L. (2004). *International Outcomes of Learning in Mathematics Literacy and Problem Solving: PISA 2003 Results From the U.S. Perspective* (NCES 2005-003). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Martin, M.O., Mullis, I.V.S., and Chrostowski, S.J. (Eds.). (2004). *TIMSS 2003 Technical Report*. Chestnut Hill, MA: Boston College.

Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., and Chrostowski, S.J. (2004). *TIMSS 2003 International Science Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grade*. Chestnut Hill, MA: Boston College.

Martin, M.O., Mullis, I.V.S., and Kennedy, A.M. (Eds.). (2003). *PIRLS 2001 Technical Report*. Chestnut Hill, MA: Boston College.

McLanahan, S.S., and Sandefur, G. (1994). *Growing Up With a Single Parent: What Hurts, What Helps?* Cambridge, MA: Harvard University Press.

Mullis, I.V.S., Kennedy, A.M., Martin, M.O., and Sainsbury, M. (2004a). *PIRLS 2006 Assessment Framework and Specifications*. Chestnut Hill, MA: Boston College.

Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., and Chrostowski, S.J. (2004b). *TIMSS 2003 International Mathematics Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grade*. Chestnut Hill, MA: Boston College.

Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., and Kennedy, A.M. (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools in 35 Countries*. Chestnut Hill, MA: Boston College.

Ogle, L.T., Sen, A., Pahlke, E., Jocelyn, L., Kastberg, D., Roey, S., and Williams, T. (2003). *International Comparisons in Fourth-Grade Reading Literacy: Findings From the Progress in International Reading Literacy Study (PIRLS) of 2001* (NCES 2003-073). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Organization for Economic Cooperation and Development (OECD). (1999). *Measuring Student Knowledge and Skills: A New Framework for Assessment*. Paris: Author.

Organization for Economic Cooperation and Development (OECD). (2001). *Knowledge and Skills for Life: First Results From the OECD Program for International Student Assessment*. Paris: Author.

Organization for Economic Cooperation and Development (OECD). (2004a). *Education at a Glance: OECD Indicators 2004*. Paris: Author.

Organization for Economic Cooperation and Development (OECD). (2004b). *Learning for Tomorrow's World: First Results From PISA 2003*. Paris: Author.

Schmid, C.L. (2001). Educational Achievement, Language-Minority Students, and the New Second Generation. *Sociology of Education, 74*: 71–87.

Shavit, Y., and Blossfield, H.P. (Eds.). (1993). *Persistent Inequality: Changing the Educational Stratification in Thirteen Countries*. Boulder, CO: Westview.

Statistics Canada. (2005). *The Adult Literacy and Lifeskills Survey: A Technical Report*. Ottawa: Author.

Statistics Canada and Organization for Economic Cooperation and Development (OECD). (2005). *Learning a Living: First Results of the Adult Literacy and Life Skills Survey*. Ottawa and Paris: Authors.

United Nations Development Program (UNDP). (2005). *Human Development Report 2005. International Development at a Crossroads: Aid, Trade, and Security in an Unequal World*. New York: Oxford University Press.

## For more information, visit the following websites:

TIMSS: http://nces.ed.gov/timss or http://www.timss.org

PIRLS: http://nces.ed.gov/surveys/pirls or http://www.pirls.org

PISA: http://nces.ed.gov/surveys/pisa or http://www.pisa.oecd.org

ALL: http://nces.ed.gov/surveys/all

International Comparisons: http://nces.ed.gov/surveys/international

## Technical Notes and Methodology

### Data Source and Estimates

The data in this special analysis were obtained from statistical samples of entire populations. Estimating the size of the total population or subpopulations from a data source based on a sample of the entire population requires consideration of several factors before the estimates become meaningful. However conscientious an organization may be in collecting data from a sample of a population, there will always be some margin of error in estimating the size of the actual total population or subpopulation because the data are available from only a portion of the total population. Consequently, data from samples can provide only an estimate of the true or actual value. The margin of error or the range of the estimate depends on several factors, such as the amount of variation in the responses, the size and representativeness of the sample, and the size of the subgroup for which the estimate is computed. The magnitude of this margin of error is measured by what statisticians call the "standard error" of an estimate.

### Standard Errors

The standard error for each estimate in this special analysis was calculated in order to determine the "margin of error" for these estimates. The standard errors for all the estimated means and percentages reported in the figures and tables of the special analysis can be found in the main printed volume of *The Condition of Education 2006*, appendix 3, Standard Error Tables and on *The Condition of Education* website.

An estimate with a smaller standard error provides a more reliable estimate of the true value than an estimate with a higher standard error. Standard errors tend to diminish in size as the size of the sample (or subsample) increases. Consequently, for the same data, standard errors will almost always be smaller for groups who represent a larger proportion of the population.

### Data Analysis and Interpretation

Due to standard errors, caution is warranted when drawing conclusions about the size of one population estimate in comparison to another or whether a time series of population estimates is increasing, decreasing, or staying about the same. Although one estimate of the population size may be larger than another, a statistical test may reveal that there is no measurable difference between the two estimates due to their uncertainty.

Whether differences in means or percentages are statistically significant can be determined using the standard errors of the estimates. When differences are statistically significant, the probability that the difference occurred by chance is usually small; for example, it might be about 5 times out of 100. Some details about the method primarily used for determining whether the difference between two means is statistically significant are presented in *The Condition of Education's* introduction to appendix 3, Standard Error Tables, available at http://nces.ed.gov/programs/coe/guide/g3b.asp.

For this special analysis, differences between means or percentages (including increases or decreases) are stated only when they are statistically significant. To determine whether differences reported are statistically significant, two-tailed *t* tests, at the .05 level, were used. The *t* test formula for determining statistical significance was adjusted when the samples being compared were dependent. When the difference between means or percentages was not statistically significant, tests of equivalence were run. An equivalence test determines the probability (generally at the .15 level) that the means or percentages are statistically equivalent: that is, with the margin of error that the two estimates are not substantively different. When the difference was found to be equivalent, language such as *x* and *y* "were similar" or "about the same" was used. Otherwise, the two estimates were reported as being "not measurably different."

## Rounding and Other Considerations

Although values reported in the supplemental tables are rounded to one decimal place (e.g., 76.5 percent), values reported in this special analysis are rounded to whole numbers (with any value of 0.5 or above rounded to the next highest whole number). Due to rounding, total percentages sometimes differ from the sum of the reported parts, which may, for example, equal 99 or 101 percent, rather than the percentage distribution's total of 100 percent.