Appendix A: Sample Science Items From PISA 2006

This section presents sample items used in the PISA 2006 science assessment. These items serve to illustrate the various competencies and types of scientific knowledge measured by PISA, as well as the different difficulty levels at which students were tested. For more information about the science literacy subject area or additional examples of science literacy items, refer to *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006* (OECD 2006).

Exhibit A-1 summarizes the distribution of the sample items across the PISA knowledge areas and competency types, along with their associated difficulty. Grand Canyon question 3, for example, tests student knowledge of science in earth and space systems under the *explaining phenomena scientifically* competency. This question has a difficulty of 411 (level 2) on the combined science literacy scale, requiring students to know that fossils from organisms that lived long ago may be exposed when sea levels recede.

Exhibit A-1. Map of selected science items in PISA 2006

			Competency		
			Identifying scientific issues	**Explaining phenomena scientifically**	**Using scientific evidence**
Knowledge	**Knowledge of science** (scientific content)	**Physical systems**		Acid Rain Q1 (506)	Acid Rain Q2 (460)
		Living systems			
		Earth and space systems		Grand Canyon Q2 (451) Grand Canyon Q3 (411)	
		Technology systems			
	Knowledge about science (scientific process)	**Scientific inquiry**	Acid Rain Q3 (513) (partial credit) Acid Rain Q3 (717) (full credit) Sunscreens Q1 (588) Sunscreens Q2 (499) Sunscreens Q3 (574) Grand Canyon Q1 (485)		
		Scientific explanation			Sunscreens Q4 (616) (partial credit) Sunscreens Q4 (629) (full credit)

NOTE: Numbers in parentheses refer to the score or proficiency level associated with the item. To reach a particular proficiency level, a student must correctly answer a majority of items at that level. Students were classified into science literacy levels according to their scores. Exact cut point scores are as follows: below level 1 (a score less than or equal to 334.94); level 1 (a score greater than 334.94 and less than or equal to 409.54); level 2 (a score greater than 409.54 and less than or equal to 484.14); level 3 (a score greater than 484.14 and less than or equal to 558.73); level 4 (a score greater than 558.73 and less than or equal to 633.33); level 5 (a score greater than 633.33 and less than or equal to 707.93); and level 6 (a score greater than 707.93).
SOURCE: Organization for Economic Cooperation and Development (OECD). (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006.* Paris: Author.

Exhibit A-2. Example A of PISA 2006 science assessment

ACID RAIN

Below is a photo of statues called Caryatids that were built on the Acropolis in Athens more than 2500 years ago. The statues are made of a type of rock called marble. Marble is composed of calcium carbonate.

In 1980, the original statues were transferred inside the museum of the Acropolis and were replaced by replicas. The original statues were being eaten away by acid rain.

Question 1: ACID RAIN

Normal rain is slightly acidic because it has absorbed some carbon dioxide from the air. Acid rain is more acidic than normal rain because it has absorbed gases like sulfur oxides and nitrogen oxides as well.

Where do these sulfur oxides and nitrogen oxides in the air come from?

Sulfur oxides and nitrogen oxides are put in the air from pollution and burning fossil fuels. **(full credit)**

Sulfur oxides and nitrogen oxides come from the pollution in the air. **(partial credit)**

The effect of acid rain on marble can be modeled by placing chips of marble in vinegar overnight. Vinegar and acid rain have about the same acidity level. When a marble chip is placed in vinegar, bubbles of gas form. The mass of the dry marble chip can be found before and after the experiment.

Question 2: ACID RAIN

A marble chip has a mass of 2.0 grams before being immersed in vinegar overnight. The chip is removed and dried the next day. What will the mass of the dried marble chip be?

- (A) Less than 2.0 grams
- (B) Exactly 2.0 grams
- (C) Between 2.0 and 2.4 grams
- (D) More than 2.4 grams

Question 3: ACID RAIN

Students who did this experiment also placed marble chips in pure (distilled) water overnight.

Explain why the students include this step in their experiment.

To provide a control. Maybe the liquid is the problem with marble being eaten away. **(full credit)**

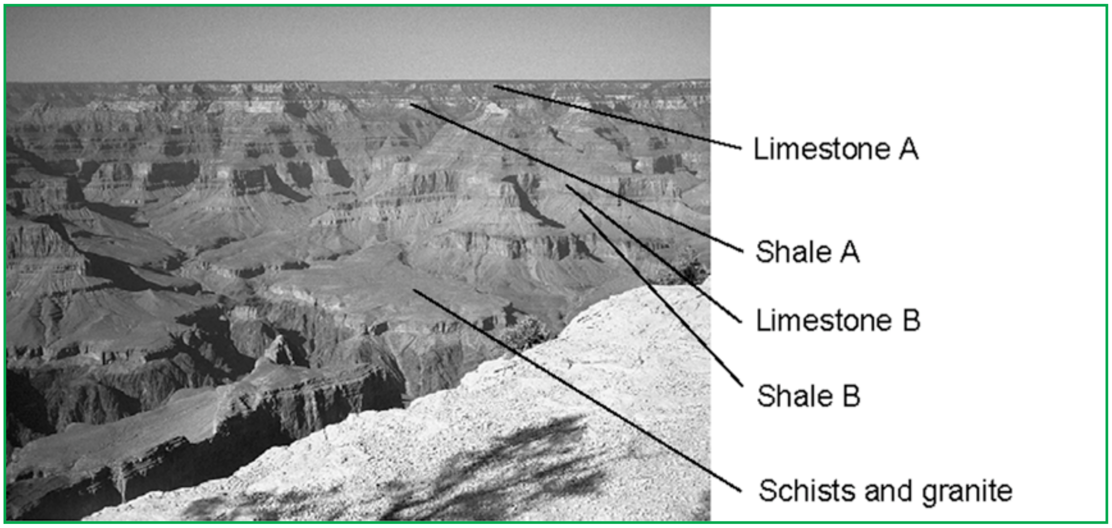
To see the difference between acidic and non-acidic water. **(partial credit)**

Exhibit A-3. Example B of PISA 2006 science assessment

THE GRAND CANYON

The Grand Canyon is located in a desert in the USA. It is a very large and deep canyon containing many layers of rock. Sometime in the past, movements in the Earth's crust lifted these layers up. The Grand Canyon is now 1.6 km deep in parts. The Colorado River runs through the bottom of the canyon.

See the picture below of the Grand Canyon taken from its south rim. Several different layers of rock can be seen in the walls of the canyon.



Question 1: THE GRAND CANYON

About five million people visit the Grand Canyon national park every year. There is concern about the damage that is being caused to the park by so many visitors.

Can the following questions be answered by scientific investigation? Circle "Yes" or "No" for each question.

Can this question be answered by scientific investigation?	Yes or No?
How much erosion is caused by use of the walking tracks?	● / No
Is the park area as beautiful as it was 100 years ago?	Yes / ●

Question 2: THE GRAND CANYON

The temperature in the Grand Canyon ranges from below 0° C to over 40° C. Although it is a desert area, cracks in the rocks sometimes contain water. How do these temperature changes and the water in rock cracks help to speed up the breakdown of rocks?

- (A) Freezing water dissolves warm rocks.
- (B) Water cements rocks together.
- (C) Ice smoothes the surface of rocks.
- ● Freezing water expands in the rock cracks.

Question 3: THE GRAND CANYON

There are many fossils of marine animals, such as clams, fish and corals, in the Limestone A layer of the Grand Canyon. What happened millions of years ago that explains why such fossils are found there?

- (A) In ancient times, people brought seafood to the area from the ocean.
- (B) Oceans were once much rougher and sea life washed inland on giant waves.
- ● An ocean covered this area at that time and then receded later.
- (D) Some sea animals once lived on land before migrating to the sea.

Exhibit A-4. Example C of PISA 2006 science assessment

SUNSCREENS

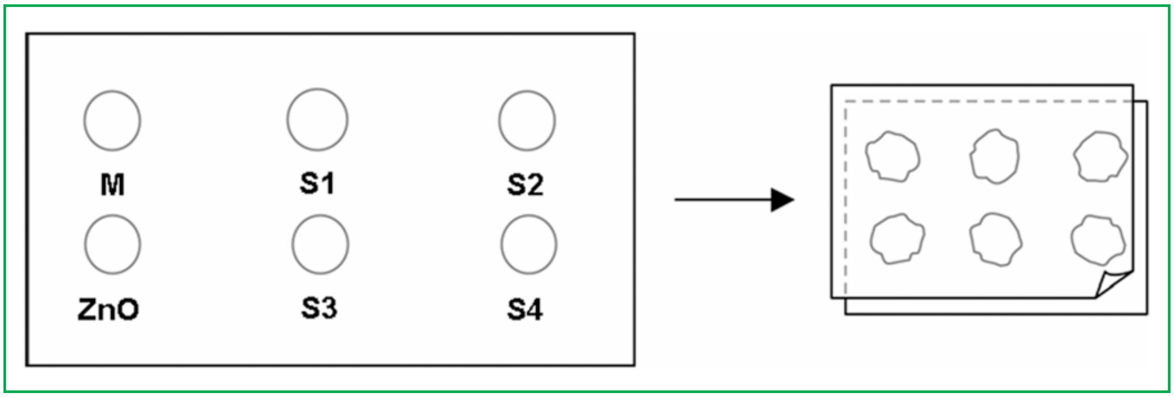
Mimi and Dean wondered which sunscreen product provides the best protection for their skin. Sunscreen products have a *Sun Protection Factor (SPF)* that shows how well each product absorbs the ultraviolet radiation component of sunlight. A high SPF sunscreen protects skin for longer than a low SPF sunscreen.

Mimi thought of a way to compare some different sunscreen products. She and Dean collected the following:

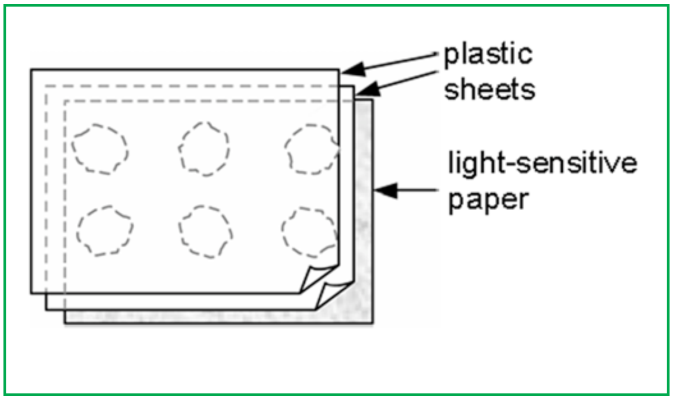
- two sheets of clear plastic that do not absorb sunlight;
- one sheet of light-sensitive paper;
- mineral oil (M) and a cream containing zinc oxide (ZnO); and
- four different sunscreens that they called S1, S2, S3, and S4.

Mimi and Dean included mineral oil because it lets most of the sunlight through, and zinc oxide because it almost completely blocks sunlight.

Dean placed a drop of each substance inside a circle marked on one sheet of plastic, then put the second plastic sheet over the top. He placed a large book on top of both sheets and pressed down.



Mimi then put the plastic sheets on top of the sheet of light-sensitive paper. Light-sensitive paper changes from dark gray to white (or very light gray), depending on how long it is exposed to sunlight. Finally, Dean placed the sheets in a sunny place.



Question 1: SUNSCREENS

Which one of these statements is a scientific description of the role of the mineral oil and the zinc oxide in comparing the effectiveness of the sunscreens?

- (A) Mineral oil and zinc oxide are both factors being tested.
- (B) Mineral oil is a factor being tested and zinc oxide is a reference substance.
- (C) Mineral oil is a reference substance and zinc oxide is a factor being tested.
- ● Mineral oil and zinc oxide are both reference substances.

Question 2: SUNSCREENS

Which one of these questions were Mimi and Dean trying to answer?

- ● How does the protection for each sunscreen compare with the others?
- (B) How do sunscreens protect your skin from ultraviolet radiation?
- (C) Is there any sunscreen that gives less protection than mineral oil?
- (D) Is there any sunscreen that gives more protection than zinc oxide?

Question 3: SUNSCREENS

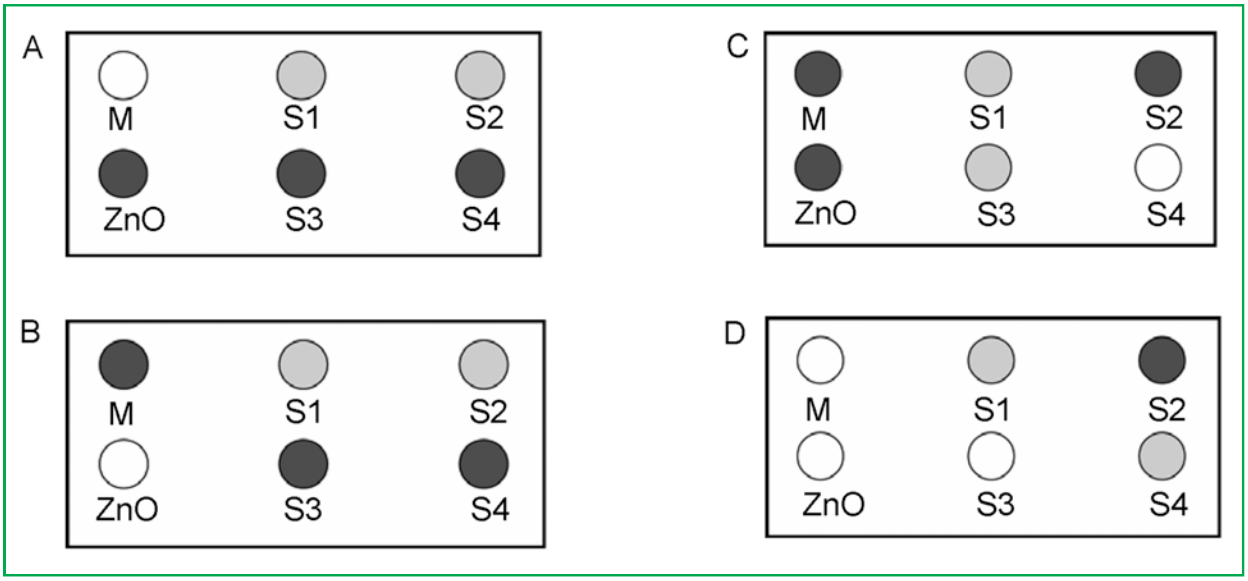
Why was the second sheet of plastic pressed down?

- (A) To stop the drops from drying out.
- (B) To spread the drops out as far as possible.
- (C) To keep the drops inside the marked circles.
- ● To make the drops the same thickness.

Question 4: SUNSCREENS

The light-sensitive paper is a dark gray and fades to a lighter gray when it is exposed to some sunlight, and to white when exposed to a lot of sunlight.

Which one of these diagrams shows a pattern that might occur? Explain why you chose it.



Answer: A.

Explanation: Mineral oil lets in a lot of sunlight, so that spot on the paper should be the lightest. Zinc oxide almost completely blocks sunlight, so that should be the darkest spot on the paper. **(full credit)**

Answer: A.

Explanation: Because ZnO blocks the light and M absorbs it. **(partial credit)**

Appendix B: Technical Notes

The Program for International Student Assessment (PISA) is a system of international assessments that measures 15-year-olds' performance in reading literacy, mathematics literacy, and science literacy. PISA was first implemented in 2000 and is carried out every 3 years by the Organization for Economic Cooperation and Development (OECD). In this third cycle, PISA 2006, science literacy was the major focus. This appendix describes features of the PISA 2006 survey methodology, including sample design, test design, scoring, data reliability, and analysis variables. For further details about the assessment and any of the topics discussed here, see the OECD's *PISA 2006 Technical Report* (Adams in press) and the *PISA 2003 Technical Report* (Adams 2004).

International Requirements for Sampling, Data Collection, and Response Rates

To provide valid estimates of student achievement and characteristics, the sample of PISA students had to be selected in a way that represented the full population of 15-year-old students in each jurisdiction. The international desired population in each jurisdiction consisted of 15-year-olds attending both publicly and privately controlled schools in grade 7 and higher. A minimum of 4,500 students from a minimum of 150 schools was required. Within schools, a sample of 35 students was to be selected in an equal probability sample unless fewer than 35 students age 15 were available (in which case all students were selected). International standards required that students in the sample be 15 years and 3 months to 16 years and 2 months at the beginning of the testing period. The testing period suggested by the OECD was between March 1, 2006, and August 31, 2006, and was required not to exceed 42 days.[1] Each jurisdiction collected its own data, following international guidelines and specifications.

The school response rate target was 85 percent for all jurisdictions. A minimum of 65 percent of schools from the original sample of schools were required to participate for a jurisdiction's data to be included in the international database. Jurisdictions were allowed to use replacement schools (selected during the sampling process) to increase the response rate once the 65 percent benchmark had been reached.

PISA 2006 also required a minimum participation rate of 80 percent of sampled students from schools within each jurisdiction. A student was considered to be a participant if he or she participated in the first testing session or a follow-up or makeup testing session. Data from jurisdictions not meeting this requirement could be excluded from international reports.

Exclusion guidelines allowed for 5 percent at the school level for approved reasons (for example, remote regions or very small schools) and 2 percent for special education schools. Overall estimated student exclusions were to be under 5 percent. PISA's intent was to be as inclusive as possible. A special 1-hour test booklet was developed for use in special education classrooms, and jurisdictions could choose whether or not to use the booklet. The United States chose not to use this special test booklet.

[1] The United States, the United Kingdom (except Scotland), and Bulgaria were given permission to move the testing dates to the fall in an effort to improve response rates. The range of eligible birthdates was adjusted so that the mean age remained the same. In 2003, the United States conducted PISA in the spring and fall and found no significant difference in student performance between the two time points.

Schools used the following international guidelines on possible student exclusions:

- **Functionally disabled students.** These were students with a moderate to severe permanent physical disability such that they cannot perform in the PISA testing environment
- **Intellectually disabled students.** These were students with a mental or emotional disability and who have been tested as cognitively delayed or who are considered in the professional opinion of qualified staff to be cognitively delayed such that they cannot perform in the PISA testing situation.
- **Students with insufficient language experience.** These were students who meet the three criteria of not being native speakers in the assessment language, having limited proficiency in the assessment language, and receiving less than 1 year of instruction in the assessment language.

Quality monitors from the PISA Consortium visited a sample of schools in every jurisdiction to ensure that testing procedures were carried out in a consistent manner.

Sampling, Data Collection, and Response Rates in the United States

The PISA 2006 school sample was drawn for the United States in June 2005 by the international PISA Consortium. Unlike the 2000 PISA sample, which had a three-stage design, the U.S. sample for 2006 followed the model used in 2003, which was a two-stage sampling process with the first stage a sample of schools and the second stage a sample of students within schools. For PISA 2000, the U.S. school sample had the selection of a sample of geographic Primary Sampling Units (PSUs) as the first stage of selection. The sample was not clustered at the geographic level for PISA 2006 or PISA 2003. This change was made in an effort to reduce the design effects observed in the 2000 data and to spread the respondent burden across school districts as much as possible. The sample design for PISA 2006 was a stratified systematic sample, with sampling probabilities proportional to measures of school size. The PISA sample was stratified into two explicit groups: large schools and small schools. The frame was implicitly stratified (i.e., sorted for sampling) by five categorical stratification variables: grade span of the school (five levels), control of school (public or private), region of the country (Northeast, Central, West, Southeast)[2], type of location relative to populous areas (eight levels), and proportion of non-White students (above or below 15 percent). The last variable used for sorting within the implicit stratification was by estimated enrollment of 15-year-olds based on grade enrollments.

Following the PISA guidelines at the same time as the PISA sample was selected, replacement schools were identified by assigning the two schools neighboring the sampled school in the frame as replacements. There were several constraints on the assignment of substitutes. One sampled school was not allowed to substitute for another, and a given school could not be assigned to substitute for more than one sampled school. Furthermore, substitutes were required to be in the same implicit stratum as the sampled school. If the sampled school was the first or last school in the stratum, then the second school following or preceding the sampled school was identified as the substitute. One was designated a first replacement and the other a second replacement. If an original school refused to participate, the first replacement was then contacted. If that school also refused to participate, the second school was then contacted.

The U.S. PISA 2006 school sample consisted of 236 schools. This number was increased from the international minimum requirement of 150 to offset school nonresponse and reduce design effects. The schools were selected with probability proportionate to the school's estimated enrollment of 15-year-olds from the school frame with 2003–04 school year data. The data for public schools were from the 2003–04 Common Core of Data (CCD), and the data for private schools were from the 2003–04 Private School Universe Survey (PSS). Any school containing at least one 7th- through 12th-grade class

[2] The Northeast region consists of Connecticut, Delaware, the District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The Central region consists of Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, Wisconsin, and South Dakota. The West region consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oklahoma, Oregon, Texas, Utah, Washington, and Wyoming. The Southeast region consists of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia.

as of school year 2003–04 was included in the school sampling frame. Participating schools provided lists of 15-year-old students typically in August or September 2006, and a sample of 42 students was selected within each school in an equal probability sample. The overall sample design for the United States was intended to approximate a self-weighting sample of students as much as possible, with each 15-year-old student having an equal probability of being selected.

In the United States, for a variety of reasons reported by school administrators (such as increased testing requirements at the national, state, and local levels; concerns about the timing of the PISA assessment; and loss of learning time), many schools in the original sample declined to participate. The United States has had difficulty meeting the minimum response rate standards in prior years and, in 2003, opened a second data collection period in the fall of 2003 with the agreement of the PISA Consortium. A bias analysis conducted in 2003 found no statistically significant session effects between the spring and fall assessments. To improve response rates and better accommodate school schedules, the PISA 2006 data collection was scheduled from September to November 2006 with the agreement of the PISA Consortium. After experiencing similar difficulties in 2003, the United Kingdom (except Scotland) and Bulgaria also opted for a fall data collection period for PISA 2006.

Of the 236 original sampled schools, 209 were eligible (18 schools did not have any 15-year-olds enrolled, 5 had closed, and 4 were alternative schools for behavioral issues where students returned to a base school after a short period of time), and 145 agreed to participate. The weighted school response rate before replacement was 69 percent, placing the United States in the "intermediate" response rate category. The weighted school response rate before replacement is given by the formula

$$\text{weighted school response rate before replacement} = \frac{\sum_{i \in Y} W_i E_i}{\sum_{i \in (Y \cup N)} W_i E_i},$$

where Y denotes the set of responding original sample schools with age-eligible students; N denotes the set of eligible nonresponding original sample schools; W_i denotes the base weight for school i; $W_i = 1/P_i$, where P_i denotes the school selection probability for school i; and E_i denotes the enrollment size of age-eligible students, as indicated in the sampling frame.

In addition to the 145 participating original schools, 21 replacement schools also participated for a total of 166 participating schools, or a 79 percent overall response rate.[3] The participation of the additional schools did not change the classification of the United States in the intermediate response rate category.

A total of 6,796 students were sampled for the assessment. Of these students, 37 were deemed ineligible because of their enrolled grades or birthdays and 326 were deemed ineligible because they had left the school. These students were removed from the sample. Of the eligible 6,433 sampled students, an additional 254 were excluded using the decision criteria described earlier, for a weighted exclusion rate of 3.8 percent at the student level. Combined with the 0.5 percent of students excluded at the school level, before sampling, the overall exclusion rate for the United States was 4.3 percent.

Of the 6,179 remaining sampled students, a total of 5,611 participated in the assessment in the United States. An overall weighted student response of 91 percent was achieved.

A bias analysis was conducted in the United States to address potential problems in the data owing to school nonresponse (Krotki and Bland 2008). To compare PISA respondents and nonrespondents, it was necessary to match the sample of schools back to the sample frame to detect as many characteristics as possible that might provide information about the presence of nonresponse bias. Comparing frame characteristics for respondents and nonrespondents is not always a good measure of nonresponse bias if the characteristics are unrelated or weakly related to more substantive items in the survey; however, this was the only approach available given that no comparable school- or student-level achievement data were available. Frame

[3] Response rates reported here are based on the formula used in the international report and are not consistent with NCES standards. A more conservative way to calculate the response rate would be to include replacement schools that participated in the denominator as well as the numerator, and to add replacement schools that were hard refusals to the denominator. This results in a response rate of 67.5 percent.

characteristics were taken from the 2003–04 CCD for public schools and from the 2003–04 PSS for private schools. For categorical variables, response rates by characteristics were calculated. The hypothesis of independence between the characteristics and response status was tested using a Rao-Scott modified chi-square statistic. For continuous variables, summary means were calculated.

The 95 percent confidence interval for the difference between the mean for respondents and the overall mean was tested to see whether or not it included zero. In addition to these tests, logistic regression models were employed to identify whether any of the frame characteristics were significant in predicting response status. All analyses were performed using SUDAAN, a statistical software package. The school base weights used in these analyses did not include a nonresponse adjustment factor. The base weight for each original school was the reciprocal of its selection probability. The base weight for each replacement school was set equal to the base weight of the original school it replaced.

Characteristics available for public and private schools included public/private affiliation, community type, region, number of age-eligible students enrolled, total number of students, and percentage of various racial/ethnic groups (Asian or Pacific Islander, non-Hispanic; Black, non-Hispanic; Hispanic; American Indian or Alaska Native, non-Hispanic; and White, non-Hispanic). The percentage of students eligible for free or reduced-price lunch was available for public schools only. For the original sample of schools, only one variable, community type (urban, suburban, or rural), showed a relationship to response status in tests of independence; school location in an urban fringe area or large town was associated with nonresponse. Using the same analytic procedure for the final sample (including replacement schools), tests of independence again showed that responding schools were less likely to be located in urban fringe areas or large towns. This same variable was found to be significant in the logistic regression model predicting response.

The international consortium adjusted the school base weights for nonresponse, as discussed in the section on weighting. Three variables were used that had been identified as stratification variables at the time of sampling: school control (public/private), census region, and community type (urban, suburban, rural). Because the nonresponse adjustments were done by the international consortium, the nonresponse bias analysis of the U.S. data was not used to inform the nonresponse weight adjustments. Thus, there was not an explicit nonresponse adjustment for this identified source of bias.

Test Development

The development of the PISA 2006 assessment instruments was an interactive process among the PISA Consortium, various expert committees, and OECD members. The assessment was developed by international experts and PISA Consortium test developers, and items were reviewed by representatives of each jurisdiction for possible bias and relevance to PISA's goals. The intention was to reflect the national, cultural, and linguistic variety among OECD jurisdictions. The assessment included items submitted by participating jurisdictions as well as items that were developed by the Consortium's test developers.

The final assessment consisted of 140 science items, 48 mathematics items, and 28 reading items allocated to 13 test booklets. Each booklet was made up of 4 test clusters. Altogether there were 7 science clusters (S1–S7), 4 mathematics clusters (M1–M4), and 2 reading clusters (R1–R2). The clusters were allocated in a rotated design to the 13 booklets. The average number of items per cluster was 20 items for science, 12 items for mathematics, and 14 items for reading. Each cluster was designed to average 30 minutes of test material. Each student took one booklet, with about 2 hours worth of testing material. Approximately one-third of the science literacy items were multiple choice, one-third were closed or short response types (for which students wrote an answer that was simply either correct or incorrect), and about one-third were open constructed responses (for which students wrote answers that were graded by trained scorers using an international scoring guide). In PISA 2006, every student answered science items. Mathematics and reading items were spread throughout other booklets. The United States did not use the optional 1-hour test booklet that included lower difficulty items designed for use in special education classrooms. This booklet was used by seven jurisdictions: Austria, Belgium, the

Czech Republic, Germany, the Netherlands, Slovakia, and Slovenia. For more information on assessment design, see the OECD's *PISA 2006 Technical Report* (Adams in press).

In addition to the cognitive assessment, students also received a 30-minute questionnaire designed to provide information about their backgrounds, attitudes, and experiences in school. Principals in schools where PISA was administered also received a 20- to 30-minute questionnaire about their schools. Results from the school survey are not discussed in this report but are available in *PISA 2006: Science Competencies for Tomorrow's World* (Vols. 1 and 2) (OECD, 2007a, 2007b).

Translation

Source versions of all instruments (assessment booklets, questionnaires, and manuals) were prepared in English and French and translated into the primary language or languages of instruction in each jurisdiction. PISA recommended that jurisdictions prepare and consolidate independent translations from both source versions and provided precise translation guidelines that included a description of the features each item was measuring and statistical analysis from the field trial. In cases for which one source language was used, independent translations were required and discrepancies reconciled. In addition, it was sometimes necessary to adapt the instrument for cultural purposes, even in nations such as the United States that use English as the primary language of instruction. For example, words such as "lift" might be adapted to "elevator" for the United States. The PISA Consortium verified the national translation and adaptation of all instruments. Electronic copies of printed materials were sent to the PISA Consortium for a final visual check prior to data collection.

Test Printing

An error was made in printing the final test booklets in the United States and the pagination of the booklets was consistently off by one page. The international consortium intended for the first page to be printed on the inside of the back cover; in the United States it was printed on the typical first page of plain white paper. As a result, some of the instructions in the reading section were incorrect. In some passages, students were incorrectly instructed to refer to the passage on the "opposite page" when the passage now appeared on the previous page. Because of the small number of items in the reading section, it was not possible to recalibrate the score to exclude the affected items. No incorrect page references appeared in the mathematics or science sections of the assessments. However, in some instances math and science items could be more difficult because the question required information provided previously that now required the student to turn back a page. In a few instances, items could be somewhat easier because of the pagination. ACER examined the potential impact of this on the math and science scales and estimated the scores would change by one point if the items that may have been affected by pagination were removed. Because one point is within the equating error of the scale, the original scales were retained using the results from all mathematics and science items.

Test Administration and Quality Assurance

PISA 2006 emphasized the use of standardized procedures in all jurisdictions. Each jurisdiction collected its own data, based on comprehensive manuals and training sessions provided by the PISA Consortium to explain the survey's implementation, including precise instructions for the work of school coordinators and scripts for test administrators to use in testing sessions. Test administration in the United States was carried out by professional staff trained according to the international guidelines. School staff were asked only to assist with listing students, identifying space for testing in the school, and specifying any parental consent procedures needed for sampled students. Students were allowed to use calculators, and U.S. students were provided calculators; however, no information on the availability of calculators was collected internationally.

At some schools, the PISA test was administered to students outside of normal school hours to address schools' concerns about the potential negative effect on students of the loss of instructional time. Tests were administered during normal school hours at 88 schools (53 percent), after normal school hours at 4 schools (2 percent), and on Saturday mornings at 74 schools (45 percent).

No differences were found between the schools that administered the test during out-of-school hours and the schools that opted for traditional in-school testing. Tests for differences by a variety of school characteristics (school control, locale, region, school size, school racial composition, and percentage of students receiving free or reduced-price lunch) demonstrated no significant results. Tests for differences in student test scores were implemented at both the school and student levels, and no measurable differences were found between the two groups of schools. Finally, a regression analysis of test scores as a function of selected school characteristics found no significant effect of the type of administration on the final test scores (Krotki and Bland 2008).

Members of the PISA Consortium visited all national centers to review data collection procedures, and members of the PISA Consortium also visited a randomly selected subsample of approximately 10 percent of the schools to ensure that procedures were being carried out in accordance with international guidelines. For a detailed description of the quality assurance procedures, see the OECD's *PISA 2006 Technical Report* (Adams in press).

Scoring

At least one-third of the PISA assessment was devoted to items requiring constructed responses. The process of scoring these items was an important step in ensuring the quality and comparability of the PISA data. Detailed guidelines were developed for the scoring guides themselves, training materials to recruit scorers, and workshop materials used for the training of national scorers. Prior to the national training, the PISA Consortium organized training sessions to present the material and train the scoring coordinators from the participating jurisdictions, who trained the national scorers.

For each test item, the scoring guide described the intent of the question and how to score the students' responses to each item. This description included the credit labels—full credit, partial credit, or no credit—attached to the possible categories of response. In addition, the scoring guides included real examples of students' responses accompanied by a rationale for their classification for purposes of clarity and illustration.

To examine the consistency of this marking process in more detail within each jurisdiction and to estimate the magnitude of the variance components associated with the use of scorers, the PISA Consortium conducted an interscorer reliability study on a subsample of assessment booklets. Homogeneity analysis was applied to the national sets of multiple scoring and compared with the results of the field trial. A full description of this process and the results can be found in the OECD's *PISA 2006 Technical Report* (Adams in press).

Data Entry and Cleaning

Data entry was the responsibility of the national project manager from each nation. The data collected for PISA 2006 were entered into data files with a common international format, as specified in the *PISA 2006 Main Study Management Manual, Version 3* (Australian Council for Educational Research [ACER] 2006). Data entry was completed using specialized software that allowed data to be merged into Keyquest, a common data processing software application developed by the ACER for use by participating nations. The software facilitated the checking and correction of data by providing various data consistency checks. The data were then sent to ACER for cleaning. ACER's role at this point was to check that the international data structure was followed, check the identification system within and between files, correct single case problems manually, and apply standard cleaning procedures to questionnaire files. Results of the data cleaning process were documented and shared with the national project managers and included specific questions when required. The national project manager then provided ACER with revisions to coding or solutions for anomalies. ACER then compiled background univariate statistics and preliminary classical and Rasch Item Analysis. Detailed information on the entire data entry and cleaning process can be found in the OECD's *PISA 2006 Technical Report* (Adams in press).

Weighting

The use of sampling weights is necessary for the computation of statistically sound, nationally representative estimates. Adjusted survey weights adjust for the probabilities of selection for individual schools and students, for school or student nonresponse, or for errors in estimating the size of the school or the number of 15-year-olds in the school at the time of sampling. Survey weighting for all jurisdictions participating in PISA 2006 was carried out by Westat, as part of the PISA Consortium.

The internationally defined weighting specifications for PISA 2006 included two base weights and five adjustments. The school base weight was defined as the reciprocal of the school's probability of selection. (For replacement schools, the school base weight was set equal to the original school it replaced.) The student base weight was given as the reciprocal of the probability of selection for each selected student from within a school.

The product of these base weights was then adjusted for school and student nonresponse. The school nonresponse adjustment was done individually for each jurisdiction using the implicit and explicit strata defined as part of the sample design. In the case of the United States, three variables were used: school control, census region, and community type. The student nonresponse adjustment was done within cells based first on their final school nonresponse cell and their explicit stratum, and within that, grade and gender were used as possible. Grade and gender were collected for students in all jurisdictions on the student tracking form. Trimming factors at the school and student levels were also used (one school weight was trimmed for the United States data; no student weights were trimmed). All PISA analyses were conducted using these adjusted sampling weights. For more information on the nonresponse adjustments and trimming factors, see the OECD's *PISA 2006 Technical Report* (Adams in press).

Scaling of Student Test Data

Thirteen versions of the PISA test booklet were created, each containing a slightly different subset of items. The fact that each student completed only a subset of items means that classic test scores, such as the percent correct, are not accurate measures of student performance. Instead, scaling techniques were used to establish a common scale for all students. For PISA 2006, item response theory (IRT) was used to estimate average scores for science, mathematics, and reading literacy for each jurisdiction.

IRT identifies patterns of response and uses statistical models to predict the probability of answering an item correctly as a function of the students' proficiency in answering other questions. PISA 2006 used a mixed coefficients multinomial logit IRT model. This model is similar in principle to the more familiar two-parameter IRT model. With this method, the performance of a sample of students in a subject area or sub-area can be summarized on a simple scale or series of scales, even when students are administered different items.

Scores for students are estimated as plausible values because each student completed only a subset of items. Five plausible values were estimated for each student for each scale. These values represent the distribution of potential scores for all students in the population with similar characteristics and identical patterns of item response. Statistics describing performance on the PISA science and mathematics scales are based on plausible values.[4]

Proficiency Levels

In addition to a range of scale scores as the basic form of measurement, PISA also describes student proficiency in science literacy in terms of six described levels. Increasing levels represent the knowledge, skills, and capabilities needed to perform tasks of increasing complexity. As a result, the findings are reported in terms of percentages of the student population at each of the predefined levels.

Each of the four science literacy scales—the combined scale and the three subscales—is divided into six levels. Descriptions were developed to characterize typical student performance at each level. A seventh level (below level 1) was established to include students whose abilities could not be accurately described based on their responses. Exhibit 1 in the body of the report summarizes the knowledge and skills that students need to demonstrate to be classified into one of the six levels on the combined science literacy scale. Similarly, exhibit B-1 in this appendix presents the

[4] For theoretical and empirical justification of the procedures employed, see Mislevy (1988).

Exhibit B-1. Description of general competencies and examples of tasks students should be able to do, by science literacy subscale and proficiency level: 2006

	Task descriptions		
Proficiency level	**Identifying scientific issues**	**Explaining phenomena scientifically**	**Using scientific evidence**
Level 1	Students at this level should be able to suggest appropriate sources of information on scientific topics. They should be able to identify a quantity that is undergoing variation in an experiment. In specific contexts they should be able to recognize whether that variable can be measured using familiar measuring tools or not.	Students at this level should be able to recognize simple cause-and-effect relationships given relevant cues. The knowledge drawn upon is a singular scientific fact that is drawn from experience or has widespread popular currency.	In response to a question, students at this level should be able to extract information from a fact sheet or diagram pertinent to a common context. They should be able to extract information from bar graphs where the requirement is simple comparisons of bar heights. In common, experienced contexts students at this level should be able to attribute an effect to a cause.
Level 2	Students at this level should be able to determine if scientific measurement can be applied to a given variable in an investigation. They should be able to recognize the variable being manipulated (changed) by the investigator. Students should be able to appreciate the relationship between a simple model and the phenomenon it is modeling. In researching topics students should be able to select appropriate key words for a search.	Students at this level should be able to recall an appropriate, tangible, scientific fact applicable in a simple and straightforward context and should be able to use it to explain or predict an outcome.	Students at this level should be able to recognize the general features of a graph if they are given appropriate cues and can point to an obvious feature in a graph or simple table in support of a given statement. They should be able to recognize if a set of given characteristics applies to the function of everyday artifacts in making choices about their use.
Level 3	Students at this level should be able to make judgments about whether an issue is open to scientific measurement and, consequently, to scientific investigation. Given a description of an investigation, they should be able to identify the change and measured variables.	Students at this level should be able to apply one or more concrete or tangible scientific ideas/concepts in the development of an explanation of a phenomenon. This is enhanced when there are specific cues given or options available from which to choose. When developing an explanation, cause-and-effect relationships are recognized and simple, explicit scientific models may be drawn upon.	Students at this level should be able to select a piece of relevant information from data in answering a question or in providing support for or against a given conclusion. They should be able to draw a conclusion from an uncomplicated or simple pattern in a dataset. Students should be able to also determine, in simple cases, if enough information is present to support a given conclusion.
Level 4	Students at this level should be able to identify the change and measured variables in an investigation and at least one variable that is being controlled. They should be able to suggest appropriate ways of controlling that variable. The question being investigated in straightforward investigations can be articulated.	Students at this level should have an understanding of scientific ideas, including scientific models, with a significant level of abstraction. They should be able to apply a general, scientific concept containing such ideas in the development of an explanation of a phenomenon.	Students at this level should be able to interpret a dataset expressed in a number of formats, such as tabular, graphic, and diagrammatic, by summarizing the data and explaining relevant patterns. They should be able to use the data to draw relevant conclusions. Students should also be able to determine whether the data support assertions about a phenomenon.
Level 5	Students at this level understand the essential elements of a scientific investigation and thus should be able to determine if scientific methods can be applied in a variety of quite complex, and often abstract contexts. Alternatively, by analyzing a given experiment they should be able to identify the question being investigated and explain how the methodology relates to that question.	Students at this level should be able to draw on knowledge of two or three scientific concepts and identify the relationship between them in developing an explanation of a contextual phenomenon.	Students at this level should be able to interpret data from related datasets presented in various formats. They should be able to identify and explain differences and similarities in the datasets and draw conclusions based on the combined evidence presented in those datasets.
Level 6	Students at this level should demonstrate an ability to understand and articulate the complex modeling inherent in the design of an investigation.	Students at this level should be able to draw on a range of abstract scientific knowledge and concepts and the relationships between these in developing explanations of processes within systems.	Students at this level should demonstrate an ability to compare and differentiate among competing explanations by examining supporting evidence. They should be able to formulate arguments by synthesizing evidence from multiple sources.

NOTE: To reach a particular proficiency level, a student must correctly answer a majority of items at that level. Students were classified into science literacy levels according to their scores. Exact cut point scores are as follows: below level 1 (a score less than or equal to 334.94); level 1 (a score greater than 334.94 and less than or equal to 409.54); level 2 (a score greater than 409.54 and less than or equal to 484.14); level 3 (a score greater than 484.14 and less than or equal to 558.73); level 4 (a score greater than 558.73 and less than or equal to 633.33); level 5 (a score greater than 633.33 and less than or equal to 707.93); and level 6 (a score greater than 707.93).
SOURCE: Organization for Economic Cooperation and Development (OECD). (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006.* Paris: Author; Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

proficiency descriptions for each of the six levels on the science subscales. Exact cut scores for the levels are as follows: below level 1 (a score less than or equal to 334.94); level 1 (a score greater than 334.94 and less than or equal to 409.54); level 2 (a score greater than 409.54 and less than or equal to 484.14); level 3 (a score greater than 484.14 and less than or equal to 558.73); level 4 (a score greater than 558.73 and less than or equal to 633.33); level 5 (a score greater than 633.33 and less than or equal to 707.93); and level 6 (a score greater than 707.93).

To determine the performance levels and cut scores on the literacy scales, IRT techniques were used. With IRT techniques, it is possible to simultaneously estimate the ability of all students taking the PISA assessment, as well as the difficulty of all PISA items. Then estimates of student ability and item difficulty can be mapped on a single continuum. The relative ability of students taking a particular test can be estimated by considering the percentage of test items they get correct. The relative difficulty of items in a test can be estimated by considering the percentage of students getting each item correct. In PISA, all students within a level are expected to answer at least half of the items from that level correctly. Students at the bottom of a level are able to provide the correct answers to about 52 percent of all items from that level, have a 62 percent chance of success on the easiest items from that level, and have a 42 percent chance of success on the hardest items from that level. Students in the middle of a level have a 62 percent chance of correctly answering items of average difficulty for that level (an overall response probability of 62 percent). Students at the top of a level are able to provide the correct answers to about 70 percent of all items from that level, have a 78 percent chance of success on the easiest items from that level, and have a 62 percent chance of success on the hardest items from that level. Students just below the top of a level would score less than 50 percent on an assessment at the next higher level. Students at a particular level demonstrate not only the knowledge and skills associated with that level but also the proficiencies defined by lower levels. Thus, all students proficient at level 3 are also proficient at levels 1 and 2. Patterns of responses for students below level 1 suggest that these students are unable to answer at least half of the items from level 1 correctly. For details about the approach to defining and describing the PISA levels and establishing the cut scores, see the OECD's *PISA 2006 Technical Report* (Adams in press) and the *PISA 2003 Technical Report* (Adams 2004).

Data Limitations

As with any study, there are limitations to PISA 2006 that researchers should take into consideration. Estimates produced using data from PISA 2006 are subject to two types of error: nonsampling and sampling errors. Nonsampling errors can be due to errors made in the collection and processing of data. Sampling errors can occur because the data were collected from a sample rather than a complete census of the population.

Nonsampling Errors

"Nonsampling error" is a term used to describe variations in the estimates that may be caused by population coverage limitations, nonresponse bias, and measurement error, as well as data collection, processing, and reporting procedures. For example, the sampling frame was limited to regular public and private schools in the 50 states and the District of Columbia and cannot be used to represent Puerto Rico or other jurisdictions. The sources of nonsampling errors are typically problems such as unit and item nonresponse, the differences in respondents' interpretations of the meaning of survey questions, response differences related to the particular time the survey was conducted, and mistakes in data preparation. Some of these issues (particularly unit nonresponse) are discussed above in the section on U.S. sampling and data collection. Another example of nonsampling error that affected this data collection was the printing error, described earlier in the Test Printing section.

Sampling Errors

Sampling errors occur when a discrepancy between a population characteristic and the sample estimate arises because not all members of the target population are sampled for the survey. The size of the sample relative to the population and the variability of the population characteristics both influence the magnitude of sampling error. The particular sample of 15-year-old students from fall 2006 was just one of many possible samples that could have been selected. Therefore, estimates produced from the PISA 2006 sample may differ from estimates that would have been produced had another sample of students been selected. This type of variability is called sampling error because it arises from using a sample of 15-year-old students in 2006 rather than all 15-year-old students in that year.

One potential source of sampling error for PISA 2006 is that the weight for a replacement school was based on the weight for the school originally selected. These schools were typically very similar in size and other characteristics (the replacement schools were adjacent to the original school on the sorted list of schools), however, there could be some error associated with this method. A second potential source of sampling error could occur if the enrollment lists used for sampling were not up to date.

The standard error is a measure of the variability owing to sampling when estimating a statistic. The approach used for calculating sampling variances in PISA was the Fay method of Balanced Repeated Replication (BRR). This method of producing standard errors uses information about the sample design to produce more accurate standard errors than would be produced using simple random sample assumptions. Thus, the standard errors that are reported here can be used as a measure of the precision expected from this particular sample.

Standard errors for all of the estimates are in appendix C of this report. These standard errors can be used to produce confidence intervals. In keeping with NCES standards, 95 percent confidence intervals are used for this report. A 95 percent confidence interval is interpreted as a 95 percent chance that the true average in the population lies within the range of 1.96 times the standard error above or below the estimated score.

Missing Data

There are four kinds of missing data at the item level. "Nonresponse" data occurs when a respondent is expected to answer an item but no response is given. Responses that are "missing or invalid" occur in multiple-choice items for which an invalid response is given. The missing or invalid code is not used for open-ended questions. An item is "not applicable" when it is not possible for the respondent to answer the question. Finally, items that are "not reached" are consecutive missing values starting from the end of each test session. All four kinds of missing data are coded differently in the PISA 2006 database.

Background data were not imputed for cases with missing data, and those cases were not included in instances where they had missing data. Item response rates for variables discussed in this report were all over 85 percent. Response rates for sex were 100 percent in all participating jurisdictions and the response rate for race/ethnicity in the United States was 98 percent.

Descriptions of Background Variables

In this report, PISA 2006 results are provided for groups of students with different demographic characteristics. Definitions of subpopulations are as follows:

Sex: Results are reported separately for male students and female students.

Race/ethnicity: In the United States, students' race/ethnicity was obtained through student responses to a two-part question in the student questionnaire. Students were asked first whether they were Hispanic or Latino and then whether they were members of the following racial groups: White (non-Hispanic), Black (non-Hispanic), Asian (non-Hispanic), American Indian or Alaska Native (non-Hispanic), or Native Hawaiian/Other Pacific Islander (non-Hispanic). Multiple responses to the race classification question were allowed. Results are shown separately for White (non-Hispanic) students, Black (non-Hispanic) students, Hispanic students, Asian (non-Hispanic) students, American Indian or Alaska Native (non-Hispanic) students, Native Hawaiian/Other Pacific Islander (non-Hispanic) students, and non-Hispanic students who selected more than one race. Students identifying themselves as Hispanic and one or more race were included in the Hispanic group, rather than in a racial group.

Full PISA 2006 student and school questionnaires are available at http://nces.ed.gov/surveys/pisa and http://www.pisa.oecd.org.

Confidentiality and Disclosure Limitations

The PISA 2006 data are hierarchical and include school and student data from the participating schools. Confidentiality analyses for the United States were designed to provide reasonable assurance that public-use data files issued by the PISA Consortium would not allow identification of individual U.S. schools or students when compared against other public-use data collections. Disclosure limitation included identifying and masking potential disclosure risk to PISA schools and including an additional measure of uncertainty to school and student identification through random swapping of data elements within the student and school files.

Statistical Procedures

Tests of Significance

Comparisons made in the text of this report have been tested for statistical significance. For example, in the commonly made comparison of jurisdiction averages against the average of the United States, tests of statistical significance were used to establish whether or not the observed differences from the U.S. average were statistically significant.

The estimation of the standard errors that are required in order to undertake the tests of significance is complicated by the complex sample and assessment designs, both of which generate error variance. Together they mandate a set of statistically complex procedures for estimating the correct standard errors. As a consequence, the estimated standard errors contain a sampling variance component estimated by BRR. Where the assessments are concerned, there is an additional imputation variance component arising from the assessment design. Details on the BRR procedures used can be found in the *PISA 2006 Technical Report* (Adams in press) and the *PISA 2003 Technical Report* (Adams 2004).

In almost all instances, the tests for significance used were standard *t* tests. These fell into two categories according to the nature of the comparison being made: comparisons of independent samples and comparisons of nonindependent samples. In PISA, jurisdiction samples are independent.

In simple comparisons of independent averages, such as the average score of jurisdiction 1 with that of jurisdiction 2, the following formula was used to compute the *t* statistic:

$$t = est_1 - est_2 \,/\, \text{SQRT}\,[(se_1)^2 + (se_2)^2],$$

where est_1 and est_2 are the estimates being compared (e.g., averages of jurisdiction 1 and jurisdiction 2) and se_1 and se_2 are the corresponding standard errors of these averages.

The second type of comparison used in this report occurred when comparing differences of nonsubset, nonindependent groups. When this occurs, the correlation and related covariance between the groups must be taken into account, such as when comparing the average scores of males versus females within the United States.

How are scores such as those for males and females correlated? Suppose that in the school sample, a coeducational school attended by low achievers is replaced by a coeducational school attended by high achievers. The jurisdiction mean will increase slightly, as well as the means for males and females. If such a school replacement process is continued, the average scores of males and the average scores of females will likely increase in a similar pattern. Indeed, a coeducational school attended by high-achieving males is usually also attended by high-achieving females. Therefore, the covariance between the males' scores and the females' scores is likely to be positive.

To determine whether the performance of females differs from the performance of males, the standard error of the difference that takes into account the covariance between the females' scores and the males' scores needs to be estimated. The estimation of the covariance requires the selection of several samples and then the analysis of the variation of the males' means in conjunction with the females' means. Such a procedure is, of course, unrealistic. Therefore, as for any computation of a standard error in PISA, replication methods using the supplied replicate weights were used to estimate the standard error of a difference. Use of the replicate weights implicitly incorporates the covariance between the two estimates into the estimate of the standard error of the difference.

To test such comparisons, the following formula was used to compute the *t* statistic:

$$t = est_{grp1} - est_{grp2} / se\ (est_{grp1} - est_{grp2}),$$

where est_{grp1} and est_{grp2} are the nonindependent group estimates being compared and $se\ (est_{grp1} - est_{grp2})$ is the standard error of the difference calculated using BRR to account for any covariance between the estimates for the two nonindependent groups.

Effect Size

Tests of statistical significance are, in part, influenced by sample sizes. To provide the reader with an increased understanding of the importance of the significant difference between student populations in the United States, effect sizes are included in the report. Effect sizes use standard deviations, rather than standard errors, and are therefore not influenced by the size of the student samples. Following Cohen (1988) and Rosnow and Rosenthal (1996), effect size is calculated by finding the difference between the means of two groups and dividing that result by the pooled standard deviation of the two groups:

$$d = \frac{est_{grp1} - est_{grp2}}{sd_{pooled}},$$

where est_{grp1} and est_{grp2} are the student group estimates being compared and sd_{pooled} is the pooled standard deviation of the groups being compared. The formula for the pooled standard deviation is as follows (Rosnow and Rosenthal 1996):

$$sd_{pooled} = \sqrt{\frac{sd_1^2 + sd_2^2}{2}},$$

where sd_1 and sd_2 are the standard deviations of the groups being compared.

Appendix C: Reference Tables

Table C-1. Percentage distribution of 15-year-old students, by grade level and jurisdiction: 2006

	7th		8th		9th		10th		11th		12th		Not reported/ unknown	
Jurisdiction	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*
OECD average[1]	**0.9**	***0.06***	**5.9**	***0.11***	**37.5**	***0.18***	**48.8**	***0.19***	**11.5**	***0.08***	**0.8**	***0.04***	**1.7**	**—**
OECD jurisdictions														
Australia	#	*†*	0.1	*0.03*	8.9	*0.59*	70.6	*0.93*	20.3	*0.77*	0.1	*0.05*	#	*†*
Austria	0.3	*0.11*	6.4	*0.70*	44.6	*1.13*	48.7	*1.06*	#	*†*	#	*†*	#	*†*
Belgium	0.4	*0.09*	4.4	*0.32*	31.1	*0.79*	63.2	*0.79*	1.0	*0.11*	#	*†*	#	*†*
Canada	#	*†*	1.7	*0.22*	13.3	*0.59*	83.8	*0.65*	1.2	*0.14*	#	*†*	#	*†*
Czech Republic	0.7	*0.14*	3.5	*0.39*	44.3	*1.32*	51.5	*1.49*	#	*†*	#	*†*	#	*†*
Denmark	0.2	*0.07*	12.0	*0.57*	85.3	*0.73*	1.4	*0.24*	1.1	*0.36*	#	*†*	#	*†*
Finland	0.2	*0.06*	11.7	*0.52*	88.1	*0.52*	#	*†*	#	*†*	#	*†*	#	*†*
France	#	*†*	5.2	*0.45*	34.8	*1.19*	57.5	*1.21*	2.4	*0.26*	#	*†*	#	*†*
Germany	1.5	*0.25*	11.9	*0.58*	54.5	*0.65*	28.2	*0.82*	0.3	*0.06*	#	*†*	3.6	*0.32*
Greece	0.5	*0.14*	2.1	*0.37*	5.3	*0.76*	78.7	*1.02*	13.3	*0.56*	#	*†*	#	*†*
Hungary	2.2	*0.45*	5.5	*0.62*	65.7	*0.85*	26.6	*0.54*	#	*†*	#	*†*	#	*†*
Iceland	#	*†*	#	*†*	0.2	*0.08*	99.2	*0.11*	0.6	*0.08*	#	*†*	#	*†*
Ireland	#	*†*	2.7	*0.38*	58.5	*0.79*	21.2	*1.27*	17.5	*1.08*	#	*†*	#	*†*
Italy	0.3	*0.10*	1.5	*0.43*	15.0	*0.58*	80.4	*0.69*	2.8	*0.24*	#	*†*	#	*†*
Japan	#	*†*	#	*†*	#	*†*	100.0	*0.00*	#	*†*	#	*†*	#	*†*
Korea, Republic of	#	*†*	#	*†*	2.0	*0.57*	97.3	*0.58*	0.7	*0.11*	#	*†*	#	*†*
Luxembourg	0.2	*0.07*	11.8	*0.26*	53.4	*0.43*	34.4	*0.41*	0.1	*0.04*	#	*†*	#	*†*
Mexico	2.3	*0.23*	8.1	*0.77*	33.2	*1.92*	48.5	*1.90*	5.1	*0.36*	2.0	*0.16*	0.9	*0.29*
Netherlands	0.1	*0.09*	3.7	*0.39*	44.9	*1.09*	50.7	*1.17*	0.4	*0.10*	#	*†*	#	*†*
New Zealand	#	*†*	#	*†*	#	*†*	6.2	*0.36*	89.4	*0.46*	4.4	*0.31*	#	*†*
Norway	#	*†*	#	*†*	0.5	*0.11*	99.0	*0.33*	0.5	*0.31*	#	*†*	#	*†*
Poland	0.6	*0.16*	3.8	*0.34*	95.1	*0.41*	0.6	*0.08*	#	*†*	#	*†*	#	*†*
Portugal	6.4	*0.67*	12.8	*0.68*	28.9	*1.12*	49.6	*1.53*	0.2	*0.05*	#	*†*	2.1	*0.63*
Slovak Republic	0.7	*0.21*	2.2	*0.42*	38.5	*2.09*	58.7	*2.20*	#	*†*	#	*†*	#	*†*
Spain	0.1	*0.04*	7.0	*0.47*	33.0	*0.79*	59.8	*0.87*	#	*†*	#	*†*	#	*†*
Sweden	#	*†*	1.9	*0.21*	95.9	*0.38*	2.2	*0.32*	#	*†*	#	*†*	#	*†*
Switzerland	0.8	*0.12*	16.1	*0.78*	62.6	*1.46*	20.3	*1.65*	0.3	*0.13*	#	*†*	#	*†*
Turkey	0.8	*0.32*	4.5	*0.90*	38.4	*1.73*	53.7	*1.88*	2.6	*0.27*	#	*†*	#	*†*
United Kingdom	#	*†*	#	*†*	#	*†*	0.9	*0.10*	98.4	*0.13*	0.7	*0.06*	#	*†*
United States	0.8	*0.74*	1.0	*0.87*	10.7	*0.78*	70.9	*1.42*	16.5	*0.75*	0.1	*0.05*	#	*†*
Non-OECD jurisdictions														
Argentina	3.9	*0.83*	9.4	*0.76*	17.0	*1.35*	64.4	*2.11*	3.0	*0.40*	0.6	*0.55*	1.7	*0.98*
Azerbaijan	0.5	*0.11*	5.5	*0.55*	53.5	*1.48*	39.0	*1.54*	0.6	*0.13*	0.5	*0.40*	0.5	*0.21*
Bulgaria	0.3	*0.14*	7.1	*0.96*	74.3	*1.17*	18.2	*0.90*	#	*†*	#	*†*	#	*†*
Brazil	11.6	*0.69*	22.0	*1.25*	47.8	*1.24*	18.0	*0.86*	0.6	*0.18*	#	*†*	#	*†*
Chile	1.0	*0.31*	3.3	*0.52*	18.9	*0.99*	70.8	*1.19*	6.1	*0.46*	#	*†*	#	*†*
Chinese Taipei	#	*†*	#	*†*	36.3	*1.30*	63.6	*1.32*	0.1	*0.08*	#	*†*	#	*†*
Colombia	6.4	*0.96*	12.3	*0.91*	22.2	*0.83*	37.8	*1.39*	21.4	*2.14*	#	*†*	#	*†*
Croatia	#	*†*	0.4	*0.26*	77.1	*0.48*	22.6	*0.43*	#	*†*	#	*†*	#	*†*
Estonia	3.3	*0.37*	25.6	*0.84*	69.4	*0.86*	1.8	*0.17*	#	*†*	#	*†*	#	*†*
Hong Kong-China	2.4	*0.22*	9.3	*0.54*	25.2	*0.46*	63.0	*0.93*	0.1	*0.15*	#	*†*	#	*†*
Indonesia	0.1	*0.05*	12.0	*1.68*	40.0	*2.97*	43.5	*3.76*	4.4	*0.63*	#	*†*	#	*†*
Israel	#	*†*	0.3	*0.07*	14.6	*1.05*	84.7	*1.07*	0.4	*0.10*	#	*†*	#	*†*
Jordan	0.1	*0.08*	1.3	*0.18*	8.1	*0.58*	90.5	*0.73*	#	*†*	#	*†*	#	*†*
Kyrgyz Republic	0.2	*0.10*	7.7	*0.59*	67.6	*1.22*	24.2	*1.35*	0.4	*0.13*	#	*†*	#	*†*
Latvia	2.6	*0.64*	16.4	*0.78*	77.7	*1.14*	3.0	*0.40*	#	*†*	#	*†*	0.4	*0.18*
Liechtenstein	#	*†*	16.7	*0.63*	72.0	*0.57*	11.0	*0.55*	0.3	*0.30*	#	*†*	#	*†*
Lithuania	0.9	*0.15*	12.1	*0.81*	80.0	*0.87*	6.8	*0.48*	#	*†*	#	*†*	0.2	*0.16*
Macao-China	7.7	*0.16*	20.6	*0.21*	34.7	*0.18*	36.5	*0.13*	0.6	*0.04*	#	*†*	#	*†*
Qatar	2.3	*0.10*	5.3	*0.13*	14.1	*0.13*	62.6	*0.17*	15.6	*0.15*	0.2	*0.06*	#	*†*
Republic of Montenegro	#	*†*	0.3	*0.15*	85.8	*0.22*	13.9	*0.17*	#	*†*	#	*†*	#	*†*
Republic of Serbia	0.1	*0.06*	1.8	*0.57*	96.6	*0.61*	1.6	*0.19*	#	*†*	#	*†*	#	*†*
Romania	0.7	*0.36*	13.5	*2.02*	82.9	*1.91*	2.9	*0.39*	#	*†*	#	*†*	#	*†*
Russian Federation	0.6	*0.14*	6.7	*0.89*	29.9	*1.58*	61.6	*2.00*	1.2	*0.22*	#	*†*	#	*†*
Slovenia	#	*†*	0.2	*0.11*	3.5	*0.33*	90.6	*0.35*	5.8	*0.21*	#	*†*	#	*†*
Thailand	#	*†*	1.3	*0.35*	30.5	*1.05*	65.2	*1.13*	3.0	*0.47*	#	*†*	#	*†*
Tunisia	11.4	*0.56*	16.7	*0.75*	21.1	*1.00*	46.6	*1.59*	4.3	*0.32*	#	*†*	#	*†*
Uruguay	7.5	*0.90*	9.8	*0.70*	17.3	*1.02*	58.9	*1.51*	6.6	*0.63*	#	*†*	#	*†*

— Not available.
† Not applicable.
# Rounds to zero.
[1] In computing the OECD average, the average for each column (grade in this case) is computed by averaging the estimates in the column but excluding those instances where no cases were reported (shown here as '#': rounds to zero). Therefore, the percentage distribution sums to greater than 100 (i.e., 107.1).
NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Standard error is noted by *s.e.* Detail may not sum to totals because of rounding.
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-2. Average scores of 15-year-old students on combined science literacy scale and science literacy subscales, by jurisdiction: 2006

	Combined science literacy scale		Science literacy subscales: Identifying scientific issues		Science literacy subscales: Explaining phenomena scientifically		Science literacy subscales: Using scientific evidence	
Jurisdiction	**Average**	***s.e.***	**Average**	***s.e.***	**Average**	***s.e.***	**Average**	***s.e.***
OECD average	**500**	***0.5***	**499**	***0.5***	**500**	***0.5***	**499**	***0.6***
OECD jurisdictions								
Australia	527	*2.3*	535	*2.3*	520	*2.3*	531	*2.4*
Austria	511	*3.9*	505	*3.7*	516	*4.0*	505	*4.7*
Belgium	510	*2.5*	515	*2.7*	503	*2.5*	516	*3.0*
Canada	534	*2.0*	532	*2.3*	531	*2.1*	542	*2.2*
Czech Republic	513	*3.5*	500	*4.2*	527	*3.5*	501	*4.1*
Denmark	496	*3.1*	493	*3.0*	501	*3.3*	489	*3.6*
Finland	563	*2.0*	555	*2.3*	566	*2.0*	567	*2.3*
France	495	*3.4*	499	*3.5*	481	*3.2*	511	*3.9*
Germany	516	*3.8*	510	*3.8*	519	*3.7*	515	*4.6*
Greece	473	*3.2*	469	*3.0*	476	*3.0*	465	*4.0*
Hungary	504	*2.7*	483	*2.6*	518	*2.6*	497	*3.4*
Iceland	491	*1.6*	494	*1.7*	488	*1.5*	491	*1.7*
Ireland	508	*3.2*	516	*3.3*	505	*3.2*	506	*3.4*
Italy	475	*2.0*	474	*2.2*	480	*2.0*	467	*2.3*
Japan	531	*3.4*	522	*4.0*	527	*3.1*	544	*4.2*
Korea, Republic of	522	*3.4*	519	*3.7*	512	*3.3*	538	*3.7*
Luxembourg	486	*1.1*	483	*1.1*	483	*1.1*	492	*1.1*
Mexico	410	*2.7*	421	*2.6*	406	*2.7*	402	*3.1*
Netherlands	525	*2.7*	533	*3.3*	522	*2.7*	526	*3.3*
New Zealand	530	*2.7*	536	*2.9*	522	*2.8*	537	*3.3*
Norway	487	*3.1*	489	*3.1*	495	*3.0*	473	*3.6*
Poland	498	*2.3*	483	*2.5*	506	*2.5*	494	*2.7*
Portugal	474	*3.0*	486	*3.1*	469	*2.9*	472	*3.6*
Slovak Republic	488	*2.6*	475	*3.2*	501	*2.7*	478	*3.3*
Spain	488	*2.6*	489	*2.4*	490	*2.4*	485	*3.0*
Sweden	503	*2.4*	499	*2.6*	510	*2.9*	496	*2.6*
Switzerland	512	*3.2*	515	*3.0*	508	*3.3*	519	*3.4*
Turkey	424	*3.8*	427	*3.4*	423	*4.1*	417	*4.3*
United Kingdom	515	*2.3*	514	*2.3*	517	*2.3*	514	*2.5*
United States	489	*4.2*	492	*3.8*	486	*4.3*	489	*5.0*
Non-OECD jurisdictions								
Argentina	391	*6.1*	395	*5.7*	386	*6.0*	385	*7.0*
Azerbaijan	382	*2.8*	353	*3.1*	412	*3.0*	344	*4.0*
Brazil	390	*2.8*	398	*2.8*	390	*2.7*	378	*3.6*
Bulgaria	434	*6.1*	427	*6.3*	444	*5.8*	417	*7.5*
Chile	438	*4.3*	444	*4.1*	432	*4.1*	440	*5.1*
Chinese Taipei	532	*3.6*	509	*3.7*	545	*3.7*	532	*3.7*
Colombia	388	*3.4*	402	*3.4*	379	*3.4*	383	*3.9*
Croatia	493	*2.4*	494	*2.6*	492	*2.5*	490	*3.0*
Estonia	531	*2.5*	516	*2.6*	541	*2.6*	531	*2.7*
Hong Kong-China	542	*2.5*	528	*3.2*	549	*2.5*	542	*2.7*
Indonesia	393	*5.7*	393	*5.6*	395	*5.1*	386	*7.3*
Israel	454	*3.7*	457	*3.9*	443	*3.6*	460	*4.7*
Jordan	422	*2.8*	409	*2.8*	438	*3.1*	405	*3.3*
Kyrgyz Republic	322	*2.9*	321	*3.2*	334	*3.1*	288	*3.8*
Latvia	490	*3.0*	489	*3.3*	486	*2.9*	491	*3.4*
Liechtenstein	522	*4.1*	522	*3.7*	516	*4.1*	535	*4.3*
Lithuania	488	*2.8*	476	*2.7*	494	*3.0*	487	*3.1*
Macao-China	511	*1.1*	490	*1.2*	520	*1.2*	512	*1.2*
Qatar	349	*0.9*	352	*0.8*	356	*1.0*	324	*1.2*
Republic of Montenegro	412	*1.1*	401	*1.2*	417	*1.1*	407	*1.3*
Republic of Serbia	436	*3.0*	431	*3.0*	441	*3.1*	425	*3.7*
Romania	418	*4.2*	409	*3.6*	426	*4.0*	407	*6.0*
Russian Federation	479	*3.7*	463	*4.2*	483	*3.4*	481	*4.2*
Slovenia	519	*1.1*	517	*1.4*	523	*1.5*	516	*1.3*
Thailand	421	*2.1*	413	*2.5*	420	*2.1*	423	*2.6*
Tunisia	386	*3.0*	384	*3.8*	383	*2.9*	382	*3.7*
Uruguay	428	*2.7*	429	*3.0*	423	*2.9*	429	*3.1*

NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-3. Scores of 15-year-old students on combined science literacy scale at 10th and 90th percentiles, by jurisdiction: 2006

	Percentiles			
	10th		90th	
Jurisdiction	Score	*s.e.*	Score	*s.e.*
OECD average	**375**	***0.9***	**622**	***0.7***
OECD jurisdictions				
Australia	395	*3.4*	653	*2.9*
Austria	378	*6.2*	633	*3.6*
Belgium	374	*5.4*	634	*2.3*
Canada	410	*3.7*	651	*2.4*
Czech Republic	385	*5.2*	641	*4.3*
Denmark	373	*4.8*	615	*3.7*
Finland	453	*3.3*	673	*2.9*
France	359	*5.5*	623	*4.0*
Germany	381	*7.0*	642	*3.2*
Greece	353	*5.4*	589	*4.1*
Hungary	388	*4.2*	617	*3.1*
Iceland	364	*3.1*	614	*2.9*
Ireland	385	*4.4*	630	*3.7*
Italy	351	*2.8*	598	*2.6*
Japan	396	*6.2*	654	*3.1*
Korea, Republic of	403	*5.7*	635	*4.7*
Luxembourg	358	*2.8*	609	*2.8*
Mexico	306	*4.2*	516	*3.0*
Netherlands	395	*5.4*	646	*3.4*
New Zealand	389	*4.5*	667	*3.3*
Norway	365	*5.6*	610	*3.5*
Poland	381	*2.9*	615	*3.3*
Portugal	357	*4.8*	588	*2.9*
Slovak Republic	368	*3.7*	609	*4.1*
Spain	370	*3.7*	604	*3.0*
Sweden	381	*4.0*	622	*2.6*
Switzerland	378	*4.9*	636	*3.8*
Turkey	325	*3.2*	540	*9.7*
United Kingdom	376	*4.3*	652	*2.9*
United States	349	*5.9*	628	*4.3*
Non-OECD jurisdictions				
Argentina	259	*9.0*	520	*6.5*
Azerbaijan	316	*2.4*	456	*6.4*
Brazil	281	*3.2*	510	*5.6*
Bulgaria	300	*7.1*	577	*8.2*
Chile	323	*4.1*	560	*6.5*
Chinese Taipei	402	*5.0*	651	*2.7*
Colombia	280	*4.5*	496	*4.6*
Croatia	383	*3.8*	604	*3.2*
Estonia	422	*3.8*	640	*3.3*
Hong Kong-China	418	*6.1*	655	*3.5*
Indonesia	307	*3.5*	488	*11.8*
Israel	310	*5.2*	601	*4.5*
Jordan	309	*4.0*	537	*4.5*
Kyrgyz Republic	220	*3.8*	428	*5.0*
Latvia	380	*4.2*	597	*3.5*
Liechtenstein	393	*12.8*	643	*9.4*
Lithuania	370	*3.2*	604	*4.2*
Macao-China	409	*2.5*	611	*1.8*
Qatar	253	*1.4*	462	*2.6*
Republic of Montenegro	312	*2.1*	517	*3.0*
Republic of Serbia	327	*4.0*	545	*3.8*
Romania	314	*5.0*	526	*5.7*
Russian Federation	364	*5.4*	596	*3.9*
Slovenia	391	*2.8*	647	*3.3*
Thailand	325	*3.4*	524	*3.8*
Tunisia	283	*3.4*	495	*6.0*
Uruguay	306	*4.9*	550	*3.6*

NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-4. Standard deviations of the average scores of 15-year-old students on combined science literacy scale, by jurisdiction: 2006

Jurisdiction	Standard deviation	*s.e.*
OECD average	**95**	***0.3***
OECD jurisdictions		
Australia	100	*1.0*
Austria	98	*2.4*
Belgium	100	*2.0*
Canada	94	*1.1*
Czech Republic	98	*2.0*
Denmark	93	*1.4*
Finland	86	*1.0*
France	102	*2.1*
Germany	100	*2.0*
Greece	92	*2.0*
Hungary	88	*1.6*
Iceland	97	*1.2*
Ireland	94	*1.5*
Italy	96	*1.3*
Japan	100	*2.0*
Korea, Republic of	90	*2.4*
Luxembourg	97	*0.9*
Mexico	81	*1.5*
Netherlands	96	*1.6*
New Zealand	107	*1.4*
Norway	96	*2.0*
Poland	90	*1.1*
Portugal	89	*1.7*
Slovak Republic	93	*1.8*
Spain	91	*1.0*
Sweden	94	*1.4*
Switzerland	99	*1.7*
Turkey	83	*3.2*
United Kingdom	107	*1.5*
United States	106	*1.7*
Non-OECD jurisdictions		
Argentina	101	*2.6*
Azerbaijan	56	*1.9*
Brazil	89	*1.9*
Bulgaria	107	*3.2*
Chile	92	*1.8*
Chinese Taipei	94	*1.6*
Colombia	85	*1.8*
Croatia	86	*1.4*
Estonia	84	*1.1*
Hong Kong-China	92	*1.9*
Indonesia	70	*3.3*
Israel	111	*2.0*
Jordan	90	*1.9*
Kyrgyz Republic	84	*2.0*
Latvia	84	*1.3*
Liechtenstein	97	*3.1*
Lithuania	90	*1.6*
Macao-China	78	*0.8*
Qatar	84	*0.8*
Republic of Montenegro	80	*0.9*
Republic of Serbia	85	*1.6*
Romania	81	*2.4*
Russian Federation	90	*1.4*
Slovenia	98	*1.0*
Thailand	77	*1.5*
Tunisia	82	*2.0*
Uruguay	94	*1.8*

NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-5. Percentage distribution of 15-year-old students on combined science literacy scale, by proficiency level and jurisdiction: 2006

Jurisdiction	Below Level 1		Level 1		Level 2		Level 3		Level 4		Level 5		Level 6	
	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*
OECD average	**5.2**	***0.11***	**14.1**	**0.15**	**24.0**	**0.17**	**27.4**	**0.17**	**20.3**	**0.16**	**7.7**	**0.10**	**1.3**	**0.04**
OECD jurisdictions														
Australia	3.0	*0.25*	9.8	*0.46*	20.2	*0.63*	27.7	*0.51*	24.6	*0.53*	11.8	*0.53*	2.8	*0.26*
Austria	4.3	*0.88*	12.0	*0.98*	21.8	*1.05*	28.3	*1.05*	23.6	*1.12*	8.8	*0.69*	1.2	*0.20*
Belgium	4.8	*0.72*	12.2	*0.62*	20.8	*0.84*	27.6	*0.84*	24.5	*0.77*	9.1	*0.47*	1.0	*0.17*
Canada	2.2	*0.27*	7.8	*0.47*	19.1	*0.64*	28.8	*0.58*	27.7	*0.65*	12.0	*0.52*	2.4	*0.25*
Czech Republic	3.5	*0.57*	12.1	*0.84*	23.4	*1.17*	27.8	*1.09*	21.7	*0.92*	9.8	*0.86*	1.8	*0.32*
Denmark	4.3	*0.64*	14.1	*0.75*	26.0	*1.07*	29.3	*1.04*	19.5	*0.91*	6.1	*0.66*	0.7	*0.18*
Finland	0.5	*0.13*	3.6	*0.45*	13.6	*0.68*	29.1	*1.07*	32.2	*0.89*	17.0	*0.72*	3.9	*0.35*
France	6.6	*0.71*	14.5	*1.05*	22.8	*1.12*	27.2	*1.09*	20.9	*1.00*	7.2	*0.60*	0.8	*0.17*
Germany	4.1	*0.68*	11.3	*0.96*	21.4	*1.06*	27.9	*1.08*	23.6	*0.95*	10.0	*0.62*	1.8	*0.24*
Greece	7.2	*0.86*	16.9	*0.88*	28.9	*1.19*	29.4	*1.01*	14.2	*0.83*	3.2	*0.33*	0.2	*0.09*
Hungary	2.7	*0.33*	12.3	*0.83*	26.0	*1.15*	31.1	*1.07*	21.0	*0.87*	6.2	*0.57*	0.6	*0.16*
Iceland	5.8	*0.50*	14.7	*0.84*	25.9	*0.71*	28.3	*0.92*	19.0	*0.74*	5.6	*0.49*	0.7	*0.18*
Ireland	3.5	*0.47*	12.0	*0.82*	24.0	*0.91*	29.7	*0.98*	21.4	*0.87*	8.3	*0.62*	1.1	*0.19*
Italy	7.3	*0.46*	18.0	*0.62*	27.6	*0.78*	27.4	*0.61*	15.1	*0.58*	4.2	*0.31*	0.4	*0.09*
Japan	3.2	*0.45*	8.9	*0.73*	18.5	*0.86*	27.5	*0.85*	27.0	*1.14*	12.4	*0.63*	2.6	*0.33*
Korea, Republic of	2.5	*0.49*	8.7	*0.77*	21.2	*1.05*	31.8	*1.17*	25.5	*0.91*	9.2	*0.83*	1.1	*0.29*
Luxembourg	6.5	*0.39*	15.6	*0.65*	25.4	*0.66*	28.6	*0.93*	18.1	*0.71*	5.4	*0.34*	0.5	*0.11*
Mexico	18.2	*1.22*	32.8	*0.89*	30.8	*0.95*	14.8	*0.66*	3.2	*0.34*	0.3	*0.09*	#	†
Netherlands	2.3	*0.38*	10.7	*0.88*	21.1	*0.98*	26.9	*0.87*	25.8	*1.04*	11.5	*0.81*	1.7	*0.24*
New Zealand	4.0	*0.43*	9.7	*0.58*	19.7	*0.80*	25.1	*0.71*	23.9	*0.81*	13.6	*0.74*	4.0	*0.37*
Norway	5.9	*0.84*	15.2	*0.84*	27.3	*0.79*	28.5	*0.99*	17.1	*0.72*	5.5	*0.44*	0.6	*0.13*
Poland	3.2	*0.36*	13.8	*0.63*	27.5	*0.94*	29.4	*1.02*	19.3	*0.80*	6.1	*0.44*	0.7	*0.14*
Portugal	5.8	*0.76*	18.7	*1.05*	28.8	*0.92*	28.8	*1.22*	14.7	*0.88*	3.0	*0.35*	0.1	*0.05*
Slovak Republic	5.2	*0.60*	15.0	*0.87*	28.0	*0.96*	28.1	*0.99*	17.9	*1.02*	5.2	*0.49*	0.6	*0.14*
Spain	4.7	*0.44*	14.9	*0.69*	27.4	*0.77*	30.2	*0.68*	17.9	*0.75*	4.5	*0.38*	0.3	*0.10*
Sweden	3.8	*0.44*	12.6	*0.64*	25.2	*0.88*	29.5	*0.90*	21.1	*0.90*	6.8	*0.47*	1.1	*0.21*
Switzerland	4.5	*0.52*	11.6	*0.56*	21.8	*0.87*	28.2	*0.81*	23.5	*1.07*	9.1	*0.78*	1.4	*0.27*
Turkey	12.9	*0.83*	33.7	*1.31*	31.3	*1.42*	15.1	*1.06*	6.2	*1.15*	0.9	*0.32*	#	†
United Kingdom	4.8	*0.49*	11.9	*0.61*	21.8	*0.71*	25.9	*0.68*	21.8	*0.62*	10.9	*0.53*	2.9	*0.31*
United States	7.6	*0.94*	16.8	*0.88*	24.2	*0.94*	24.0	*0.79*	18.3	*0.97*	7.5	*0.62*	1.5	*0.25*
Non-OECD jurisdictions														
Argentina	28.3	*2.34*	27.9	*1.39*	25.6	*1.27*	13.6	*1.29*	4.1	*0.63*	0.4	*0.14*	#	†
Azerbaijan	19.4	*1.50*	53.1	*1.57*	22.4	*1.41*	4.7	*0.86*	0.4	*0.15*	#	†	#	†
Brazil	27.9	*0.99*	33.1	*0.96*	23.8	*0.93*	11.3	*0.88*	3.4	*0.42*	0.5	*0.21*	#	†
Bulgaria	18.3	*1.72*	24.3	*1.32*	25.2	*1.23*	18.8	*1.14*	10.3	*1.13*	2.6	*0.51*	0.4	*0.18*
Chile	13.1	*1.12*	26.7	*1.54*	29.9	*1.18*	20.1	*1.44*	8.4	*1.01*	1.8	*0.32*	0.1	*0.06*
Chinese Taipei	1.9	*0.29*	9.7	*0.82*	18.6	*0.86*	27.3	*0.80*	27.9	*1.03*	12.9	*0.77*	1.7	*0.24*
Colombia	26.2	*1.71*	34.0	*1.55*	27.2	*1.53*	10.6	*1.04*	1.9	*0.35*	0.2	*0.05*	#	†
Croatia	3.0	*0.43*	14.0	*0.71*	29.3	*0.91*	31.0	*0.99*	17.7	*0.86*	4.6	*0.44*	0.5	*0.12*
Estonia	1.0	*0.23*	6.7	*0.57*	21.0	*0.88*	33.7	*0.96*	26.2	*0.94*	10.1	*0.71*	1.4	*0.27*
Hong Kong-China	1.7	*0.36*	7.0	*0.68*	16.9	*0.81*	28.7	*0.95*	29.7	*0.95*	13.9	*0.80*	2.1	*0.30*
Indonesia	20.3	*1.71*	41.3	*2.23*	27.5	*1.46*	9.5	*1.99*	1.4	*0.53*	#	†	#	†
Israel	14.9	*1.18*	21.2	*1.01*	24.0	*0.95*	20.8	*0.96*	13.8	*0.80*	4.4	*0.49*	0.8	*0.18*
Jordan	16.2	*0.86*	28.2	*0.86*	30.8	*0.83*	18.7	*0.81*	5.6	*0.66*	0.6	*0.20*	#	†
Kyrgyz Republic	58.2	*1.56*	28.2	*1.13*	10.0	*0.81*	2.9	*0.39*	0.7	*0.18*	#	†	#	†
Latvia	3.6	*0.49*	13.8	*0.98*	29.0	*1.19*	32.9	*0.95*	16.6	*0.96*	3.8	*0.39*	0.3	*0.09*
Liechtenstein	2.6	*0.99*	10.3	*2.11*	21.0	*2.84*	28.7	*2.58*	25.2	*2.54*	10.0	*1.77*	2.2	*0.84*
Lithuania	4.3	*0.44*	16.0	*0.83*	27.4	*0.91*	29.8	*0.85*	17.5	*0.85*	4.5	*0.60*	0.4	*0.15*
Macao-China	1.4	*0.24*	8.9	*0.50*	26.0	*0.97*	35.7	*1.14*	22.8	*0.73*	5.0	*0.34*	0.3	*0.09*
Qatar	47.6	*0.62*	31.5	*0.63*	13.9	*0.49*	5.0	*0.35*	1.6	*0.14*	0.3	*0.09*	#	†
Republic of Montenegro	17.3	*0.79*	33.0	*1.20*	31.0	*0.91*	14.9	*0.65*	3.6	*0.37*	0.3	*0.11*	#	†
Republic of Serbia	11.9	*0.91*	26.6	*1.18*	32.3	*1.26*	21.8	*1.18*	6.6	*0.57*	0.8	*0.18*	#	†
Romania	16.0	*1.53*	30.9	*1.55*	31.8	*1.62*	16.6	*1.24*	4.2	*0.77*	0.5	*0.14*	#	†
Russian Federation	5.2	*0.65*	17.0	*1.08*	30.2	*0.93*	28.3	*1.32*	15.1	*1.09*	3.7	*0.46*	0.5	*0.13*
Slovenia	2.8	*0.34*	11.1	*0.72*	23.1	*0.68*	27.6	*1.08*	22.5	*1.13*	10.7	*0.57*	2.2	*0.29*
Thailand	12.6	*0.80*	33.5	*1.03*	33.2	*0.88*	16.3	*0.80*	4.0	*0.42*	0.4	*0.12*	#	†
Tunisia	27.7	*1.12*	35.1	*0.94*	25.0	*0.97*	10.2	*0.98*	1.9	*0.45*	0.1	*0.06*	#	†
Uruguay	16.7	*1.25*	25.4	*1.09*	29.8	*1.50*	19.7	*1.07*	6.9	*0.54*	1.3	*0.21*	0.1	*0.07*

† Not applicable.

Rounds to zero.

NOTE: To reach a particular proficiency level, a student must correctly answer a majority of items at that level. Students were classified into science literacy levels according to their scores. Exact cut point scores are as follows: below level 1 (a score less than or equal to 334.94); level 1 (a score greater than 334.94 and less than or equal to 409.54); level 2 (a score greater than 409.54 and less than or equal to 484.14); level 3 (a score greater than 484.14 and less than or equal to 558.73); level 4 (a score greater than 558.73 and less than or equal to 633.33); level 5 (a score greater than 633.33 and less than or equal to 707.93); and level 6 (a score greater than 707.93). The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.* Detail may not sum to totals because of rounding.

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-6. Average scores of 15-year-old students on combined science literacy scale, by jurisdiction: 2000, 2003, and 2006

Jurisdiction	2000 Average	2000 s.e.	2003 Average	2003 s.e.	2006 Average	2006 s.e.
OECD average	**500**	***0.7***	**500**	***0.6***	**500**	***0.5***
OECD jurisdictions						
Australia	528	*3.5*	525	*2.1*	527	*2.3*
Austria	519	*2.6*	491	*3.4*	511	*3.9*
Belgium	496	*4.3*	509	*2.4*	510	*2.5*
Canada	529	*1.6*	519	*2.0*	534	*2.0*
Czech Republic	511	*2.4*	523	*3.4*	513	*3.5*
Denmark	481	*2.8*	475	*3.0*	496	*3.1*
Finland	538	*2.5*	548	*1.9*	563	*2.0*
France	501	*3.2*	511	*3.0*	495	*3.4*
Germany	487	*2.4*	502	*3.6*	516	*3.8*
Greece	461	*4.9*	481	*3.8*	473	*3.2*
Hungary	496	*4.2*	503	*2.8*	504	*2.7*
Iceland	496	*2.2*	495	*1.5*	491	*1.6*
Ireland	513	*3.2*	505	*2.7*	508	*3.2*
Italy	478	*3.1*	487	*3.1*	475	*2.0*
Japan	550	*5.5*	548	*4.1*	531	*3.4*
Korea, Republic of	552	*2.7*	538	*3.5*	522	*3.4*
Luxembourg	443	*2.3*	483	*1.5*	486	*1.1*
Mexico	422	*3.2*	405	*3.5*	410	*2.7*
Netherlands[1]	—	—	524	*3.2*	525	*2.7*
New Zealand	528	*2.4*	521	*2.4*	530	*2.7*
Norway	500	*2.8*	484	*2.9*	487	*3.1*
Poland	483	*5.1*	498	*2.9*	498	*2.3*
Portugal	459	*4.0*	468	*3.5*	474	*3.0*
Slovak Republic	—	†	495	*3.7*	488	*2.6*
Spain	491	*3.0*	487	*2.6*	488	*2.6*
Sweden	512	*2.5*	506	*2.7*	503	*2.4*
Switzerland	496	*4.5*	513	*3.7*	512	*3.2*
Turkey	—	†	434	*5.9*	424	*3.8*
United Kingdom[2]	532	*2.7*	—	—	515	*2.3*
United States	500	*7.3*	491	*3.1*	489	*4.2*
Non-OECD jurisdictions						
Argentina	—	†	—	†	391	*6.1*
Azerbaijan	—	†	—	†	382	*2.8*
Brazil	—	†	390	*4.3*	390	*2.8*
Bulgaria	—	†	—	†	434	*6.1*
Chile	—	†	—	†	438	*4.3*
Chinese Taipei	—	†	—	†	532	*3.6*
Colombia	—	†	—	†	388	*3.4*
Croatia	—	†	—	†	493	*2.4*
Estonia	—	†	—	†	531	*2.5*
Hong Kong-China	—	†	540	*4.3*	542	*2.5*
Indonesia	—	†	395	*3.2*	393	*5.7*
Israel	—	†	—	†	454	*3.7*
Jordan	—	†	—	†	422	*2.8*
Kyrgyz Republic	—	†	—	†	322	*2.9*
Latvia	460	*5.6*	489	*3.9*	490	*3.0*
Liechtenstein	476	*7.1*	525	*4.3*	522	*4.1*
Lithuania	—	†	—	†	488	*2.8*
Macao-China	—	†	525	*3.0*	511	*1.1*
Qatar	—	†	—	†	349	*0.9*
Republic of Montenegro[3]	—	†	436	*3.5*	412	*1.1*
Republic of Serbia[3]	—	†	436	*3.5*	436	*3.0*
Romania	—	†	—	†	418	*4.2*
Russian Federation	460	*4.7*	489	*4.1*	479	*3.7*
Slovenia	—	†	—	†	519	*1.1*
Thailand	—	†	429	*2.7*	421	*2.1*
Tunisia	—	†	385	*2.6*	386	*3.0*
Uruguay	—	†	438	*2.9*	428	*2.7*

— Not available.
† Not applicable.
[1] Although the Netherlands participated in PISA in 2000, technical problems with its sample prevent its results from being discussed here.
[2] Because of low response rates, 2003 data for the United Kingdom are not discussed in this report.
[3] The Republics of Montenegro and Serbia were a united jurisdiction under the PISA 2003 assessment.
NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions with data available. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance in 2006 may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2000, 2003, and 2006.

Table C-7. Average scores of 15-year-old students on mathematics literacy scale, by jurisdiction: 2003 and 2006

Jurisdiction	2003		2006	
	Average	*s.e.*	Average	*s.e.*
OECD average	**500**	***0.6***	**498**	***0.5***
OECD jurisdictions				
Australia	524	*2.2*	520	*2.2*
Austria	506	*3.3*	505	*3.7*
Belgium	529	*2.3*	520	*3.0*
Canada	533	*1.8*	527	*2.0*
Czech Republic	517	*3.6*	510	*3.6*
Denmark	514	*2.7*	513	*2.6*
Finland	544	*1.9*	548	*2.3*
France	511	*2.5*	496	*3.2*
Germany	503	*3.3*	504	*3.9*
Greece	445	*3.9*	459	*3.0*
Hungary	490	*2.8*	491	*2.9*
Iceland	515	*1.4*	506	*1.8*
Ireland	503	*2.5*	501	*2.8*
Italy	466	*3.1*	462	*2.3*
Japan	534	*4.0*	523	*3.3*
Korea, Republic of	542	*3.2*	547	*3.8*
Luxembourg	493	*1.0*	490	*1.1*
Mexico	385	*3.6*	406	*2.9*
Netherlands	538	*3.1*	531	*2.6*
New Zealand	524	*2.3*	522	*2.4*
Norway	495	*2.4*	490	*2.6*
Poland	490	*2.5*	495	*2.4*
Portugal	466	*3.4*	466	*3.1*
Slovak Republic	498	*3.4*	492	*2.8*
Spain	485	*2.4*	480	*2.3*
Sweden	509	*2.6*	502	*2.4*
Switzerland	527	*3.4*	530	*3.2*
Turkey	423	*6.7*	424	*4.9*
United Kingdom[1]	—	—	495	*2.1*
United States	483	*3.0*	474	*4.0*
Non-OECD jurisdictions				
Argentina	—	†	381	*6.2*
Azerbaijan	—	†	476	*2.3*
Brazil	356	*4.8*	370	*2.9*
Bulgaria	—	†	413	*6.1*
Chile	—	†	411	*4.6*
Chinese Taipei	—	†	549	*4.1*
Colombia	—	†	370	*3.8*
Croatia	—	†	467	*2.4*
Estonia	—	†	515	*2.7*
Hong Kong-China	550	*4.5*	547	*2.7*
Indonesia	360	*3.9*	391	*5.6*
Israel	—	†	442	*4.3*
Jordan	—	†	384	*3.3*
Kyrgyz Republic	—	†	311	*3.4*
Latvia	483	*3.7*	486	*3.0*
Liechtenstein	536	*4.1*	525	*4.2*
Lithuania	—	†	486	*2.9*
Macao-China	527	*2.9*	525	*1.3*
Qatar	—	†	318	*1.0*
Republic of Montenegro[2]	437	*3.8*	399	*1.4*
Republic of Serbia[2]	437	*3.8*	435	*3.5*
Romania	—	†	415	*4.2*
Russian Federation	468	*4.2*	476	*3.9*
Slovenia	—	†	504	*1.0*
Thailand	417	*3.0*	417	*2.3*
Tunisia	359	*2.5*	365	*4.0*
Uruguay	422	*3.3*	427	*2.6*

— Not available.
† Not applicable.
[1] Because of low response rates, 2003 data for the United Kingdom are not discussed in this report.
[2] The Republics of Montenegro and Serbia were a united jurisdiction under the PISA 2003 assessment.
NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions with data available. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance in 2006 may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2003 and 2006.

Table C-8. Scores of 15-year-old students on mathematics literacy scale at 10th and 90th percentiles, by jurisdiction: 2006

	Percentiles			
	10th		90th	
Jurisdiction	Score	*s.e.*	Score	*s.e.*
OECD average	**379**	***0.9***	**615**	***0.8***
OECD jurisdictions				
Australia	406	*2.7*	633	*3.3*
Austria	373	*6.3*	630	*3.8*
Belgium	381	*6.6*	650	*2.4*
Canada	416	*3.3*	635	*2.3*
Czech Republic	376	*4.7*	644	*4.8*
Denmark	404	*4.3*	621	*3.4*
Finland	444	*3.4*	652	*2.8*
France	369	*5.4*	617	*3.8*
Germany	375	*6.8*	632	*3.8*
Greece	341	*5.6*	575	*4.1*
Hungary	377	*3.9*	609	*5.0*
Iceland	391	*3.6*	618	*3.2*
Ireland	396	*4.4*	608	*3.2*
Italy	341	*3.3*	584	*4.2*
Japan	404	*5.5*	638	*3.6*
Korea, Republic of	426	*6.1*	664	*6.9*
Luxembourg	368	*3.5*	610	*2.7*
Mexico	299	*4.9*	514	*3.3*
Netherlands	412	*5.0*	645	*3.3*
New Zealand	401	*4.1*	643	*4.0*
Norway	373	*3.8*	609	*3.3*
Poland	384	*3.4*	610	*3.7*
Portugal	348	*5.2*	583	*2.8*
Slovak Republic	370	*5.1*	611	*4.4*
Spain	366	*2.8*	593	*2.9*
Sweden	387	*4.2*	617	*2.8*
Switzerland	401	*4.7*	652	*3.7*
Turkey	316	*4.0*	550	*12.4*
United Kingdom	381	*3.3*	612	*3.2*
United States	358	*5.8*	593	*4.8*
Non-OECD jurisdictions				
Argentina	249	*9.8*	508	*7.6*
Azerbaijan	419	*2.2*	536	*3.6*
Brazil	255	*4.5*	487	*5.8*
Bulgaria	287	*7.2*	543	*8.4*
Chile	302	*4.3*	527	*6.6*
Chinese Taipei	409	*6.2*	677	*3.4*
Colombia	258	*5.6*	482	*3.8*
Croatia	361	*3.3*	576	*3.6*
Estonia	411	*4.3*	618	*3.2*
Hong Kong-China	423	*6.4*	665	*3.5*
Indonesia	293	*3.9*	498	*9.4*
Israel	304	*6.9*	581	*5.0*
Jordan	279	*4.3*	489	*5.0*
Kyrgyz Republic	204	*5.0*	423	*5.9*
Latvia	378	*5.2*	590	*3.4*
Liechtenstein	402	*11.1*	643	*9.5*
Lithuania	369	*4.3*	602	*4.9*
Macao-China	416	*3.1*	632	*2.4*
Qatar	212	*2.2*	438	*2.7*
Republic of Montenegro	291	*3.0*	510	*2.4*
Republic of Serbia	318	*5.0*	553	*3.9*
Romania	307	*7.4*	523	*7.1*
Russian Federation	363	*4.8*	592	*5.3*
Slovenia	390	*2.1*	623	*2.7*
Thailand	317	*3.5*	524	*3.7*
Tunisia	250	*3.9*	488	*7.8*
Uruguay	296	*4.4*	551	*5.5*

NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-9. Average scores of 15-year-old students on combined science literacy scale, by sex and jurisdiction: 2006

Jurisdiction	Male		Female		Male-female difference	
	Average	*s.e.*	Average	*s.e.*	Average	*s.e.*
OECD average	**501**	***0.7***	**499**	***0.6***	**2**	***0.7***
OECD jurisdictions						
Australia	527	*3.2*	527	*2.7*	#	†
Austria	515	*4.2*	507	*4.9*	8	*4.9*
Belgium	511	*3.3*	510	*3.2*	1	*4.1*
Canada	536	*2.5*	532	*2.1*	4	*2.2*
Czech Republic	515	*4.2*	510	*4.8*	5	*5.6*
Denmark	500	*3.6*	491	*3.4*	9	*3.2*
Finland	562	*2.6*	565	*2.4*	-3	*2.9*
France	497	*4.3*	494	*3.6*	3	*4.0*
Germany	519	*4.6*	512	*3.8*	7	*3.7*
Greece	468	*4.5*	479	*3.4*	-11	*4.7*
Hungary	507	*3.3*	501	*3.5*	6	*4.2*
Iceland	488	*2.6*	494	*2.1*	-6	*3.4*
Ireland	508	*4.3*	509	*3.3*	#	†
Italy	477	*2.8*	474	*2.5*	3	*3.5*
Japan	533	*4.9*	530	*5.1*	3	*7.4*
Korea, Republic of	521	*4.8*	523	*3.9*	-2	*5.5*
Luxembourg	491	*1.8*	482	*1.8*	9	*2.9*
Mexico	413	*3.2*	406	*2.6*	7	*2.2*
Netherlands	528	*3.2*	521	*3.1*	7	*3.0*
New Zealand	528	*3.9*	532	*3.6*	-4	*5.2*
Norway	484	*3.8*	489	*3.2*	-4	*3.4*
Poland	500	*2.7*	496	*2.6*	3	*2.5*
Portugal	477	*3.7*	472	*3.2*	5	*3.3*
Slovak Republic	491	*3.9*	485	*3.0*	6	*4.7*
Spain	491	*2.9*	486	*2.7*	4	*2.4*
Sweden	504	*2.7*	503	*2.9*	1	*3.0*
Switzerland	514	*3.3*	509	*3.6*	6	*2.7*
Turkey	418	*4.6*	430	*4.1*	-12	*4.1*
United Kingdom	520	*3.0*	510	*2.8*	10	*3.4*
United States	489	*5.1*	489	*4.0*	1	*3.5*
Non-OECD jurisdictions						
Argentina	384	*6.5*	397	*6.8*	-13	*5.6*
Azerbaijan	379	*3.1*	386	*2.7*	-8	*2.0*
Brazil	395	*3.2*	386	*2.9*	9	*2.3*
Bulgaria	426	*6.6*	443	*6.9*	-17	*5.8*
Chile	448	*5.4*	426	*4.4*	22	*4.8*
Chinese Taipei	536	*4.3*	529	*5.1*	7	*6.0*
Colombia	393	*4.1*	384	*4.1*	9	*4.6*
Croatia	492	*3.3*	494	*3.1*	-2	*4.1*
Estonia	530	*3.1*	533	*2.9*	-4	*3.1*
Hong Kong-China	546	*3.5*	539	*3.5*	7	*4.9*
Indonesia	399	*8.2*	387	*3.7*	12	*6.3*
Israel	456	*5.6*	452	*4.2*	3	*6.5*
Jordan	408	*4.5*	436	*3.3*	-29	*5.3*
Kyrgyz Republic	319	*3.6*	325	*3.0*	-6	*3.0*
Latvia	486	*3.5*	493	*3.2*	-7	*3.1*
Liechtenstein	516	*7.6*	527	*6.3*	-11	*11.1*
Lithuania	483	*3.1*	493	*3.1*	-9	*2.8*
Macao-China	513	*1.8*	509	*1.6*	4	*2.7*
Qatar	334	*1.2*	365	*1.3*	-32	*1.9*
Republic of Montenegro	411	*1.7*	413	*1.7*	-2	*2.6*
Republic of Serbia	433	*3.3*	438	*3.8*	-5	*3.8*
Romania	417	*4.1*	419	*4.8*	-2	*3.3*
Russian Federation	481	*4.1*	478	*3.7*	3	*2.7*
Slovenia	515	*2.0*	523	*1.9*	-8	*3.2*
Thailand	411	*3.4*	428	*2.5*	-17	*3.9*
Tunisia	383	*3.2*	388	*3.5*	-5	*3.4*
Uruguay	427	*4.0*	430	*2.7*	-3	*4.0*

† Not applicable.
# Rounds to zero.
NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Differences were computed using unrounded numbers. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-10. Percentage distribution of 15-year-old students at each proficiency level on combined science literacy scale, by sex and jurisdiction: 2006

	Below level 1				Level 1				Level 2				Level 3			
	Male		Female		Male		Female		Male		Female		Male		Female	
Jurisdiction	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*
OECD average	**5.6**	***0.15***	**4.7**	***0.13***	**14.1**	***0.19***	**14.0**	***0.19***	**23.4**	***0.23***	**24.7**	***0.23***	**26.4**	***0.22***	**28.5**	***0.23***
OECD jurisdictions																
Australia	3.6	*0.35*	2.5	*0.33*	10.3	*0.58*	9.3	*0.60*	19.7	*0.85*	20.8	*0.80*	26.6	*0.91*	28.9	*0.62*
Austria	3.6	*0.76*	5.0	*1.30*	11.6	*1.26*	12.5	*1.18*	22.7	*1.60*	20.8	*1.54*	27.5	*1.35*	29.0	*1.52*
Belgium	5.0	*1.03*	4.6	*0.66*	12.9	*1.00*	11.4	*0.81*	20.8	*1.05*	20.8	*1.14*	25.6	*0.91*	29.9	*1.40*
Canada	2.4	*0.37*	1.9	*0.33*	8.1	*0.68*	7.5	*0.69*	18.1	*0.70*	20.0	*0.85*	27.5	*0.74*	30.2	*0.91*
Czech Republic	2.6	*0.47*	4.7	*0.95*	11.7	*1.02*	12.5	*1.18*	24.5	*1.60*	22.0	*1.45*	28.0	*1.40*	27.5	*1.40*
Denmark	4.2	*0.66*	4.5	*0.83*	13.6	*1.00*	14.5	*1.03*	24.8	*1.24*	27.1	*1.27*	28.6	*1.16*	30.0	*1.43*
Finland	0.6	*0.21*	0.4	*0.17*	4.3	*0.61*	2.8	*0.49*	14.6	*0.83*	12.6	*0.91*	28.0	*1.26*	30.3	*1.27*
France	7.5	*0.96*	5.8	*0.74*	14.5	*1.18*	14.6	*1.23*	22.2	*1.39*	23.4	*1.39*	25.3	*1.50*	28.9	*1.34*
Germany	4.4	*0.84*	3.7	*0.67*	10.5	*1.09*	12.1	*1.19*	21.6	*1.23*	21.1	*1.26*	25.9	*1.21*	29.9	*1.47*
Greece	9.3	*1.28*	5.1	*0.81*	18.9	*1.29*	14.9	*0.95*	27.2	*1.24*	30.7	*1.83*	26.4	*1.37*	32.5	*1.53*
Hungary	2.8	*0.50*	2.6	*0.49*	12.8	*1.09*	11.9	*1.16*	25.2	*1.45*	26.9	*1.61*	28.7	*1.28*	33.6	*1.74*
Iceland	6.9	*0.69*	4.7	*0.66*	15.5	*1.02*	14.0	*1.11*	25.8	*1.37*	25.9	*1.23*	26.0	*1.61*	30.5	*1.46*
Ireland	4.1	*0.68*	3.0	*0.51*	12.5	*1.28*	11.5	*0.91*	23.2	*1.24*	24.8	*1.68*	28.8	*1.22*	30.6	*1.58*
Italy	8.0	*0.71*	6.5	*0.52*	17.5	*0.88*	18.5	*0.83*	25.9	*0.96*	29.3	*0.98*	27.4	*0.86*	27.4	*0.91*
Japan	3.6	*0.59*	2.8	*0.72*	9.2	*1.03*	8.5	*1.04*	18.1	*1.07*	18.8	*1.23*	25.8	*1.07*	29.2	*1.25*
Korea, Republic of	3.2	*0.70*	1.8	*0.45*	9.2	*0.99*	8.3	*1.05*	20.8	*1.57*	21.5	*1.13*	30.2	*1.42*	33.3	*1.43*
Luxembourg	7.0	*0.57*	6.1	*0.58*	15.1	*1.02*	16.1	*1.01*	23.8	*1.25*	27.0	*1.07*	27.2	*1.31*	29.9	*1.10*
Mexico	17.4	*1.55*	18.9	*1.28*	32.1	*1.27*	33.4	*1.10*	30.5	*1.39*	31.0	*1.08*	15.8	*0.85*	13.9	*0.76*
Netherlands	2.4	*0.47*	2.2	*0.49*	9.9	*0.99*	11.5	*1.16*	20.7	*1.38*	21.6	*1.21*	27.3	*1.23*	26.6	*1.25*
New Zealand	5.0	*0.69*	3.1	*0.45*	10.3	*0.84*	9.1	*0.77*	19.4	*1.16*	20.0	*1.18*	24.1	*1.17*	26.0	*0.96*
Norway	7.3	*1.19*	4.3	*0.68*	15.1	*0.91*	15.3	*1.12*	26.5	*1.13*	28.1	*1.08*	27.7	*1.12*	29.4	*1.57*
Poland	3.7	*0.54*	2.7	*0.43*	13.6	*0.77*	13.9	*0.83*	26.9	*1.51*	28.1	*1.04*	28.6	*1.36*	30.3	*1.21*
Portugal	5.9	*0.93*	5.6	*0.88*	18.3	*1.53*	19.0	*1.14*	28.3	*1.19*	29.3	*1.17*	27.9	*1.46*	29.8	*1.49*
Slovak Republic	5.5	*0.94*	4.8	*0.70*	14.6	*1.07*	15.5	*1.33*	27.0	*1.45*	29.2	*1.23*	27.4	*1.45*	28.8	*1.32*
Spain	5.2	*0.54*	4.3	*0.55*	14.4	*0.89*	15.4	*0.87*	26.4	*1.03*	28.3	*1.19*	29.7	*0.99*	30.7	*0.89*
Sweden	4.1	*0.62*	3.4	*0.48*	13.1	*0.90*	12.0	*0.86*	24.0	*1.14*	26.4	*1.56*	28.6	*1.40*	30.4	*1.44*
Switzerland	4.6	*0.63*	4.4	*0.52*	10.9	*0.64*	12.2	*0.80*	20.8	*1.06*	22.8	*1.06*	28.5	*1.06*	27.8	*1.04*
Turkey	15.2	*1.21*	10.1	*1.12*	35.0	*1.60*	32.2	*1.79*	29.0	*1.60*	34.1	*1.86*	13.8	*1.16*	16.6	*1.39*
United Kingdom	5.3	*0.74*	4.3	*0.48*	11.4	*0.90*	12.4	*0.88*	20.5	*0.80*	23.0	*1.03*	24.1	*0.86*	27.7	*1.02*
United States	8.3	*1.23*	6.8	*0.85*	17.4	*1.28*	16.2	*1.06*	22.3	*1.18*	26.2	*1.16*	23.4	*1.10*	24.6	*1.02*
Non-OECD jurisdictions																
Argentina	30.7	*2.65*	26.2	*2.45*	28.1	*1.74*	27.8	*1.66*	25.1	*1.61*	26.0	*1.52*	12.2	*1.30*	14.9	*1.70*
Azerbaijan	22.4	*1.82*	16.1	*1.67*	52.2	*1.83*	54.1	*1.81*	20.2	*1.61*	24.8	*1.77*	4.7	*0.96*	4.7	*0.95*
Brazil	26.8	*1.17*	28.9	*1.21*	31.6	*1.24*	34.4	*1.11*	24.9	*1.16*	22.8	*1.20*	11.9	*1.17*	10.7	*0.88*
Bulgaria	21.2	*2.14*	15.2	*1.80*	25.5	*1.57*	23.1	*1.86*	23.4	*1.59*	27.0	*1.76*	17.3	*1.33*	20.5	*1.61*
Chile	10.8	*1.17*	15.7	*1.42*	25.0	*1.95*	28.6	*1.54*	29.7	*1.53*	30.1	*1.46*	22.2	*1.54*	17.6	*1.62*
Chinese Taipei	2.0	*0.38*	1.9	*0.41*	9.7	*0.97*	9.7	*1.06*	17.4	*0.92*	19.9	*1.38*	26.4	*1.21*	28.3	*1.13*
Colombia	25.2	*1.91*	27.0	*2.01*	32.2	*1.76*	35.5	*2.12*	27.6	*1.98*	26.9	*1.96*	12.3	*1.52*	9.1	*1.21*
Croatia	3.4	*0.68*	2.6	*0.50*	14.8	*1.02*	13.1	*1.18*	28.8	*1.18*	29.7	*1.47*	29.7	*1.17*	32.3	*1.44*
Estonia	1.2	*0.37*	0.7	*0.21*	7.4	*0.76*	6.0	*0.68*	21.0	*1.11*	21.0	*1.15*	33.2	*1.18*	34.2	*1.45*
Hong Kong-China	1.9	*0.54*	1.5	*0.33*	7.3	*0.85*	6.7	*0.83*	15.9	*1.09*	17.9	*1.13*	26.8	*1.13*	30.5	*1.56*
Indonesia	18.7	*2.18*	22.0	*1.61*	39.9	*3.15*	42.7	*1.95*	28.0	*2.01*	27.0	*1.66*	11.5	*3.04*	7.3	*1.21*
Israel	16.0	*1.62*	13.8	*1.41*	21.3	*1.33*	21.1	*1.14*	21.7	*1.07*	26.3	*1.42*	19.6	*1.09*	22.0	*1.37*
Jordan	21.6	*1.40*	10.8	*1.03*	29.2	*1.36*	27.1	*1.17*	27.8	*1.22*	33.7	*1.04*	16.2	*1.21*	21.2	*1.28*
Kyrgyz Republic	60.0	*1.85*	56.6	*1.69*	26.2	*1.31*	29.9	*1.30*	9.7	*0.97*	10.4	*0.93*	3.2	*0.60*	2.6	*0.43*
Latvia	4.0	*0.64*	3.2	*0.61*	15.1	*1.14*	12.7	*1.15*	29.3	*1.62*	28.7	*1.51*	31.9	*1.52*	33.9	*1.32*
Liechtenstein	3.0	*1.73*	2.3	*1.22*	10.2	*3.89*	10.3	*2.63*	22.8	*4.58*	19.4	*3.56*	31.0	*4.43*	26.7	*3.35*
Lithuania	4.9	*0.55*	3.8	*0.57*	17.2	*0.99*	14.8	*1.26*	27.9	*1.24*	26.8	*1.38*	28.5	*1.20*	31.1	*1.25*
Macao-China	1.8	*0.32*	1.0	*0.27*	9.5	*0.70*	8.2	*0.67*	24.2	*0.99*	27.8	*1.41*	34.4	*1.52*	36.9	*1.54*
Qatar	57.7	*0.96*	37.3	*0.92*	26.2	*0.99*	36.9	*0.96*	9.5	*0.72*	18.3	*0.88*	4.3	*0.38*	5.9	*0.57*
Republic of Montenegro	17.7	*0.95*	16.8	*1.15*	33.1	*1.55*	32.8	*1.77*	30.4	*1.43*	31.6	*1.11*	15.0	*0.91*	14.8	*0.91*
Republic of Serbia	12.9	*1.09*	10.9	*1.19*	27.9	*1.44*	25.3	*1.48*	31.0	*1.56*	33.5	*1.52*	20.6	*1.48*	23.1	*1.43*
Romania	17.6	*1.57*	14.3	*1.93*	30.7	*1.73*	31.2	*1.96*	29.4	*1.56*	34.2	*2.46*	16.9	*1.63*	16.2	*1.54*
Russian Federation	5.6	*0.77*	4.9	*0.72*	17.0	*1.15*	16.9	*1.35*	29.3	*1.35*	31.1	*1.09*	27.5	*1.97*	29.1	*1.28*
Slovenia	3.2	*0.40*	2.4	*0.51*	12.1	*0.97*	10.1	*0.75*	24.0	*1.14*	22.3	*0.86*	26.6	*1.56*	28.6	*1.13*
Thailand	17.1	*1.56*	9.3	*0.92*	34.7	*1.41*	32.6	*1.19*	29.1	*1.32*	36.3	*1.18*	14.9	*1.05*	17.3	*1.08*
Tunisia	29.3	*1.43*	26.2	*1.36*	34.2	*1.27*	35.9	*1.40*	24.5	*1.20*	25.4	*1.30*	9.8	*1.19*	10.6	*1.27*
Uruguay	18.2	*1.84*	15.3	*1.26*	25.8	*1.85*	25.0	*1.19*	27.8	*1.92*	31.7	*1.93*	18.9	*1.45*	20.5	*1.46*

See notes at end of table.

Table C-10. Percentage distribution of 15-year-old students at each proficiency level on combined science literacy scale, by sex and jurisdiction: 2006—Continued

	Level 4				Level 5				Level 6			
	Male		Female		Male		Female		Male		Female	
Jurisdiction	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*	Percent	*s.e.*
OECD average	**20.5**	***0.21***	**20.2**	***0.21***	**8.5**	***0.15***	**6.9**	***0.13***	**1.5**	***0.06***	**1.0**	***0.05***
OECD jurisdictions												
Australia	24.2	*0.72*	25.0	*0.71*	12.3	*0.73*	11.2	*0.68*	3.3	*0.43*	2.4	*0.29*
Austria	23.3	*1.38*	24.1	*1.62*	9.7	*0.97*	7.9	*0.89*	1.6	*0.30*	0.8	*0.21*
Belgium	24.5	*0.92*	24.4	*1.21*	9.9	*0.67*	8.3	*0.70*	1.3	*0.24*	0.6	*0.22*
Canada	28.1	*0.94*	27.2	*0.85*	12.9	*0.61*	11.2	*0.77*	2.8	*0.29*	2.0	*0.32*
Czech Republic	21.4	*1.42*	22.2	*1.27*	9.9	*1.02*	9.6	*1.10*	2.0	*0.41*	1.6	*0.35*
Denmark	21.0	*1.24*	18.1	*1.12*	7.0	*0.92*	5.2	*0.68*	0.8	*0.26*	0.6	*0.28*
Finland	30.8	*1.07*	33.7	*1.16*	17.0	*0.96*	16.9	*0.97*	4.6	*0.50*	3.3	*0.48*
France	20.9	*1.29*	20.8	*1.26*	8.5	*0.80*	6.0	*0.83*	1.1	*0.33*	0.5	*0.16*
Germany	23.8	*1.35*	23.3	*1.07*	11.5	*1.03*	8.4	*0.73*	2.2	*0.37*	1.4	*0.38*
Greece	14.2	*1.06*	14.1	*1.09*	3.7	*0.50*	2.7	*0.46*	0.3	*0.14*	0.1	*0.08*
Hungary	22.0	*1.14*	19.8	*1.28*	7.6	*0.94*	4.8	*0.74*	0.8	*0.21*	0.4	*0.20*
Iceland	19.2	*1.11*	18.8	*1.00*	5.8	*0.71*	5.4	*0.75*	0.8	*0.22*	0.7	*0.29*
Ireland	21.1	*1.10*	21.6	*1.23*	8.9	*0.92*	7.6	*0.75*	1.4	*0.31*	0.9	*0.29*
Italy	15.8	*0.73*	14.4	*0.73*	4.9	*0.44*	3.6	*0.41*	0.6	*0.13*	0.3	*0.09*
Japan	26.5	*1.51*	27.5	*1.58*	13.7	*0.93*	11.2	*0.91*	3.3	*0.48*	2.0	*0.35*
Korea, Republic of	25.5	*1.32*	25.5	*1.31*	9.9	*1.13*	8.6	*0.89*	1.3	*0.37*	0.9	*0.33*
Luxembourg	19.6	*1.07*	16.5	*0.91*	6.6	*0.63*	4.1	*0.53*	0.8	*0.23*	0.3	*0.16*
Mexico	3.8	*0.41*	2.6	*0.36*	0.3	*0.12*	0.2	*0.10*	#	†	#	†
Netherlands	24.9	*1.30*	26.8	*1.27*	13.0	*1.13*	9.9	*0.75*	2.0	*0.44*	1.3	*0.31*
New Zealand	22.8	*1.10*	24.9	*1.07*	14.0	*0.98*	13.3	*1.05*	4.4	*0.67*	3.6	*0.50*
Norway	16.7	*1.22*	17.5	*1.20*	6.0	*0.68*	4.9	*0.69*	0.7	*0.20*	0.5	*0.18*
Poland	19.1	*1.09*	19.5	*1.06*	7.2	*0.65*	5.0	*0.62*	0.9	*0.25*	0.5	*0.18*
Portugal	15.5	*1.00*	14.0	*1.20*	3.9	*0.60*	2.2	*0.31*	0.1	*0.09*	#	†
Slovak Republic	18.8	*1.35*	17.0	*1.32*	6.0	*0.79*	4.4	*0.55*	0.8	*0.29*	0.4	*0.18*
Spain	18.7	*0.95*	17.1	*0.87*	5.1	*0.49*	4.0	*0.44*	0.5	*0.16*	0.2	*0.08*
Sweden	21.5	*1.13*	20.6	*1.27*	7.3	*0.69*	6.2	*0.75*	1.2	*0.34*	1.0	*0.26*
Switzerland	24.0	*1.24*	23.0	*1.27*	9.7	*0.87*	8.4	*0.95*	1.4	*0.33*	1.4	*0.35*
Turkey	6.2	*1.28*	6.1	*1.23*	0.9	*0.37*	0.9	*0.40*	#	†	#	†
United Kingdom	22.5	*0.81*	21.1	*0.97*	12.3	*0.78*	9.4	*0.71*	3.7	*0.48*	2.1	*0.39*
United States	18.6	*1.33*	18.0	*1.01*	8.4	*0.84*	6.7	*0.78*	1.6	*0.30*	1.5	*0.35*
Non-OECD jurisdictions												
Argentina	3.4	*0.58*	4.6	*0.92*	0.4	*0.16*	0.5	*0.22*	#	†	#	†
Azerbaijan	0.4	*0.15*	0.4	*0.19*	#	†	#	†	#	†	#	†
Brazil	4.0	*0.55*	2.8	*0.50*	0.7	*0.31*	0.4	*0.18*	0.1	*0.06*	#	†
Bulgaria	9.2	*1.18*	11.4	*1.49*	2.8	*0.62*	2.4	*0.52*	0.5	*0.23*	0.4	*0.18*
Chile	9.9	*1.29*	6.6	*1.02*	2.3	*0.54*	1.2	*0.42*	0.1	*0.08*	0.1	*0.11*
Chinese Taipei	28.8	*1.18*	26.9	*1.51*	13.8	*1.08*	12.0	*1.13*	2.0	*0.42*	1.4	*0.30*
Colombia	2.5	*0.53*	1.4	*0.42*	0.2	*0.11*	0.1	*0.09*	#	†	#	†
Croatia	17.9	*1.01*	17.5	*1.24*	4.7	*0.52*	4.4	*0.63*	0.7	*0.16*	0.4	*0.15*
Estonia	25.4	*1.40*	27.0	*1.25*	10.2	*0.90*	10.0	*0.98*	1.6	*0.32*	1.2	*0.33*
Hong Kong-China	30.4	*1.27*	29.1	*1.25*	14.7	*1.06*	13.0	*1.17*	2.8	*0.50*	1.3	*0.27*
Indonesia	1.8	*0.84*	1.0	*0.40*	#	†	#	†	#	†	#	†
Israel	14.7	*1.25*	12.9	*0.96*	5.4	*0.75*	3.5	*0.47*	1.3	*0.28*	0.3	*0.16*
Jordan	4.6	*0.88*	6.5	*0.75*	0.6	*0.26*	0.7	*0.21*	#	†	#	†
Kyrgyz Republic	1.0	*0.30*	0.5	*0.16*	#	†	#	†	#	†	#	†
Latvia	15.4	*1.24*	17.7	*1.18*	4.0	*0.60*	3.7	*0.46*	0.3	*0.13*	0.2	*0.11*
Liechtenstein	20.8	*4.08*	29.0	*3.72*	10.6	*2.79*	9.5	*2.34*	1.5	*1.17*	2.8	*1.35*
Lithuania	16.9	*1.09*	18.1	*1.18*	4.1	*0.62*	5.0	*0.76*	0.4	*0.21*	0.4	*0.20*
Macao-China	23.5	*1.60*	22.0	*1.11*	6.2	*0.58*	3.8	*0.56*	0.3	*0.18*	0.2	*0.12*
Qatar	1.9	*0.24*	1.4	*0.22*	0.4	*0.14*	0.2	*0.10*	#	†	#	†
Republic of Montenegro	3.5	*0.50*	3.8	*0.54*	0.3	*0.16*	0.2	*0.17*	#	†	#	†
Republic of Serbia	6.5	*0.71*	6.6	*0.69*	1.0	*0.27*	0.6	*0.21*	#	†	#	†
Romania	4.6	*0.83*	3.9	*1.02*	0.7	*0.24*	0.2	*0.13*	#	†	#	†
Russian Federation	15.6	*1.45*	14.6	*1.12*	4.4	*0.69*	3.0	*0.42*	0.7	*0.21*	0.3	*0.15*
Slovenia	21.5	*1.54*	23.5	*1.44*	10.2	*0.95*	11.2	*1.00*	2.4	*0.52*	1.9	*0.38*
Thailand	3.8	*0.62*	4.1	*0.55*	0.5	*0.20*	0.4	*0.14*	#	†	#	†
Tunisia	2.0	*0.47*	1.8	*0.58*	0.1	*0.09*	0.1	*0.12*	#	†	#	†
Uruguay	7.3	*0.67*	6.5	*0.67*	1.7	*0.35*	0.9	*0.30*	0.2	*0.12*	0.1	*0.06*

† Not applicable.

Rounds to zero.

NOTE: To reach a particular proficiency level, a student must correctly answer a majority of items at that level. Students were classified into science literacy levels according to their scores. Exact cut point scores are as follows: below level 1 (a score less than or equal to 334.94); level 1 (a score greater than 334.94 and less than or equal to 409.54); level 2 (a score greater than 409.54 and less than or equal to 484.14); level 3 (a score greater than 484.14 and less than or equal to 558.73); level 4 (a score greater than 558.73 and less than or equal to 633.33); level 5 (a score greater than 633.33 and less than or equal to 707.93); and level 6 (a score greater than 707.93). The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.* Detail may not sum to totals because of rounding.

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-11. Average scores of 15-year-old students on science literacy subscales, by sex and jurisdiction: 2006

	Identifying scientific issues						Explaining phenomena scientifically					
	Male		Female		Male-female difference		Male		Female		Male-female difference	
Jurisdiction	Average	*s.e.*	Average	*s.e.*	Average	*s.e.*	Average	*s.e.*	Average	*s.e.*	Average	*s.e.*
OECD average	**490**	***0.7***	**508**	***0.6***	**-17**	***0.7***	**508**	***0.7***	**493**	***0.6***	**15**	***0.7***
OECD jurisdictions												
Australia	525	*3.2*	546	*2.6*	-21	*3.6*	527	*3.1*	513	*2.7*	13	*3.6*
Austria	495	*4.2*	516	*4.7*	-22	*4.6*	526	*4.4*	507	*4.7*	19	*4.8*
Belgium	508	*3.8*	523	*3.1*	-14	*4.3*	510	*3.4*	494	*3.1*	16	*4.1*
Canada	525	*2.7*	539	*2.4*	-14	*2.4*	539	*2.6*	522	*2.3*	17	*2.5*
Czech Republic	492	*4.8*	511	*5.3*	-19	*5.7*	537	*4.3*	516	*4.6*	21	*5.7*
Denmark	488	*3.5*	499	*3.2*	-11	*3.2*	512	*3.8*	491	*3.7*	21	*3.4*
Finland	542	*2.7*	568	*2.6*	-26	*2.8*	571	*2.5*	562	*2.5*	9	*3.0*
France	491	*4.6*	507	*3.7*	-16	*4.7*	489	*4.2*	474	*3.4*	15	*4.1*
Germany	502	*4.5*	518	*3.9*	-16	*3.4*	529	*4.5*	508	*3.7*	21	*3.7*
Greece	453	*4.1*	485	*3.1*	-31	*4.3*	478	*4.3*	475	*3.0*	3	*4.2*
Hungary	477	*3.4*	489	*3.3*	-13	*4.1*	529	*3.2*	507	*3.6*	22	*4.4*
Iceland	479	*2.9*	509	*2.4*	-30	*4.1*	491	*2.6*	485	*2.1*	6	*3.7*
Ireland	508	*4.4*	524	*3.5*	-16	*4.6*	510	*4.4*	501	*3.5*	9	*4.6*
Italy	466	*2.9*	483	*2.5*	-17	*3.4*	487	*2.8*	472	*2.5*	15	*3.4*
Japan	513	*5.1*	531	*6.6*	-18	*8.5*	535	*4.6*	519	*4.4*	16	*6.6*
Korea, Republic of	508	*4.9*	530	*4.2*	-22	*5.7*	517	*4.8*	506	*4.0*	11	*5.7*
Luxembourg	477	*1.7*	489	*1.8*	-11	*2.8*	495	*1.8*	471	*2.0*	25	*3.0*
Mexico	418	*2.9*	425	*2.8*	-7	*2.2*	415	*3.3*	398	*2.6*	18	*2.3*
Netherlands	527	*3.8*	539	*3.5*	-12	*3.2*	531	*3.1*	512	*3.1*	18	*3.0*
New Zealand	525	*3.7*	547	*3.7*	-22	*4.9*	528	*4.0*	517	*3.6*	11	*5.2*
Norway	478	*3.9*	501	*3.3*	-24	*3.7*	498	*3.9*	492	*3.2*	6	*3.9*
Poland	476	*2.8*	490	*2.7*	-13	*2.5*	514	*2.9*	498	*2.8*	17	*2.7*
Portugal	480	*3.6*	493	*3.4*	-13	*3.1*	477	*3.6*	462	*3.0*	16	*3.2*
Slovak Republic	465	*4.5*	485	*3.6*	-20	*5.1*	512	*4.0*	490	*3.0*	22	*4.7*
Spain	482	*2.7*	496	*2.6*	-15	*2.1*	499	*2.8*	481	*2.7*	18	*2.6*
Sweden	491	*2.9*	507	*3.1*	-16	*3.0*	516	*3.0*	504	*3.5*	12	*3.1*
Switzerland	510	*3.1*	520	*3.3*	-10	*2.4*	517	*3.4*	498	*3.9*	18	*2.8*
Turkey	414	*4.1*	443	*3.6*	-29	*3.8*	423	*4.7*	423	*4.5*	1	*4.1*
United Kingdom	510	*2.9*	517	*2.8*	-7	*3.2*	527	*3.0*	506	*2.7*	21	*3.5*
United States	484	*4.6*	500	*3.8*	-16	*3.6*	492	*5.3*	480	*4.0*	13	*3.6*
Non-OECD jurisdictions												
Argentina	381	*5.8*	408	*6.4*	-27	*5.2*	387	*6.4*	386	*7.0*	0	*5.8*
Azerbaijan	349	*3.3*	357	*3.3*	-8	*2.3*	408	*3.3*	417	*3.0*	-9	*1.9*
Brazil	394	*3.2*	402	*3.0*	-7	*2.5*	400	*3.0*	382	*2.9*	19	*2.4*
Bulgaria	411	*6.6*	445	*7.1*	-34	*5.6*	442	*6.5*	447	*6.5*	-5	*5.8*
Chile	445	*5.0*	443	*4.1*	3	*4.5*	448	*5.1*	414	*4.1*	34	*4.6*
Chinese Taipei	506	*4.4*	512	*5.0*	-6	*5.8*	554	*4.3*	535	*5.3*	19	*6.1*
Colombia	401	*4.4*	404	*4.0*	-3	*4.8*	388	*4.3*	371	*4.3*	18	*4.8*
Croatia	480	*3.5*	507	*3.1*	-27	*4.1*	498	*3.2*	487	*3.3*	11	*4.1*
Estonia	504	*3.1*	528	*2.6*	-25	*2.8*	544	*3.2*	537	*3.0*	6	*3.3*
Hong Kong-China	520	*4.1*	535	*4.5*	-15	*5.9*	560	*3.5*	539	*3.3*	21	*4.6*
Indonesia	397	*8.0*	389	*3.6*	8	*6.0*	403	*7.0*	386	*3.8*	17	*5.7*
Israel	451	*5.9*	463	*4.0*	-12	*6.6*	451	*5.4*	436	*4.0*	16	*6.4*
Jordan	393	*4.6*	425	*2.8*	-32	*5.1*	427	*4.6*	448	*4.1*	-21	*6.0*
Kyrgyz Republic	311	*3.6*	330	*3.3*	-20	*2.9*	335	*3.9*	333	*2.9*	2	*3.0*
Latvia	473	*3.7*	504	*3.5*	-31	*3.1*	491	*3.6*	481	*3.2*	10	*3.3*
Liechtenstein	508	*7.0*	534	*5.7*	-26	*10.3*	519	*7.5*	513	*6.4*	6	*11.1*
Lithuania	463	*2.9*	489	*3.0*	-26	*2.7*	499	*3.3*	490	*3.4*	9	*3.1*
Macao-China	483	*1.9*	498	*1.6*	-15	*2.6*	527	*2.0*	513	*1.6*	14	*2.7*
Qatar	334	*1.2*	371	*1.3*	-37	*2.1*	342	*1.4*	371	*1.6*	-29	*2.3*
Republic of Montenegro	393	*2.0*	409	*1.8*	-16	*2.9*	421	*1.8*	412	*1.7*	9	*2.7*
Republic of Serbia	420	*3.3*	441	*3.6*	-21	*3.7*	444	*3.7*	438	*3.8*	6	*4.1*
Romania	401	*3.6*	418	*4.4*	-17	*3.5*	431	*4.3*	421	*4.5*	10	*3.6*
Russian Federation	453	*4.6*	472	*4.1*	-20	*2.6*	493	*4.0*	474	*3.4*	19	*2.6*
Slovenia	504	*2.0*	530	*2.0*	-27	*2.8*	528	*2.3*	518	*2.2*	10	*3.3*
Thailand	394	*3.7*	427	*2.8*	-33	*4.1*	418	*3.4*	421	*2.2*	-3	*3.6*
Tunisia	373	*3.9*	394	*4.2*	-21	*3.4*	386	*3.1*	381	*3.5*	5	*3.1*
Uruguay	418	*4.2*	439	*2.8*	-21	*3.9*	429	*4.0*	418	*3.1*	11	*4.0*

See notes at end of table.

Table C-11. Average scores of 15-year-old students on science literacy subscales, by sex and jurisdiction: 2006—Continued

	Using scientific evidence					
	Male		Female		Male-female difference	
Jurisdiction	Average	*s.e.*	Average	*s.e.*	Average	*s.e.*
OECD average	**498**	***0.8***	**501**	***0.7***	**-3**	***0.8***
OECD jurisdictions						
Australia	530	*3.4*	533	*3.0*	-3	*4.2*
Austria	509	*4.9*	500	*6.2*	9	*6.1*
Belgium	512	*3.8*	521	*3.8*	-9	*4.7*
Canada	541	*2.7*	542	*2.3*	-1	*2.3*
Czech Republic	501	*5.0*	500	*5.4*	1	*6.5*
Denmark	490	*4.1*	487	*4.0*	3	*3.8*
Finland	564	*3.0*	571	*2.7*	-7	*3.3*
France	509	*5.0*	513	*4.2*	-4	*4.7*
Germany	517	*5.6*	513	*4.5*	4	*4.3*
Greece	456	*5.6*	475	*3.7*	-20	*5.4*
Hungary	497	*4.1*	498	*4.5*	-1	*5.2*
Iceland	487	*3.1*	495	*2.5*	-7	*4.4*
Ireland	503	*4.8*	509	*3.5*	-7	*4.8*
Italy	466	*3.2*	468	*3.1*	-2	*4.2*
Japan	543	*5.8*	545	*6.4*	-2	*8.9*
Korea, Republic of	535	*5.2*	542	*4.5*	-8	*6.4*
Luxembourg	493	*2.0*	490	*2.2*	3	*3.5*
Mexico	404	*3.7*	401	*3.0*	3	*2.7*
Netherlands	527	*3.8*	524	*3.7*	3	*3.5*
New Zealand	532	*4.4*	541	*4.3*	-10	*5.8*
Norway	469	*4.2*	476	*3.9*	-7	*3.8*
Poland	492	*3.0*	495	*3.0*	-3	*2.8*
Portugal	473	*4.2*	471	*4.0*	2	*3.8*
Slovak Republic	478	*4.8*	478	*3.6*	#	*†*
Spain	484	*3.4*	485	*3.1*	-1	*2.5*
Sweden	494	*3.1*	499	*3.2*	-5	*3.4*
Switzerland	520	*3.6*	517	*3.9*	2	*2.9*
Turkey	410	*5.2*	426	*4.6*	-16	*4.7*
United Kingdom	517	*3.1*	510	*3.1*	6	*3.8*
United States	486	*6.1*	491	*4.6*	-5	*4.1*
Non-OECD jurisdictions						
Argentina	374	*7.4*	396	*7.7*	-23	*6.2*
Azerbaijan	342	*4.5*	347	*3.9*	-6	*2.4*
Brazil	382	*3.9*	375	*3.8*	6	*2.7*
Bulgaria	404	*8.0*	430	*8.2*	-26	*6.7*
Chile	447	*6.2*	431	*5.2*	16	*5.3*
Chinese Taipei	532	*4.5*	532	*5.1*	#	*†*
Colombia	386	*4.5*	381	*4.8*	5	*4.9*
Croatia	488	*4.1*	493	*3.5*	-5	*4.8*
Estonia	529	*3.2*	533	*3.0*	-5	*3.3*
Hong Kong-China	544	*3.8*	541	*4.0*	2	*5.5*
Indonesia	388	*10.2*	383	*5.0*	5	*7.3*
Israel	456	*6.7*	464	*5.4*	-8	*7.6*
Jordan	385	*5.5*	424	*3.6*	-39	*6.3*
Kyrgyz Republic	280	*4.7*	295	*3.9*	-15	*3.7*
Latvia	484	*4.1*	497	*3.6*	-13	*3.6*
Liechtenstein	524	*8.2*	544	*6.8*	-20	*12.2*
Lithuania	478	*3.7*	495	*3.3*	-17	*3.0*
Macao-China	512	*2.0*	511	*1.6*	#	*†*
Qatar	307	*1.5*	341	*1.9*	-35	*2.5*
Republic of Montenegro	403	*2.0*	411	*2.0*	-8	*3.1*
Republic of Serbia	419	*4.0*	431	*4.8*	-11	*4.9*
Romania	403	*6.0*	412	*6.7*	-9	*4.6*
Russian Federation	478	*4.5*	483	*4.4*	-5	*3.1*
Slovenia	510	*2.3*	522	*2.0*	-12	*3.4*
Thailand	409	*4.2*	433	*2.7*	-24	*4.5*
Tunisia	377	*4.1*	387	*4.3*	-10	*3.9*
Uruguay	425	*4.0*	433	*3.5*	-8	*4.1*

† Not applicable.
# Rounds to zero.
NOTE: The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because the Program for International Student Assessment (PISA) is principally an OECD study, the results for non-OECD jurisdictions are displayed separately from those of the OECD jurisdictions and are not included in the OECD average. Differences were computed using unrounded numbers. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Table C-12. Average scores of U.S. 15-year-old students on combined science literacy scale, by race/ethnicity: 2006

Race/ethnicity	Average	*s.e.*
U.S. average	**489**	***4.2***
White, non-Hispanic	523	*3.0*
Black, non-Hispanic	409	*8.8*
Hispanic	439	*4.7*
Asian, non-Hispanic	499	*9.7*
American Indian/Alaska Native, non-Hispanic	436	*12.0*
Native Hawaiian/Other Pacific Islander, non-Hispanic	483	*24.5*
More than one race, non-Hispanic	501	*8.0*
OECD average	**500**	***0.5***

NOTE: Black includes African American, and Hispanic includes Latino. Students who identified themselves as being of Hispanic origin were classified as Hispanic, regardless of their race. To reach a particular proficiency level, a student must correctly answer a majority of items at that level. Students were classified into science literacy levels according to their scores. Exact cut point scores are as follows: below level 1 (a score less than or equal to 334.94); level 1 (a score greater than 334.94 and less than or equal to 409.54); level 2 (a score greater than 409.54 and less than or equal to 484.14); level 3 (a score greater than 484.14 and less than or equal to 558.73); level 4 (a score greater than 558.73 and less than or equal to 633.33); level 5 (a score greater than 633.33 and less than or equal to 707.93); and level 6 (a score greater than 707.93). The Organization for Economic Cooperation and Development (OECD) average is the average of the national averages of the OECD member jurisdictions. Because of an error in printing the test booklets, the United States mean performance may be misestimated by approximately 1 score point. The impact is below one standard error. For details see appendix B. Standard error is noted by *s.e.*

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2006.

Page intentionally left blank

Appendix D: PISA 2006 Expert Panelists

Rodger Bybee
Executive Director
Biological Sciences Curriculum Study
Colorado Springs, CO

John Easton
Executive Director
Consortium on Chicago School Research
Chicago, IL

Thomas Hoffer
Senior Research Scientist
National Opinion Research Center
Chicago, IL

Stan Metzenberg
Associate Professor, Department of Biology
California State University at Northridge
Northridge, CA

Brett Moulding
State Science Specialist
Utah State Office of Education
Salt Lake City, UT

Aaron Pallas
Professor of Sociology and Education
Columbia University
New York, NY

Jo Ellen Roseman
Director
American Association for the Advancement of Science
Washington, DC

Gerald Wheeler
Executive Director
National Science Teachers Association
Arlington, VA