

## Appendix A

### Technical Procedures for the NAEP 2009 Science Assessment

This appendix provides an overview of some of the technical procedures for the NAEP 2009 science assessment. Information is included about the content of the assessment, school and student samples and participation, inclusion of students with disabilities and/or English language learners, analysis procedures, and interpretation of results. Additional technical information about NAEP assessments is available on the Web at <http://www.nces.ed.gov/nationsreportcard/tdw/>.

### Development of the Science Framework

The National Assessment Governing Board oversees the creation of the NAEP frameworks that describe the specific knowledge and skills that should be assessed in each subject. The frameworks also provide the theoretical basis for the assessment, direction for what types of items should be included, and how the items should be designed and scored. While the frameworks describe the general content and design of NAEP subject area assessments, the specifications provide the detailed information used by test developers for constructing the assessments and more detailed information in scoring. Both the *Science Framework for the 2009 National Assessment of Educational Progress* and *Science Assessment and Item Specifications for the 2009 National Assessment of Educational Progress* are available on the Governing Board's website at <http://www.nagb.org/publications/frameworks.htm>.

The 2009 NAEP science framework approved by the Governing Board replaces the framework used for the 1996, 2000, and 2005 science assessments. A variety of factors made it necessary to create a new framework to guide the assessment of science in 2009 and beyond: the publication of *National Standards* for science literacy, advances in both science and cognitive research, the growth in the prevalence of national and international science assessments, advances in innovative assessment approaches, and the need to advance the state of the art so that the widest possible range of students can be fairly assessed.

The development of the new science framework involved the critical input of hundreds of individuals across the country, including some of the nation's leading scientists, science educators, policymakers, and assessment experts. Under contract to the Governing Board, WestEd and the Council of Chief State School Officers (CCSSO) spent 18 months developing the framework; this process involved committees, regional hearings, and other public forums. The Governing Board also engaged an external review panel to evaluate the draft framework and convened a public hearing to receive additional input during the development process.

The frameworks for all main NAEP assessments are periodically updated or changed to reflect current curricula and standards. Whenever changes are made to a subject framework, every effort is made to maintain the trend lines that permit the reporting of changes in student achievement over time. If, however, the changes made to an assessment are such that the results are not comparable to earlier assessments, a new trend line is started. The assessment resulting from the 2009 framework will start a new NAEP science trend.

### Framework Dimensions

Since science consists of both knowing and doing, the design of the NAEP science assessment is guided by the framework's descriptions of the science content and practices to be assessed. Students are expected to have learned science content comprised of the facts, concepts, laws, principles, and theories that have been verified by the community of scientists, as well as understand how scientists gather, organize, and evaluate empirical evidence. Each question in the 2009 science assessment was classified based on two dimensions: *science content* and *science practices*. By considering these two dimensions for each question, the framework ensures that NAEP assesses an appropriate balance of content along with a variety of ways of doing science.

## SCIENCE CONTENT

The 2009 framework organizes science content into three broad content areas reflecting the science curriculum students are generally exposed to across the K–12 curriculum, including physical science, life science, and Earth and space sciences.

- Physical science includes concepts related to properties and changes of matter, forms of energy, energy transfer and conservation, position and motion of objects, and forces affecting motion.
- Life science includes concepts related to organization and development, matter and energy transformations, interdependence, heredity and reproduction, and evolution and diversity.
- Earth and space sciences include concepts related to objects in the universe, the history of the Earth, properties of Earth materials, tectonics, energy in Earth systems, climate and weather, and biogeochemical cycles.

Because of differences in curricular emphasis, the proportion of assessment time devoted to each science practices area varies by grade. The distribution of items across the four science practices is as follows. At grade 4, assessment time is distributed approximately evenly among Physical Science, Life Science, and Earth and Space Sciences. At grade 8, there is a somewhat greater emphasis on Earth and Space Sciences. At grade 12, the balance shifts toward Physical Science and Life Science, with less emphasis on Earth and Space Sciences.

## SCIENCE PRACTICES

In addition to the science content, the framework assesses student understanding of how scientific knowledge is used by measuring what students are able to do with the science content. Four science practices describe how science knowledge is used—identifying science principles, using science principles, using scientific inquiry, and using technological design.

- Identifying science principles focuses on students' ability to recognize, recall, define, relate, and represent basic science principles in each of the three content areas.
- Using science principles focuses on the importance of science knowledge in making accurate predictions about and explaining observations of the natural world.
- Using scientific inquiry focuses on designing, critiquing, and evaluating scientific investigations; identifying patterns in data; and using empirical evidence to validate or criticize conclusions.
- Using technological design focuses on the systematic process of applying science knowledge and skills to propose or critique solutions to real world problems, identify trade-offs, and anticipate effects of technological design decisions.

The distribution of items across the four science practices is as follows: Identifying Science Principles and Using Science Principles (combined), 60 percent; Using Scientific Inquiry, 30 percent; and Using Technological Design, 10 percent. From grade 4 to grade 8 to grade 12, the emphasis on Using Science Principles increases and the emphasis on Identifying Science Principles decreases.

**Table A-1. Percentage distribution of target and actual assessment time in NAEP science, by field of science and grade: 2009**

Grade	Physical science		Life science		Earth and space sciences	
	Target	Actual	Target	Actual	Target	Actual
<b>Grade 4</b>	33	33	33	34	33	33
<b>Grade 8</b>	30	26	30	33	40	41
<b>Grade 12</b>	38	38	38	36	25	26

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

## Content of the 2009 Science Assessment

Each NAEP assessment contains two major components: subject-specific cognitive items that measure the achievement of students in an academic subject; and background items that collect information from students, teachers, and school administrators about variables that are related to student achievement. Both the cognitive and background items are developed through a process that includes reviews by external advisory groups and field testing. Results from the cognitive items provide information about what students know and can do in a subject area. Information from the background items gives context to NAEP results and allows researchers to track factors associated with academic achievement.

The 2009 science assessment was made up of 143 cognitive questions at fourth grade, 162 questions at eighth grade, and 179 questions at twelfth grade. Students spent about one-half of the assessment time responding to multiple-choice questions and one-half responding to two types of constructed-response questions. Short constructed-response questions required students to write a concise explanation for a given situation or result, illustrate with a brief example, or describe a quantitative relationship in response to the question provided. Extended constructed-response questions were generally multidimensional and required students to solve a problem by applying and integrating science concepts and required that students analyze a science situation and explain a concept. Table A-2 shows the number of cognitive items administered in 2009 by item format at each grade level.

**Table A-2 . Number of NAEP science questions at grades 4, 8, and 12, by question type: 2009**

Question type	Grade 4	Grade 8	Grade 12
<b>Total</b>	<b>143</b>	<b>162</b>	<b>179</b>
Multiple-choice	97	104	120
Short constructed-response	39	36	44
Extended constructed-response	7	22	15

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Cognitive Blocks:** The assessment design allowed for broad coverage at each grade of the three science content areas and four science practices, while minimizing the time burden for any one student. This was accomplished through the use of matrix sampling of items in which each student was required to take only a small portion of the entire pool of assessment questions.

The science item pool for each grade was divided up into subsets or "blocks." In 2009, there were a total of 9 cognitive blocks at fourth grade, 10 blocks at eighth grade, and 11 blocks at twelfth grade. Each science assessment booklet contained two separately timed 25-minute blocks. Each block contained between 14 and 18 questions, depending on the balance between multiple-choice and constructed-response questions.

The procedure used to create booklets ensured that each block was paired with every other block. In addition, the procedure controlled for possible block-position effects across the set of booklets by balancing the order of the blocks within booklets. The booklets were cycled through in such a way that each booklet was used approximately an equal number of times across the entire assessment, while no more than a few students in any given assessment session received the same booklet.

Sample released questions at all three grade levels can be viewed at the NAEP website at <http://nces.ed.gov/nationsreportcard/itmrls/>. Items may be sorted by difficulty and question type.

## NAEP Samples

NAEP assesses representative samples of students rather than the entire population of students. The sample selection process utilizes a probability sample design in which each school and each student has a known probability of being selected (the probabilities are proportionate to the estimated number of students in the grade of an assessed school). Samples are selected according to a multistage design, with students drawn from within sampled public and private schools nationwide.

The 2005-06 Common Core of Data (CCD) file, a comprehensive list of operating public schools in each jurisdiction that is compiled each school year by the National Center for Education Statistics, served as the sampling frame for the selection of public schools in each state/jurisdiction. The sample of students in districts participating in the Trial Urban District Assessment (TUDA) represents an augmentation of the sample of students selected as part of the state samples. All students at more local geographic sampling levels also make up part of the broader samples. For example, the TUDA samples are included as part of the corresponding state samples, and the state samples are included as part of the national sample.

The 2005-06 Private School Survey (PSS), a mail survey of all U.S. private schools carried out biennially by the Census Bureau under contract to NCES, served as the sampling frame for private schools. While state and district results are based on samples of public schools only, the national results are based on the combined samples of public and private schools. Although information about the combined public and private school national samples is provided here for context, performance results in the State Report Generator and the District Report Generator are for public school students only.

Table A-3 shows the target populations and sample sizes in 2009 for the nation and participating states and jurisdictions at grades 4 and 8. Table A-4 shows the same information for participating urban districts for grades 4 and 8. Note that a total of 5 states/jurisdictions did not participate in the state level 2009 NAEP science assessment at both grades 4 and 8. For those 5 states/jurisdictions, a sample of 300–400 students was selected per state to ensure that these states are sufficiently represented in the national science sample.

Because each school that participated in the assessment, and each student assessed, represents only a portion of the larger population of interest, the results are weighted to make appropriate inferences between the student samples and the respective populations from which they are drawn. Sampling weights are adjusted for the disproportionate representation of some groups in the selected sample. This includes oversampling of schools with high concentrations of students from certain racial/ethnic groups and the lower sampling rates of students who attend very small schools.

**Table A-3. Student sample sizes and target populations in NAEP science at grades 4 and 8, by state/jurisdiction: 2009**

State/jurisdiction	Grade 4		Grade 8	
	Sample size	Target population	Sample size	Target population
<b>Nation</b>	<b>160,000</b>	<b>3,824,000</b>	<b>154,700</b>	<b>3,843,000</b>
Public	155,000	3,485,000	149,900	3,504,000
Private	2,800	330,000	3,100	331,000
Alabama	2,700	56,000	2,700	53,000
Alaska	—	—	—	—
Arizona	3,100	78,000	2,900	73,000
Arkansas	2,900	37,000	2,700	33,000
California	7,600	444,000	7,400	469,000
Colorado	2,800	55,000	2,800	54,000
Connecticut	2,800	41,000	2,800	42,000
Delaware	2,800	9,000	2,800	9,000
Florida	4,800	186,000	4,500	180,000
Georgia	4,100	117,000	3,600	109,000
Hawaii	2,800	13,000	2,900	12,000
Idaho	3,100	21,000	3,000	20,000
Illinois	4,300	146,000	4,200	154,000
Indiana	2,800	76,000	2,800	77,000
Iowa	2,900	34,000	2,700	33,000
Kansas	—	—	—	—
Kentucky	3,900	48,000	3,800	47,000
Louisiana	2,900	53,000	2,600	45,000
Maine	2,700	13,000	2,700	14,000
Maryland	3,600	56,000	3,500	58,000
Massachusetts	3,900	71,000	3,800	72,000
Michigan	3,500	117,000	3,500	117,000
Minnesota	3,400	62,000	3,000	60,000
Mississippi	2,800	39,000	2,900	37,000
Missouri	2,800	63,000	2,800	64,000
Montana	2,700	10,000	2,700	11,000
Nebraska	—	—	—	—
Nevada	3,100	32,000	2,900	32,000
New Hampshire	2,700	15,000	2,600	15,000
New Jersey	2,900	102,000	2,800	100,000
New Mexico	2,900	25,000	2,600	23,000
New York	4,100	194,000	3,900	198,000
North Carolina	4,600	105,000	4,500	112,000
North Dakota	2,100	7,000	2,300	7,000
Ohio	3,600	126,000	3,700	129,000
Oklahoma	2,900	46,000	2,800	44,000
Oregon	2,900	39,000	2,900	42,000
Pennsylvania	3,700	121,000	3,700	127,000
Rhode Island	2,500	10,000	2,800	11,000
South Carolina	3,000	53,000	2,900	50,000
South Dakota	2,800	9,000	2,900	9,000
Tennessee	3,000	75,000	3,000	75,000
Texas	6,500	318,000	6,200	322,000
Utah	3,400	42,000	3,000	38,000
Vermont	—	—	—	—
Virginia	3,000	86,000	2,900	90,000
Washington	3,200	75,000	2,900	75,000
West Virginia	2,800	20,000	3,000	23,000
Wisconsin	3,900	59,000	3,600	61,000
Wyoming	2,000	7,000	1,900	6,000
Other jurisdictions				
BIE <sup>1</sup>	100	3,000	100	2,000
District of Columbia	—	—	—	—
DoDEA <sup>2</sup>	2,100	7,000	1,600	5,000

— Not available. The jurisdiction did not participate.

<sup>1</sup> Bureau of Indian Education.

<sup>2</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: The sample size is rounded to the nearest hundred. The target population is rounded to the nearest thousand. Data for BIE and DoDEA schools are counted in the overall nation total, but not in the nation (public) total. Data for the District of Columbia public schools are counted, along with the states, in nation (public). Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-4. Student sample sizes and target populations for Trial Urban District Assessment (TUDA) in science at grades 4 and 8, by urban district: 2009**

District	Grade 4		Grade 8	
	Sample size	Target population	Sample size	Target population
Atlanta	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA
Jefferson County (KY)	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA

NOTE: The sample size is rounded to the nearest hundred. The target population is rounded to the nearest thousand.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

# School and Student Participation

## National Participation

To ensure unbiased samples, the National Assessment Governing Board policy on reporting requires that participation rates for original school samples be 70 percent or higher, for public and private schools respectively, to report national results separately for public and private schools. In instances where the original school participation rate falls below 85 percent, NCES statistical standards require that a nonresponse bias analysis be conducted to determine if the responding school sample is not representative of the population, thereby introducing the potential for nonresponse bias. The decision whether or not to report the results in a case where the response rate falls between 70 and 85 percent depends upon the results of this nonresponse bias analysis.

National school and student participation rates for the 2009 science assessment are presented in table A-5. Student-weighted school participation rates were 97 percent for grade 4 (100 percent for public schools and 73 percent for private schools), 97 percent for grade 8 (100 percent for public schools and 72 percent for private schools), and 83 percent for grade 12 (86 percent for public schools and 52 percent for private schools). Weighted student participation rates were 95 percent for grade 4 (95 percent for public schools and 96 percent for private schools), 93 percent for grade 8 (92 percent for public schools and 95 percent for private schools), and 80 percent for grade 12 (79 percent for public schools and 88 percent for private schools).

**Table A-5. National school and student participation rates in NAEP science, by grade and type of school: 2009**

Grade and type of school	School participation				Number of schools participating after substitution	Student participation	
	Student-weighted		School-weighted			Student-weighted percent	Number of students assessed
	Percent before substitution	Percent after substitution	Percent before substitution	Percent after substitution			
<b>Grade 4</b>							
<b>Nation</b>	<b>97</b>	<b>98</b>	<b>91</b>	<b>95</b>	<b>9,330</b>	<b>95</b>	<b>156,500</b>
Public	100	100	100	100	8,780	95	151,500
Private	73	85	68	80	370	96	2,800
<b>Grade 8</b>							
<b>Nation</b>	<b>97</b>	<b>98</b>	<b>87</b>	<b>92</b>	<b>6,920</b>	<b>93</b>	<b>151,100</b>
Public	100	100	100	100	6,440	92	146,300
Private	72	83	68	80	360	95	3,100
<b>Grade 12</b>							
<b>Nation</b>	<b>83</b>	<b>87</b>	<b>79</b>	<b>84</b>	<b>1,410</b>	<b>80</b>	<b>11,100</b>
Public	86	90	87	90	1,260	79	9,900
Private	52	66	57	69	160	88	1,200

NOTE: The national totals for schools include Department of Defense Education Activity (overseas and domestic schools) and Bureau of Indian Education schools, which are not included in either the public or private totals. The national totals for students include students in these schools. Columns of percentages have different denominators. The number of schools is rounded to the nearest ten. The number of students is rounded to the nearest hundred.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

The student-weighted school participation rates are calculated based on school sampling weights and grade-specific school enrollment figures. The denominator of the rate is the weighted total number of students represented by the initially selected schools that had eligible students enrolled. This includes both participating and nonparticipating schools. The numerator is the weighted total number of students represented by participating schools. This is calculated in two distinct ways: first, with participating schools defined as only the initially selected schools that participated in the assessment (which gives rise to the rate before substitution), and second, with all of the participating schools after substitution (giving the rate after substitution). On the other hand, the school-weighted school participation rates are calculated based only on the school sampling weights. They show the weighted total number of schools (either before or after substitution) divided by the weighted total number of schools in the initially selected sample.

## State and District Participation

Standards established by the Governing Board require that student-weighted school participation rates for the state and district samples need to be at least 85 percent for results to be reported. In 2009, all 47 states and jurisdictions and all 17 urban districts participating in the science assessment at grades 4 and 8 met this participation rate requirement (tables A-6 through A-8). Note that no school substitution was used for the state and district samples at grades 4 and 8.

**Table A-6. Public school and student participation rates in NAEP science at grade 4, by state/jurisdiction: 2009**

State/jurisdiction	School participation			Student participation	
	Student-weighted percent	School-weighted percent	Number of schools participating	Student-weighted percent	Number of students assessed
<b>Nation (public)</b>	<b>100</b>	<b>100</b>	<b>8,780</b>	<b>95</b>	<b>151,500</b>
Alabama	100	100	130	95	2,700
Alaska	—	—	—	—	—
Arizona	100	100	140	96	3,100
Arkansas	100	100	140	95	2,800
California	100	100	310	95	7,400
Colorado	100	100	150	95	2,700
Connecticut	100	100	130	94	2,700
Delaware	100	100	100	94	2,800
Florida	100	100	180	94	4,700
Georgia	100	100	170	96	4,000
Hawaii	100	100	140	95	2,800
Idaho	100	100	160	95	3,000
Illinois	100	100	220	95	4,100
Indiana	100	100	140	94	2,700
Iowa	100	100	180	95	2,800
Kansas	—	—	—	—	—
Kentucky	100	100	190	95	3,800
Louisiana	100	100	150	94	2,900
Maine	100	100	200	93	2,600
Maryland	99	99	200	95	3,500
Massachusetts	97	99	210	94	3,700
Michigan	100	100	190	94	3,400
Minnesota	100	99	170	95	3,300
Mississippi	100	100	130	95	2,800
Missouri	100	100	160	96	2,700
Montana	100	98	240	94	2,700
Nebraska	—	—	—	—	—
Nevada	100	100	130	95	3,000
New Hampshire	99	99	160	93	2,700
New Jersey	100	100	140	93	2,800
New Mexico	100	100	160	94	2,800
New York	100	100	180	93	4,000
North Carolina	100	100	190	95	4,500
North Dakota	100	100	240	96	2,000
Ohio	100	100	210	94	3,500
Oklahoma	100	100	180	96	2,800
Oregon	100	100	180	94	2,900
Pennsylvania	100	100	190	94	3,600
Rhode Island	100	100	150	95	2,500
South Carolina	100	100	130	95	2,900
South Dakota	100	100	290	96	2,700
Tennessee	100	100	140	94	2,900
Texas	100	100	270	95	6,300
Utah	100	100	150	94	3,300
Vermont	—	—	—	—	—
Virginia	100	100	130	95	2,900
Washington	100	100	160	94	3,100
West Virginia	100	100	200	94	2,800
Wisconsin	99	99	240	96	3,800
Wyoming	100	100	160	95	2,000
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	99	98	110	93	2,100

— Not available. The jurisdiction did not participate.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: The number of schools is rounded to the nearest ten. The number of students is rounded to the nearest hundred. The school participation rates are student-weighted percentages before substitution. Columns of percentages have different denominators. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-7. Public school and student participation rates in NAEP science at grade 8, by state/jurisdiction: 2009**

State/jurisdiction	School participation			Student participation	
	Student-weighted percent	School-weighted percent	Number of schools participating	Student-weighted percent	Number of students assessed
<b>Nation (public)</b>	<b>100</b>	<b>100</b>	<b>6,440</b>	<b>92</b>	<b>146,300</b>
Alabama	100	100	110	93	2,700
Alaska	—	—	—	—	—
Arizona	100	100	130	93	2,900
Arkansas	100	100	120	93	2,600
California	100	100	230	93	7,200
Colorado	100	100	120	92	2,800
Connecticut	100	100	110	91	2,800
Delaware	100	100	50	91	2,800
Florida	100	100	160	91	4,300
Georgia	100	100	120	94	3,500
Hawaii	100	100	70	91	2,800
Idaho	100	100	110	94	2,900
Illinois	100	100	200	94	4,200
Indiana	100	100	110	93	2,700
Iowa	100	100	130	93	2,700
Kansas	—	—	—	—	—
Kentucky	100	100	130	94	3,700
Louisiana	100	100	120	92	2,600
Maine	100	100	130	91	2,600
Maryland	100	100	130	92	3,400
Massachusetts	100	100	140	92	3,700
Michigan	100	100	150	92	3,400
Minnesota	100	100	140	93	3,000
Mississippi	100	100	120	93	2,800
Missouri	100	100	130	93	2,700
Montana	100	98	180	92	2,600
Nebraska	—	—	—	—	—
Nevada	100	100	90	91	2,900
New Hampshire	96	96	90	89	2,500
New Jersey	100	100	110	93	2,800
New Mexico	100	100	100	89	2,500
New York	97	98	150	90	3,800
North Carolina	100	100	150	92	4,400
North Dakota	100	100	180	95	2,200
Ohio	100	100	190	93	3,500
Oklahoma	100	100	150	93	2,700
Oregon	100	100	130	92	2,800
Pennsylvania	100	100	150	92	3,600
Rhode Island	100	100	60	93	2,700
South Carolina	100	100	110	94	2,800
South Dakota	100	100	220	95	2,800
Tennessee	100	100	120	93	3,000
Texas	99	100	170	92	5,900
Utah	100	100	110	92	2,900
Vermont	—	—	—	—	—
Virginia	100	100	110	93	2,800
Washington	100	100	130	92	2,800
West Virginia	100	100	120	93	2,900
Wisconsin	99	99	170	93	3,500
Wyoming	100	100	90	91	1,900
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	99	97	60	93	1,600

— Not available. The jurisdiction did not participate.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: The number of schools is rounded to the nearest ten. The number of students is rounded to the nearest hundred. The school participation rates are student-weighted percentages before substitution. Columns of percentages have different denominators. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-8. Public school and student participation rates for Trial Urban District Assessment (TUDA) in science, by grade and urban district: 2009**

Grade and district	School participation		Student participation	
	Student-weighted percent	Number of schools participating	Student-weighted percent	Number of students assessed
<b>Grade 4</b>				
Atlanta	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA
Jefferson County, KY	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA
<b>Grade 8</b>				
Atlanta	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA
Jefferson County, KY	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA

NOTE: The number of schools is rounded to the nearest ten. The number of students is rounded to the nearest hundred. The school participation rates are student-weighted percentages before substitution. The percentages for school-weighted and student-weighted school participation were both at 100 percent for the participating districts in 2009.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

## Inclusion of Students With Disabilities and/or English Language Learners

Testing all sampled students is the best way for NAEP to ensure that results are as representative as possible of the performance of students in the nation and in participating states/jurisdictions and districts. NAEP has always endeavored to assess all students selected as a part of its sampling process, including students who are classified by their schools as students with disabilities (SD) and/or as English language learners (ELL).

### Accommodations

Prior to 1996, no testing accommodations were provided to students taking the NAEP assessments, resulting in the exclusion of students who could not be assessed without them. As the number of identified students with disabilities and English language learners increased over the years, the exclusion of those needing accommodations to participate in NAEP threatened the stability of trend lines (excluding more students in one assessment year than in another might lead to apparent rather than real differences), and threatened to compromise NAEP samples as optimally representative of target populations. Therefore, administration procedures allowing for many of the same testing accommodations provided on state and district assessments (e.g., extra testing time or individual rather than group administration) were introduced in 1996 for national NAEP assessments and in 2000 for NAEP state assessments.

The percentages of SD/ELL students assessed with the available accommodations in 2009 are presented in table A-9. Students assessed with accommodations typically received some combination of accommodations. For example, students assessed in small groups (as compared with standard NAEP sessions of about 30 students) were also usually given extended time and are included in counts for both groups in table A-9.

**Table A-9. Percentage of fourth-, eighth-, and twelfth-grade public and nonpublic school students with disabilities (SD) and/or English language learners (ELL) assessed in NAEP science with accommodations, as a percent of all students, by SD/ELL category and type of primary accommodation: 2009**

Type of accommodation	Grade 4			Grade 8			Grade 12		
	SD and/or ELL	SD	ELL	SD and/or ELL	SD	ELL	SD and/or ELL	SD	ELL
Bilingual book	0.4	#	0.4	0.1	#	0.1	—	—	—
Bilingual dictionary	0.5	#	0.5	0.5	#	0.5	0.4	#	0.4
Large-print book	0.1	0.1	#	#	#	#	#	#	#
Extended time	9.5	7.2	3.0	8.6	7.4	1.6	6.2	5.6	0.9
Read aloud	6.2	5.1	1.6	4.5	4.1	0.7	1.9	1.8	0.2
Small group	8.5	6.8	2.3	7.4	6.7	1.1	4.9	4.5	0.6
One-on-one	0.6	0.6	0.1	0.4	0.3	#	0.3	0.3	#
Scribe/computer	0.5	0.5	#	0.3	0.3	#	0.3	0.3	#
Breaks	3.8	3.2	1.0	2.4	2.1	0.5	1.3	1.1	0.3
Magnifying device	#	#	#	#	#	#	#	#	#
School staff administers	0.7	0.6	0.2	0.4	0.3	0.1	0.2	0.1	#
Directions read aloud in Spanish	0.3	#	0.3	0.2	0.1	0.2	0.1	#	0.1
Braille version of the text	#	#	#	#	#	#	#	#	#
Other	1.1	1.0	0.2	1.1	1.0	0.1	0.8	0.8	#

— Not available.

# Rounds to zero.

NOTE: Students identified as both SD and ELL were counted only once under the combined SD and/or ELL category, but were counted separately under the SD and ELL categories.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

## Exclusion Rates

Even with the availability of accommodations, some students are excluded from the NAEP assessments by their schools. The decision to exclude any student is made by school staff who, using NAEP guidelines and each student's Individualized Education Program (IEP), decide whether the student can meaningfully be assessed.

Jurisdictions vary in their proportions of special-needs students. These variations, as well as differences in policies and practices regarding the identification and inclusion of special-needs students, lead to differences in exclusion and accommodation rates. These differences should be considered when comparing student performance over time and across jurisdictions. While the effect of exclusion is not precisely known, the validity of comparisons of performance results could be affected if exclusion rates are comparatively high or vary widely over time.

**National Exclusion Rates (public and nonpublic school students):** In 2009, twenty-one percent of students at grade four, 17 percent at grade eight, and 13 percent at grade twelve were identified as SD and/or ELL, with 2 percent excluded at grade four, 2 percent excluded at grade eight, and 3 percent excluded at grade twelve (table A-10). The percentages of SD and/or ELL students assessed with accommodations in 2009 ranged from 11 percent at grade four to 7 percent at grade twelve. The proportions of SD and/or ELL students excluded and assessed with and without accommodations as a percentage of students identified are provided in table A-11.

**Table A-10. Percentage of fourth-, eighth-, and twelfth-grade public and nonpublic school students with disabilities (SD) and/or English language learners (ELL) excluded and assessed in NAEP science, as a percent of all students: 2009**

SD/ELL category	Grade 4	Grade 8	Grade 12
<b>SD and/or ELL</b>			
Identified	21	17	13
Excluded	2	2	3
Assessed	19	15	11
Without accommodations	8	5	4
With accommodations	11	10	7
<b>SD</b>			
Identified	13	12	11
Excluded	2	2	2
Assessed	11	11	8
Without accommodations	3	2	2
With accommodations	8	9	6
<b>ELL</b>			
Identified	10	5	3
Excluded	1	#	#
Assessed	9	5	3
Without accommodations	5	3	2
With accommodations	4	2	1

# Rounds to zero.

NOTE: Students identified as both SD and ELL were counted only once under the combined SD and/or ELL category, but were counted separately under the SD and ELL categories. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-11. Percentage of fourth-, eighth-, and twelfth-grade public and nonpublic school students identified as students with disabilities (SD) and/or English language learners (ELL) excluded and assessed in NAEP science, as a percentage of all identified SD and/or ELL students, by grade and SD/ELL category: 2009**

Grade and SD/ELL category	Percentage of identified SD and/or ELL students			
	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>Grade 4</b>				
SD and/or ELL	9	91	39	52
SD	12	88	23	64
ELL	7	93	57	37
<b>Grade 8</b>				
SD and/or ELL	11	89	30	59
SD	13	87	17	70
ELL	9	91	57	34
<b>Grade 12</b>				
SD and/or ELL	19	81	28	52
SD	23	77	19	58
ELL	10	90	57	33

NOTE: Students identified as both SD and ELL were counted only once under the combined SD and/or ELL category, but were counted separately under the SD and ELL categories. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**State Exclusion Rates (public school students only):** The state percentages of fourth-graders identified as SD and/or ELL in 2009 ranged from 10 to 36 percent, and exclusion rates ranged from 1 to 3 percent (table A-12).

The state percentages of eighth-graders identified as SD and/or ELL in 2009 ranged from 10 to 25 percent, and exclusion rates ranged from 1 to 4 percent (table A-13).

Rates by state are reported separately for SD and ELL students at each grade in tables A-14 through A-17. Rates are also reported as the percentage of SD and/or ELL students identified in each state in tables A-18 through A-19.

**Table A-12. Percentage of fourth-grade public school students identified as students with disabilities and/or English language learners excluded and assessed in NAEP science, as a percentage of all students, by state/jurisdiction: 2009**

State/jurisdiction	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>Nation (public)</b>	<b>23</b>	<b>2</b>	<b>20</b>	<b>9</b>	<b>12</b>
Alabama	12	1	11	7	4
Alaska	—	—	—	—	—
Arizona	26	2	24	11	13
Arkansas	17	1	16	3	13
California	36	2	33	27	6
Colorado	21	1	20	6	14
Connecticut	18	2	16	2	14
Delaware	18	2	17	2	14
Florida	23	2	22	4	18
Georgia	14	1	13	4	9
Hawaii	20	1	19	6	13
Idaho	15	2	13	6	8
Illinois	22	2	19	5	14
Indiana	19	2	17	6	12
Iowa	18	2	17	3	13
Kansas	—	—	—	—	—
Kentucky	17	2	15	5	10
Louisiana	22	1	20	4	16
Maine	20	1	18	4	14
Maryland	19	3	16	2	14
Massachusetts	24	3	21	7	14
Michigan	17	2	15	6	8
Minnesota	21	3	19	8	10
Mississippi	10	1	9	3	6
Missouri	16	2	14	5	9
Montana	14	1	13	4	9
Nebraska	—	—	—	—	—
Nevada	30	2	28	11	17
New Hampshire	21	2	19	4	15
New Jersey	19	2	17	2	15
New Mexico	26	2	24	9	15
New York	22	1	21	1	20
North Carolina	19	2	18	5	12
North Dakota	17	3	15	4	11
Ohio	16	2	14	3	11
Oklahoma	19	3	15	5	10
Oregon	26	3	23	8	15
Pennsylvania	18	1	17	4	12
Rhode Island	22	2	20	5	15
South Carolina	19	1	18	8	10
South Dakota	16	2	14	7	8
Tennessee	16	2	14	3	11
Texas	29	3	26	16	9
Utah	19	2	17	6	11
Vermont	—	—	—	—	—
Virginia	20	2	18	5	13
Washington	21	2	19	8	11
West Virginia	17	2	16	7	9
Wisconsin	20	2	18	3	15
Wyoming	18	1	17	4	13
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	18	2	16	6	10

— Not available. The jurisdiction did not participate.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-13. Percentage of eighth-grade public school students identified as students with disabilities and/or English language learners excluded and assessed in NAEP science, as a percentage of all students, by state/jurisdiction: 2009**

State/jurisdiction	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>Nation (public)</b>	<b>18</b>	<b>2</b>	<b>16</b>	<b>5</b>	<b>10</b>
Alabama	11	1	10	7	3
Alaska	—	—	—	—	—
Arizona	16	2	14	5	9
Arkansas	16	1	14	3	11
California	25	2	24	18	6
Colorado	17	1	15	5	11
Connecticut	16	2	14	3	11
Delaware	17	1	16	2	14
Florida	19	2	17	1	16
Georgia	13	1	12	2	10
Hawaii	18	2	17	6	10
Idaho	12	1	11	4	7
Illinois	16	1	15	3	12
Indiana	16	2	14	3	12
Iowa	16	1	15	2	12
Kansas	—	—	—	—	—
Kentucky	13	2	10	2	9
Louisiana	16	1	14	2	12
Maine	19	2	17	3	14
Maryland	14	3	12	1	11
Massachusetts	21	4	17	3	14
Michigan	15	2	12	3	9
Minnesota	17	2	15	6	9
Mississippi	10	1	9	2	7
Missouri	14	1	12	3	10
Montana	14	2	12	3	9
Nebraska	—	—	—	—	—
Nevada	17	1	16	5	10
New Hampshire	21	2	19	5	14
New Jersey	18	2	16	1	14
New Mexico	21	3	18	8	11
New York	20	2	18	1	17
North Carolina	17	2	15	3	13
North Dakota	16	4	12	3	9
Ohio	15	2	13	1	12
Oklahoma	18	3	14	4	10
Oregon	18	2	16	8	9
Pennsylvania	19	2	17	2	15
Rhode Island	21	3	18	4	14
South Carolina	16	2	14	5	9
South Dakota	12	1	10	3	7
Tennessee	12	2	11	1	9
Texas	17	4	14	7	7
Utah	14	2	12	4	8
Vermont	—	—	—	—	—
Virginia	17	2	15	4	11
Washington	14	2	12	4	7
West Virginia	15	2	14	4	10
Wisconsin	18	2	16	3	13
Wyoming	15	2	13	3	10
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	13	2	11	3	7

— Not available. The jurisdiction did not participate.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-14. Percentage of fourth-grade public school students identified as students with disabilities excluded and assessed in NAEP science, by state/jurisdiction: 2009**

State/jurisdiction	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>Nation (public)</b>	<b>13</b>	<b>2</b>	<b>12</b>	<b>3</b>	<b>9</b>
Alabama	10	1	9	6	4
Alaska	—	—	—	—	—
Arizona	13	2	12	5	7
Arkansas	12	1	11	2	9
California	10	2	7	3	4
Colorado	11	1	10	2	8
Connecticut	13	2	12	2	10
Delaware	15	2	14	2	12
Florida	17	1	15	3	12
Georgia	10	1	10	3	7
Hawaii	10	1	10	2	8
Idaho	10	1	9	3	6
Illinois	15	1	14	3	10
Indiana	16	2	14	5	9
Iowa	14	1	13	2	10
Kansas	—	—	—	—	—
Kentucky	15	2	13	5	9
Louisiana	20	1	18	4	15
Maine	18	1	17	3	14
Maryland	14	2	12	2	10
Massachusetts	19	3	15	2	13
Michigan	14	2	12	4	8
Minnesota	14	2	12	5	8
Mississippi	9	1	9	3	6
Missouri	14	2	13	4	8
Montana	12	1	10	3	8
Nebraska	—	—	—	—	—
Nevada	12	2	10	3	6
New Hampshire	18	2	17	3	14
New Jersey	16	1	14	2	12
New Mexico	13	2	11	3	8
New York	16	1	15	1	14
North Carolina	15	2	13	4	9
North Dakota	16	3	13	4	10
Ohio	14	2	12	2	10
Oklahoma	15	3	12	3	9
Oregon	16	3	13	5	8
Pennsylvania	15	1	14	4	11
Rhode Island	17	2	16	3	13
South Carolina	14	1	13	6	8
South Dakota	15	2	13	6	7
Tennessee	14	2	12	3	9
Texas	10	2	8	2	5
Utah	12	2	10	4	7
Vermont	—	—	—	—	—
Virginia	14	1	13	3	10
Washington	12	2	11	3	7
West Virginia	17	2	16	7	9
Wisconsin	15	2	13	2	11
Wyoming	16	1	14	3	11
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	12	1	11	3	8

— Not available. The jurisdiction did not participate.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-15. Percentage of eighth-grade public school students identified as students with disabilities excluded and assessed in NAEP science, by state/jurisdiction: 2009**

State/jurisdiction	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>Nation (public)</b>	<b>13</b>	<b>2</b>	<b>11</b>	<b>2</b>	<b>9</b>
Alabama	10	1	9	6	3
Alaska	—	—	—	—	—
Arizona	12	2	10	2	7
Arkansas	12	1	11	2	9
California	9	1	8	3	5
Colorado	11	1	9	1	8
Connecticut	13	1	12	2	10
Delaware	15	1	14	1	13
Florida	15	1	14	1	12
Georgia	11	1	10	2	8
Hawaii	12	1	11	3	8
Idaho	9	1	8	3	5
Illinois	14	1	13	2	11
Indiana	14	2	12	1	10
Iowa	14	1	13	1	12
Kansas	—	—	—	—	—
Kentucky	12	2	9	1	8
Louisiana	15	1	13	2	12
Maine	17	2	16	3	13
Maryland	12	2	10	1	9
Massachusetts	19	3	15	2	13
Michigan	13	2	10	2	8
Minnesota	12	2	11	3	8
Mississippi	9	1	8	1	7
Missouri	13	1	12	3	9
Montana	12	2	10	1	9
Nebraska	—	—	—	—	—
Nevada	11	1	10	2	8
New Hampshire	20	2	18	5	13
New Jersey	16	2	14	1	13
New Mexico	13	3	10	3	7
New York	16	1	15	1	14
North Carolina	12	1	11	1	10
North Dakota	15	4	11	3	9
Ohio	15	2	12	1	11
Oklahoma	15	3	12	2	10
Oregon	13	2	11	5	7
Pennsylvania	17	2	16	2	14
Rhode Island	18	2	16	4	12
South Carolina	14	2	12	4	8
South Dakota	10	1	9	2	7
Tennessee	12	2	10	1	9
Texas	12	3	9	3	6
Utah	10	2	8	2	7
Vermont	—	—	—	—	—
Virginia	14	2	12	3	9
Washington	11	2	9	3	6
West Virginia	15	2	13	4	10
Wisconsin	14	2	12	2	10
Wyoming	14	1	12	3	10
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	8	1	8	1	6

— Not available. The jurisdiction did not participate.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-16. Percentage of fourth-grade public school students identified as English language learners excluded and assessed in NAEP science, by state/jurisdiction: 2009**

State/jurisdiction	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>Nation (public)</b>	<b>10</b>	<b>1</b>	<b>10</b>	<b>6</b>	<b>4</b>
Alabama	2	#	2	2	#
Alaska	—	—	—	—	—
Arizona	15	1	14	7	7
Arkansas	6	#	6	1	4
California	30	1	29	25	3
Colorado	11	#	10	4	6
Connecticut	6	1	5	1	4
Delaware	4	#	4	1	3
Florida	8	1	7	#	7
Georgia	4	#	4	1	3
Hawaii	10	1	10	4	6
Idaho	5	#	5	3	2
Illinois	8	1	7	2	5
Indiana	5	1	4	1	3
Iowa	5	#	4	1	3
Kansas	—	—	—	—	—
Kentucky	2	#	2	1	1
Louisiana	2	#	2	1	2
Maine	1	#	1	1	1
Maryland	6	1	5	#	5
Massachusetts	7	1	6	4	2
Michigan	4	#	3	3	1
Minnesota	8	1	7	4	3
Mississippi	1	#	1	#	1
Missouri	2	#	2	1	1
Montana	3	#	3	2	2
Nebraska	—	—	—	—	—
Nevada	20	1	20	8	12
New Hampshire	3	#	3	1	2
New Jersey	4	1	3	#	3
New Mexico	16	1	15	6	9
New York	8	1	7	#	7
North Carolina	6	#	6	2	4
North Dakota	2	#	1	#	1
Ohio	3	#	2	1	2
Oklahoma	4	1	4	2	2
Oregon	12	1	11	4	7
Pennsylvania	3	#	3	#	2
Rhode Island	6	1	6	2	3
South Carolina	5	#	5	3	2
South Dakota	2	#	2	1	1
Tennessee	3	#	2	#	2
Texas	21	2	19	15	5
Utah	9	1	8	3	5
Vermont	—	—	—	—	—
Virginia	6	1	6	2	4
Washington	10	1	9	4	5
West Virginia	#	#	#	#	#
Wisconsin	7	1	6	1	5
Wyoming	3	#	3	#	2
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	7	1	5	3	2

— Not available. The jurisdiction did not participate.

# Rounds to zero.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-17. Percentage of eighth-grade public school students identified as English language learners excluded and assessed in NAEP science, by state/jurisdiction: 2009**

State/jurisdiction	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>Nation (public)</b>	<b>6</b>	<b>1</b>	<b>5</b>	<b>3</b>	<b>2</b>
Alabama	1	#	1	1	#
Alaska	—	—	—	—	—
Arizona	6	1	6	3	3
Arkansas	4	#	4	1	3
California	20	1	19	16	3
Colorado	7	#	7	3	3
Connecticut	4	1	3	1	2
Delaware	2	#	2	#	2
Florida	5	1	4	#	4
Georgia	2	#	2	#	1
Hawaii	7	1	6	3	3
Idaho	4	#	4	2	2
Illinois	3	1	3	1	2
Indiana	3	#	3	1	1
Iowa	2	#	2	1	1
Kansas	—	—	—	—	—
Kentucky	1	#	1	#	1
Louisiana	1	#	1	#	1
Maine	2	#	2	1	1
Maryland	2	#	2	#	2
Massachusetts	3	1	2	1	1
Michigan	2	#	2	2	#
Minnesota	6	1	5	4	1
Mississippi	1	#	1	#	#
Missouri	1	#	1	#	#
Montana	3	#	3	2	1
Nebraska	—	—	—	—	—
Nevada	8	#	8	4	4
New Hampshire	1	#	1	1	1
New Jersey	3	1	2	#	2
New Mexico	11	1	10	5	5
New York	5	1	4	#	4
North Carolina	5	#	5	2	3
North Dakota	2	1	1	1	#
Ohio	1	#	1	#	#
Oklahoma	3	#	3	2	1
Oregon	6	#	6	3	3
Pennsylvania	2	#	2	1	1
Rhode Island	3	1	2	1	1
South Carolina	3	#	3	1	2
South Dakota	1	#	1	1	#
Tennessee	1	#	1	#	1
Texas	7	1	6	4	1
Utah	5	#	4	2	2
Vermont	—	—	—	—	—
Virginia	3	#	3	1	2
Washington	4	#	3	2	2
West Virginia	1	#	1	#	#
Wisconsin	4	1	4	1	3
Wyoming	1	#	1	#	1
Other jurisdictions					
District of Columbia	—	—	—	—	—
DoDEA <sup>1</sup>	5	1	4	2	1

— Not available. The jurisdiction did not participate.

# Rounds to zero.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-18. Percentage of fourth-grade public school students identified as students with disabilities (SD) and/or English language learners (ELL) excluded and assessed in NAEP science, as a percentage of all identified SD and/or ELL students, by state/jurisdiction: 2009**

State/jurisdiction	Percentage of identified SD and/or ELL students											
	SD and/or ELL				SD				ELL			
	Excluded	Assessed	Assessed without accom- modations	Assessed with accom- modations	Excluded	Assessed	Assessed without accom- modations	Assessed with accom- modations	Excluded	Assessed	Assessed without accom- modations	Assessed with accom- modations
<b>Nation (public)</b>	<b>9</b>	<b>91</b>	<b>39</b>	<b>52</b>	<b>13</b>	<b>87</b>	<b>23</b>	<b>64</b>	<b>7</b>	<b>93</b>	<b>57</b>	<b>37</b>
Alabama	8	92	60	31	9	91	56	36	5	95	82	14
Alaska	—	—	—	—	—	—	—	—	—	—	—	—
Arizona	7	93	43	51	12	88	34	54	4	96	49	48
Arkansas	8	92	19	73	10	90	16	73	7	93	23	71
California	7	93	76	17	23	77	32	46	4	96	85	11
Colorado	7	93	29	64	11	89	15	74	3	97	42	55
Connecticut	13	87	11	76	14	86	11	75	13	87	10	77
Delaware	9	91	13	78	10	90	11	78	7	93	18	75
Florida	7	93	15	77	8	92	19	73	8	92	4	88
Georgia	6	94	30	64	7	93	31	62	2	98	28	70
Hawaii	7	93	28	65	7	93	16	76	8	92	37	55
Idaho	11	89	37	52	14	86	27	59	8	92	56	35
Illinois	10	90	23	67	7	93	23	69	16	84	21	63
Indiana	11	89	29	61	11	89	31	58	15	85	17	69
Iowa	9	91	19	72	10	90	15	74	3	97	29	68
Kansas	—	—	—	—	—	—	—	—	—	—	—	—
Kentucky	12	88	30	57	12	88	30	58	19	81	29	52
Louisiana	7	93	19	74	7	93	18	75	7	93	27	66
Maine	7	93	20	73	8	92	18	74	3	97	44	53
Maryland	15	85	12	73	17	83	14	69	16	84	6	78
Massachusetts	14	86	28	58	17	83	13	70	11	89	61	28
Michigan	13	87	37	50	15	85	28	57	8	92	72	20
Minnesota	12	88	39	49	14	86	34	52	13	87	44	43
Mississippi	8	92	33	59	8	92	33	59	8	92	39	54
Missouri	11	89	31	58	12	88	30	57	6	94	29	65
Montana	11	89	29	61	13	87	23	65	2	98	50	49
Nebraska	—	—	—	—	—	—	—	—	—	—	—	—
Nevada	8	92	36	56	17	83	28	55	5	95	38	58
New Hampshire	8	92	19	73	9	91	17	74	3	97	25	72
New Jersey	9	91	11	80	8	92	12	80	15	85	4	81
New Mexico	8	92	33	58	14	86	21	66	6	94	39	55
New York	7	93	6	88	5	95	7	88	9	91	2	89
North Carolina	10	90	27	63	12	88	26	62	5	95	29	66
North Dakota	16	84	23	61	17	83	22	61	21	79	25	54
Ohio	11	89	18	71	13	87	17	70	13	87	19	68
Oklahoma	19	81	28	54	21	79	22	58	14	86	48	38
Oregon	11	89	32	57	17	83	29	54	6	94	33	61
Pennsylvania	8	92	22	69	8	92	23	69	6	94	16	78
Rhode Island	10	90	23	67	10	90	16	74	11	89	39	50
South Carolina	7	93	43	51	8	92	39	53	3	97	51	46
South Dakota	11	89	41	48	12	88	40	49	10	90	46	44
Tennessee	10	90	22	68	12	88	22	66	2	98	15	83
Texas	11	89	57	32	24	76	22	54	7	93	70	23
Utah	11	89	33	56	15	85	29	57	7	93	33	59
Vermont	—	—	—	—	—	—	—	—	—	—	—	—
Virginia	9	91	24	67	10	90	22	68	8	92	26	66
Washington	11	89	37	52	14	86	27	58	7	93	44	49
West Virginia	9	91	39	52	9	91	39	52	#	100	51	49
Wisconsin	10	90	16	74	12	88	14	74	9	91	16	75
Wyoming	8	92	22	71	9	91	22	70	2	98	19	79
Other jurisdictions												
District of Columbia	—	—	—	—	—	—	—	—	—	—	—	—
DoDEA <sup>1</sup>	11	89	33	56	9	91	25	67	17	83	46	37

— Not available. The jurisdiction did not participate.

# Rounds to zero.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Students identified as both SD and ELL were counted only once under the combined SD and/or ELL category, but were counted separately under the SD and ELL categories. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-19. Percentage of eighth-grade public school students identified as students with disabilities (SD) and/or English language learners (ELL) excluded and assessed in NAEP science, as a percentage of all identified SD and/or ELL students, by state/jurisdiction: 2009**

State/jurisdiction	Percentage of identified SD and/or ELL students											
	SD and/or ELL				SD				ELL			
	Excluded	Assessed	Assessed without accom- modations	Assessed with accom- modations	Excluded	Assessed	Assessed without accom- modations	Assessed with accom- modations	Excluded	Assessed	Assessed without accom- modations	Assessed with accom- modations
<b>Nation (public)</b>	<b>11</b>	<b>89</b>	<b>30</b>	<b>58</b>	<b>14</b>	<b>86</b>	<b>17</b>	<b>70</b>	<b>9</b>	<b>91</b>	<b>56</b>	<b>35</b>
Alabama	12	88	61	27	12	88	59	29	13	87	79	8
Alaska	—	—	—	—	—	—	—	—	—	—	—	—
Arizona	13	87	30	57	16	84	20	65	9	91	43	47
Arkansas	8	92	19	73	9	91	16	74	5	95	27	68
California	7	93	70	23	16	84	30	54	4	96	80	16
Colorado	9	91	27	64	12	88	12	76	5	95	47	47
Connecticut	10	90	20	70	10	90	17	73	17	83	29	53
Delaware	8	92	9	83	8	92	8	84	12	88	13	74
Florida	9	91	8	83	9	91	8	83	12	88	6	82
Georgia	11	89	16	73	13	87	14	73	#	100	25	75
Hawaii	9	91	34	56	8	92	27	65	13	87	45	42
Idaho	10	90	33	57	13	87	28	59	2	98	45	53
Illinois	9	91	18	73	6	94	16	78	22	78	25	53
Indiana	13	87	17	70	14	86	11	75	7	93	45	48
Iowa	7	93	15	78	8	92	10	82	10	90	41	49
Kansas	—	—	—	—	—	—	—	—	—	—	—	—
Kentucky	19	81	14	67	19	81	13	68	30	70	21	49
Louisiana	9	91	15	76	10	90	13	78	5	95	45	50
Maine	8	92	18	73	9	91	16	75	3	97	38	59
Maryland	18	82	9	73	20	80	9	71	10	90	6	83
Massachusetts	17	83	15	68	17	83	13	70	18	82	29	52
Michigan	17	83	23	60	19	81	15	67	11	89	74	15
Minnesota	13	87	36	50	14	86	22	63	13	87	64	23
Mississippi	10	90	17	73	10	90	15	75	12	88	50	38
Missouri	9	91	21	70	8	92	20	72	28	72	35	38
Montana	14	86	21	65	16	84	12	73	4	96	59	37
Nebraska	—	—	—	—	—	—	—	—	—	—	—	—
Nevada	8	92	31	61	12	88	19	69	4	96	43	52
New Hampshire	10	90	24	66	10	90	23	67	9	91	35	57
New Jersey	12	88	8	80	11	89	8	81	22	78	3	75
New Mexico	15	85	36	49	23	77	21	57	9	91	47	43
New York	9	91	5	86	7	93	4	89	16	84	5	79
North Carolina	10	90	16	74	11	89	9	80	8	92	34	58
North Dakota	25	75	19	55	27	73	17	57	35	65	38	27
Ohio	15	85	8	76	14	86	7	78	44	56	18	38
Oklahoma	19	81	23	58	21	79	16	63	12	88	58	31
Oregon	9	91	43	47	13	87	36	52	1	99	53	45
Pennsylvania	8	92	13	79	9	91	11	80	9	91	27	65
Rhode Island	14	86	21	66	10	90	20	69	38	62	23	39
South Carolina	13	87	31	56	14	86	28	58	8	92	39	53
South Dakota	11	89	27	62	12	88	23	65	5	95	60	35
Tennessee	13	87	11	76	14	86	11	75	16	84	12	73
Texas	21	79	41	39	26	74	25	49	15	85	64	22
Utah	12	88	27	61	15	85	18	67	5	95	43	51
Vermont	—	—	—	—	—	—	—	—	—	—	—	—
Virginia	11	89	25	64	12	88	21	67	8	92	39	53
Washington	15	85	32	52	17	83	27	56	12	88	43	45
West Virginia	10	90	26	64	11	89	24	66	#	100	73	27
Wisconsin	12	88	16	72	13	87	14	73	12	88	21	67
Wyoming	10	90	21	69	11	89	20	70	10	90	34	57
Other jurisdictions												
District of Columbia	—	—	—	—	—	—	—	—	—	—	—	—
DoDEA <sup>1</sup>	16	84	26	58	9	91	14	76	26	74	45	29

— Not available. The jurisdiction did not participate.

# Rounds to zero.

<sup>1</sup> Department of Defense Education Activity (overseas and domestic schools).

NOTE: Students identified as both SD and ELL were counted only once under the combined SD and/or ELL category, but were counted separately under the SD and ELL categories. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-20. Percentage of fourth-grade public school students with disabilities (SD) and/or English language learners (ELL) excluded and assessed in NAEP science, as a percent of all students, by SD/ELL category and urban district: 2009**

SD/ELL category and district	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>SD and/or ELL</b>					
<b>Nation (public)</b>	TBA	TBA	TBA	TBA	TBA
<b>Large city (public)</b>	TBA	TBA	TBA	TBA	TBA
Atlanta	TBA	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA	TBA
Jefferson County (KY)	TBA	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA	TBA
<b>SD</b>					
<b>Nation (public)</b>	TBA	TBA	TBA	TBA	TBA
<b>Large city (public)</b>	TBA	TBA	TBA	TBA	TBA
Atlanta	TBA	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA	TBA
Jefferson County (KY)	TBA	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA	TBA
<b>ELL</b>					
<b>Nation (public)</b>	TBA	TBA	TBA	TBA	TBA
<b>Large city (public)</b>	TBA	TBA	TBA	TBA	TBA
Atlanta	TBA	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA	TBA
Jefferson County (KY)	TBA	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA	TBA

# Rounds to zero.

NOTE: Students identified as both SD and ELL were counted only once under the combined SD and/or ELL category, but were counted separately under the SD and ELL categories. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

**Table A-21. Percentage of eighth-grade public school students with disabilities (SD) and/or English language learners (ELL) excluded and assessed in NAEP science, as a percent of all students, by SD/ELL category and urban district: 2009**

SD/ELL category and district	Identified	Excluded	Assessed	Assessed without accommodations	Assessed with accommodations
<b>SD and/or ELL</b>					
<b>Nation (public)</b>	TBA	TBA	TBA	TBA	TBA
<b>Large city (public)</b>	TBA	TBA	TBA	TBA	TBA
Atlanta	TBA	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA	TBA
Jefferson County (KY)	TBA	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA	TBA
<b>SD</b>					
<b>Nation (public)</b>	TBA	TBA	TBA	TBA	TBA
<b>Large city (public)</b>	TBA	TBA	TBA	TBA	TBA
Atlanta	TBA	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA	TBA
Jefferson County (KY)	TBA	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA	TBA
<b>ELL</b>					
<b>Nation (public)</b>	TBA	TBA	TBA	TBA	TBA
<b>Large city (public)</b>	TBA	TBA	TBA	TBA	TBA
Atlanta	TBA	TBA	TBA	TBA	TBA
Austin	TBA	TBA	TBA	TBA	TBA
Baltimore City	TBA	TBA	TBA	TBA	TBA
Boston	TBA	TBA	TBA	TBA	TBA
Charlotte	TBA	TBA	TBA	TBA	TBA
Chicago	TBA	TBA	TBA	TBA	TBA
Cleveland	TBA	TBA	TBA	TBA	TBA
Detroit	TBA	TBA	TBA	TBA	TBA
Fresno	TBA	TBA	TBA	TBA	TBA
Houston	TBA	TBA	TBA	TBA	TBA
Jefferson County (KY)	TBA	TBA	TBA	TBA	TBA
Los Angeles	TBA	TBA	TBA	TBA	TBA
Miami-Dade	TBA	TBA	TBA	TBA	TBA
Milwaukee	TBA	TBA	TBA	TBA	TBA
New York City	TBA	TBA	TBA	TBA	TBA
Philadelphia	TBA	TBA	TBA	TBA	TBA
San Diego	TBA	TBA	TBA	TBA	TBA

# Rounds to zero.

NOTE: Students identified as both SD and ELL were counted only once under the combined SD and/or ELL category, but were counted separately under the SD and ELL categories. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Science Assessment.

## Data Collection

The NAEP 2009 science assessment was conducted from January to March 2009 by contractors to the U.S. Department of Education. Data collection for NAEP involves a collaborative effort among the participating schools, school districts, states, and NAEP staff. To reduce the burden on the participating schools, NAEP field staff perform most of the work associated with the assessment. The cooperation of the schools involves enlisting a school staff member to assist in coordinating selected students and providing space to administer the assessments.

Assessment sessions are scripted so that all students are given the same instructions and opportunity to demonstrate what they know and can do. Assessment administrators conduct the sessions under the supervision of their team's assessment coordinator. Training of assessment administrators focuses on their responsibilities in the classroom and on reading the scripts verbatim to administer the sessions in a uniform manner.

NAEP procedures guarantee the anonymity of participants. The names of students are never removed from the schools. The results of NAEP are reported on the national level and by region of the country, state, and for some urban districts—not by school or individual student.

## Scoring

Four types of cognitive items were scored for the NAEP science assessment. Responses to multiple-choice questions were scored by high-speed scanners during student booklet processing. Dichotomous constructed-response (correct and incorrect), short constructed-response (correct, partial, and incorrect) and the extended constructed-response questions (those with four or five valid score points) were scored by trained personnel using high-definition images of student responses also captured during processing.

Scoring a large number of short and extended constructed-responses with a high level of accuracy and reliability within a limited time frame is essential to the success of NAEP. To ensure reliable, efficient scoring, NAEP does the following:

- develops focused, explicit scoring guides for each item that match the criteria delineated in the assessment frameworks;
- pilot tests all items and adjusts the scoring guides (if necessary) to reflect actual student responses;
- recruits qualified and experienced scorers, trains them, and verifies their ability to score particular questions through qualifying tests;
- employs an image-processing and scoring system that routes images of student responses directly to the scorers so they can focus on scoring rather than paper routing;
- monitors scorer consistency through a second scoring. This procedure randomly selects 5 percent of state samples and 25 percent of the national sample to score twice by different scorers;
- assesses the quality of scorer decision-making through constant monitoring by NAEP assessment experts; and
- documents all training, scoring, and quality control procedures in the technical reports.

For the 2009 science assessment, more than four million individual student responses were scored in all three grades (including second scoring to monitor within-year interrater reliability). There are approximately 3/4 of the 2009 science items that had 90 percent or higher exact agreement between raters of the same student responses. Note that for scoring purposes, each individual part of a multipart item or the bilingual versions of a regular item that is given to the bilingual accommodated students were scored as separate items.

## Data Analysis and Scaling

The goal of the analysis of NAEP data is to summarize the performance of groups of students. Initial analysis activities verify the accuracy of the data and data files used in the analysis and provide the first indication of aspects of the data and analysis that require special consideration and attention. The first step is to determine the percentages of students who gave various responses to each cognitive item. Next, the properties of the items are further examined using classical test theory measures of item difficulty and item discrimination. Some of these activities are conducted without student weights or with preliminary student weights, but final student weights are used whenever possible.

After the initial activities are completed, Item Response Theory (IRT) models are used to describe the relationships between the item responses provided by students and the underlying scale. The primary purpose of IRT scaling is to provide a common scale on which performance can be compared, even when students receive different blocks of items. Item parameters that are used in the models are estimated from student response data for each item. Different IRT models with different types of item parameters are used to describe multiple-choice items, dichotomous constructed-response items, and polytomous constructed-response items.

Because the NAEP matrix design gives each student a small proportion of the pool of assessment items, the assessment cannot provide reliable information about individual student performance. Traditional test scores for individual students, even those based on IRT, would result in misleading estimates of population characteristics, such as student group means and percentages of students at or above a certain scale-score level. However, it is NAEP's goal to estimate these population characteristics. NAEP's objectives can be achieved with methodologies that produce estimates of the population-level parameters using marginal estimation techniques for latent variables. Under the assumptions of the analysis models, these population estimates will be consistent in the sense that the estimates approach the population values as the sample size increases.

Prior to 2009, the overall science scale for each grade was a composite scale as a weighted average of subscales estimated for each of the science content areas. Starting with the new 2009 science framework, the overall science scale for each grade is estimated as a single scale. IRT and the NAEP marginal estimation methodology are used to estimate the overall score scale. The overall scale for each grade ranges from 0 to 300, and summarizes student performance across all three science content areas (Physical Science, Life Science, and Earth and Space Sciences) and across all three types of questions in the assessment (multiple choice, short constructed response, and extended constructed response). Summary statistics of the scale scores are estimated, and statistical tests are used to make inferences about the comparisons of results for different groups of students. Finally, NAEP scale score distributions are described via achievement levels and/or item mapping procedures. Additionally, score scales are estimated for each of the three science content areas (Physical Science, Life Science, and Earth and Space Sciences). These subscale scores are also reported on a 0 to 300 scale. For more information about NAEP analysis, IRT, and scaling see <http://nces.ed.gov/nationsreportcard/tdw/analysis/>.

## Variance Estimation

The averages and percentages in this report are estimates based on samples of students rather than on entire populations. Moreover, the collection of questions used at each grade level is only a sample of the many questions that could have been asked to assess the skills and abilities described in the NAEP framework, and each assessed student takes only a subset of the entire collection of questions. Therefore, the results are subject to a measure of uncertainty, reflected in the standard error of the estimates—a range of up to a few points above or below the score or percentage—which takes into account potential score fluctuation due to both sampling error and measurement error.

Because NAEP uses complex sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate. NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any student information that can be observed without error. However, because each student typically responds to only a few questions within any science content area, the estimated scale score for any single student would be imprecise. In this case, NAEP's marginal estimation methodology is used to describe the performance of groups of students without requiring precise estimates of individual student performance. The estimate of the variance of the students' scale score distributions (which reflect the imprecision due to lack of measurement accuracy) is computed. This component of variability is then included in the standard errors of NAEP scale scores.

## Drawing Inferences from the NAEP Results

Drawing correct inferences from NAEP assessment results depends on the use of appropriate statistical procedures for comparing assessment results for population groups of interest and following guidelines to ensure the validity of the inferences. Comparisons of different groups of students with respect to scores or percentages of a certain attribute are of primary interest to users of NAEP results. The user is cautioned to rely on the results of statistical tests, rather than on the apparent magnitude of the difference between two numbers when determining whether differences are likely to represent actual differences among the groups in the population.

***t* Test Comparison:** By convention, references to differences in NAEP reports indicate that scores or percentages from two groups are different (e.g., one group performed higher or lower than another group) only when the difference in the point estimates for the groups being compared is statistically significant at a level of .05.

Since 1998, *t* tests have been used for most NAEP comparisons. These tests are more appropriate than *z* tests (based on normal distribution approximations) when the statistics that are being compared are from distributions with proportionally larger extremes (i.e., thicker tails) than the normal distribution. One aspect of the use of *t* tests that contributes to the difficulty in their use for large-scale surveys is the determination of the appropriate degrees of freedom for the *t* distribution of interest.

**Multiple Comparison Procedures:** The *t* test used by NAEP and the certainty ascribed to intervals (e.g., a 95 percent confidence interval) are based on statistical theory that assumes that only one confidence interval is being examined or one test of statistical significance is being performed. However, in some sections of a report, many different groups may be compared (i.e., multiple sets of confidence intervals are being analyzed). In sets of confidence intervals, statistical theory indicates that certainty associated with the entire set of intervals is less than that attributable to each individual comparison from the set. To hold the significance level for the set of comparisons at a particular level (e.g., .05), adjustments—called multiple comparison procedures—must be made to the methods.

To ensure that comparisons made using NAEP data are as accurate as possible, error rates are controlled when multiple comparisons are made. When making a number of comparisons in a single analysis, such as analyzing White student performance versus the performance of Black, Hispanic, Asian/Pacific Islander, and American Indian/Alaska Native students, the probability of finding significant differences by chance, for at least one comparison, increases with the family size or number of comparisons. There are several ways to take into account how many related comparisons are being made. In NAEP, the Benjamini-Hochberg False Discovery Rate (FDR) procedure is used to control for this.

Unlike other multiple comparison procedures (e.g., the Bonferroni procedure) that control the familywise error rate (i.e., the probability of making even one false rejection in the set of comparisons), the FDR procedure controls the expected proportion of falsely rejected hypotheses. Familywise procedures are considered conservative for large families of comparisons; therefore the FDR procedure is more suitable for multiple comparisons in NAEP than other procedures. There are two exceptions where the FDR is not applied: when comparing multiple years and when comparing a state's overall results to the nation, but not to other states.

## NAEP Reporting Groups

In addition to overall results for each grade assessed, NAEP results are reported for certain student groups provided there are sufficient numbers of students and adequate school representation. Results for some student groups may not be available for certain years, grades, or jurisdictions.

**Race/Ethnicity:** The school-recorded race/ethnicity variable records the race/ethnicity of each student as reported by the student's school. When the school-recorded information is missing, student-reported data derived from the student background questions are used. The mutually exclusive racial/ethnic categories are White, Black, Hispanic, Asian/Pacific Islander, American Indian/Alaska Native, and Other. Black includes African American, Hispanic includes Latino, and Pacific Islander includes Native Hawaiian. Race categories exclude Hispanic origin unless specified. Unclassified students are those whose school-reported race/ethnicity was "other" or "unavailable" or was missing, and whose race/ethnicity category could not be determined from self-reported information.

**Gender:** The gender of the student assessed is taken from school records.

**Eligibility for the National School Lunch Program:** The school lunch variable is based on available school records. Students are classified as either currently eligible or not currently eligible for the national lunch component of the Department of Agriculture's National School Lunch Program. The classification refers only to the school year when the assessment was administered and is not based on eligibility in previous years. If school records are not available, the student is classified as "Information not available." If the school did not participate in the program, all students in that school were classified as "Information not available." Eligibility for the program is determined by students' family income in relation to the federally established poverty level. Free lunch qualification is set at 130 percent of the poverty level or below, and reduced-price lunch qualification is set at between 130 and 185 percent of the poverty level. (For the period July 1, 2008 through June 30, 2009, for a family of four, 130 percent of the poverty level was \$27,560, and 185 percent was \$39,220.) Additional information on eligibility may be found at the U.S. Department of Agriculture website at <http://www.fns.usda.gov/cnd/lunch/>.

**Type of Location:** Results for four mutually exclusive categories of school location are also reported: city, suburb, town, and rural. The categories are based on standard definitions established by the Federal Office of Management and Budget using population and geographic information from the U.S. Census Bureau. Schools are assigned to these categories in the NCES Common Core of Data based on their physical address. The classification system was revised for 2007; therefore, trend comparisons to previous years are not available. The new locale codes are based on an address's proximity to an urbanized area (a densely settled core with densely settled surrounding areas). This is a change from the original system based on metropolitan statistical areas. To distinguish the two systems, the new system is referred to as "urban-centric locale codes."

**Parental Education:** Eighth- and twelfth-graders were asked the following two questions, the responses to which were combined to derive the parental education variable:

How far in school did your mother go?

- She did not finish high school
- She graduated from high school
- She had some education after high school
- She graduated from college
- I don't know

How far in school did your father go?

- He did not finish high school
- He graduated from high school
- He had some education after high school
- He graduated from college
- I don't know

The information was combined into one parental-education reporting variable in the following way:

- If a student indicated the extent of education for only one parent, that level was included in the data. If a student indicated the extent of education for both parents, the higher of the two levels was included in the data.
- If a student responded "I don't know" for both parents, or responded "I don't know" for one parent and did not respond for the other, the parental education level was classified as "I don't know."
- If the student did not respond for either parent, the student was recorded as having provided no response.

Because fourth-graders' responses to the questions tend to be not reliable, the questions were not presented to students at grade 4 in 2009.

**Region of the Country:** Prior to 2003, NAEP results were reported for four NAEP-defined regions of the nation: Northeast, Southeast, Central, and West. To align NAEP with other federal data collections, NAEP analysis and reports have used the U.S. Census Bureau's definition of "region" beginning in 2003. The four regions defined by the U.S. Census Bureau are Northeast, South, Midwest, and West. Therefore, trend data by region are not provided for assessment years prior to 2003.

Figure A-1 shows how states are divided into these census regions. All 50 states and the District of Columbia are listed. Other jurisdictions, including the Department of Defense Education Activity schools, are not assigned to any region.

**Figure A-1. States within regions of the country defined by the U.S. Census Bureau**

<b>Northeast</b>	<b>South</b>	<b>Midwest</b>	<b>West</b>
Connecticut	Alabama	Illinois	Alaska
Maine	Arkansas	Indiana	Arizona
Massachusetts	Delaware	Iowa	California
New Hampshire	District of Columbia	Kansas	Colorado
New Jersey	Florida	Michigan	Hawaii
New York	Georgia	Minnesota	Idaho
Pennsylvania	Kentucky	Missouri	Montana
Rhode Island	Louisiana	Nebraska	Nevada
Vermont	Maryland	North Dakota	New Mexico
	Mississippi	Ohio	Oregon
	North Carolina	South Dakota	Utah
	Oklahoma	Wisconsin	Washington
	South Carolina		Wyoming
	Tennessee		
	Texas		
	Virginia		
	West Virginia		

SOURCE: U.S. Department of Commerce Economics and Statistics Administration, U.S. Census Bureau.

## **Caution in Interpretations**

The NAEP science scale makes it possible to examine relationships between students' performance and various background factors that NAEP measures. However, because NAEP assessment is a type of observational study, the relationship between achievement and another variable does not warrant any causal inferences. The results about student group performance are most useful when considered in combination with other knowledge about the student population and the educational system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.