

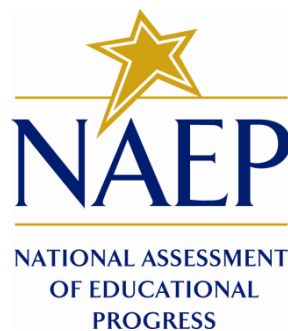
*NATIONAL CENTER FOR EDUCATION STATISTICS
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS*

Appendix G

NAEP 2011 Sample Design

*Request for Clearance for
NAEP Assessments for 2017-2019*

*OMB# 1850-NEW v.1
(previous OMB# 1850-0790 v.43)*



July 29, 2016

NAEP Technical Documentation Website

NAEP 2011 Sample Design

The sample design for NAEP 2011 included samples for various operational, special study, and pilot test assessments. Representative samples were drawn for the following operational assessments:

2011 State Assessment Sample Design

- national assessments in mathematics and reading in public and private schools at grades 4 and 8;
- national assessments in computer-based writing (WCBA) in public and private schools at grades 8 and 12;
- national assessments in science in public and private schools at grade 8;
- state-by-state and Trial Urban District Assessments (TUDA) assessments in mathematics and reading in public schools at grades 4 and 8; and
- state-by-state assessments in science in public schools at grade 8.

2011 National Assessment Sample Design

Representative samples were drawn for the following special study and pilot test assessments:

- mathematics computer-based study (MCBS) in public schools at grade 8;
- study to examine a direct link between NAEP and Trends in International Mathematics and Science Study (TIMSS) in public schools at grade 8;
- Special mathematics assessment in Puerto Rico in public and private schools at grade 4 and in public schools at grade 8; and
- pilot tests in reading and mathematics in public and private schools at grade 4, in reading and mathematics in public schools at grade 8, and in economics in public schools at grade 12.

The samples for the operational assessments were organized into four distinct components and sampled separately. The samples for the special studies and pilot tests were integrated into these various components.

- mathematics, reading, and science assessments of fourth- and eighth-grade students in public schools;
- mathematics, reading, and science assessments of fourth-grade and eighth-grade students in private schools;
- computer-based writing assessments and mathematics study of eighth-grade and twelfth-grade students in public schools; and
- computer-based writing assessments of eighth-grade and twelfth-grade students in private schools.

The national assessments were designed to achieve nationally representative samples of public and private school students in the fourth, eighth, and twelfth grades. Their target populations included all students in public, private, Bureau of Indian Education (BIE), and Department of Defense Education Activity (DoDEA) schools, who were enrolled in fourth, eighth, and twelfth grades, respectively, at the time of assessment.

For the fourth- and eighth-grade mathematics, reading, and science assessments in public schools, the NAEP state student samples and assessments constitute the NAEP national student samples and assessments. Nationally representative samples were drawn for the remaining populations of private school students in fourth and eighth grades.

The TUDA samples formed part of the corresponding state public school samples, and the state samples formed the public school grade 4 and 8 part of the national sample.

The mathematics, reading, and science samples were based on a two-stage sample design:

- selection of schools within strata, and
- selection of students within schools.

The computer-based writing and mathematics samples were based on a three-stage sample design:

- selection of primary sampling units (PSUs),
- selection of schools within strata, and
- selection of students within schools.

In the three-stage design, schools were stratified and selected within the sampled PSUs. The samples of schools were selected with probability proportional to a measure of size based on the estimated grade-specific enrollment in the schools for both designs.

The state assessments were designed to achieve representative samples of students in the fourth and eighth grades. Their target populations included all students in each participating jurisdiction, which included states, District of Columbia, BIE, DoDEA, and school districts chosen for the TUDA assessments. Each sample was designed to produce aggregate estimates with reliable precision for all the participating jurisdictions, as well as estimates for various student subpopulations of interest.

At grades 4 and 8, all BIE schools were included in the mathematics, reading, and science assessments. Also, public schools with relatively high American Indian/Alaska Native populations were oversampled in six states (Arizona, Minnesota, North Carolina, Oregon, Utah, and Washington). This was designed to enhance the reporting of results for American Indian students at the state level in those states with a sizable proportion of the nation's American Indian students for the National Indian Education Study (NIES), which was conducted in conjunction with NAEP.

All states participated in the mathematics, reading, and science assessments. By design, only BIE schools did not participate in the state science assessment, as it lacked the required number of students for the state science assessment. A small portion of students received the science assessment in BIE schools in science to supplement the national science sample.

The District of Columbia, which generally does not have enough students for an assessment in a third subject, also participated in the grade 8 science assessment. To accomplish this, each student in the District of Columbia was assigned to two of the three assessment subjects and thus tested twice over two days.

The figure below illustrates the various sample types and subjects.

Components of the NAEP 2011 samples, by assessment subject, grade, and school type: 2011

Grade	Assessment			
	Reading	Mathematics	Science	WCBA
4	Public/BIE/DoDEA			
	Private			
8	Public/BIE/DoDEA			Public
	Private			Private
12				Public
				Private

NOTE: View an accessible version of this figure.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Assessments.

Components of the NAEP samples, by assessment subject, school type, and grade: 2011

Components of the NAEP 2011 samples, by assessment subject, school type, and grade: 2011

Grade	Assessment			
	Reading	Mathematics	Science	WCBA
4	(1) (3)	(1) (3)		
8	(1) (3)	(1) (3)	(1) (3)	(2) (3)
12				(2) (3)

¹Public/Bureau of Indian Education (BIE)/Department of Defense Education Activity (DoDEA).

²Public.

³Private.

NOTE: WCBA = Writing computer-based assessment.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Assessments.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/sampdsngn_2011_accessible.aspx

Sample Design for the 2011 National Assessment

The 2011 national assessment included the following components:

- mathematics and reading assessments in public and private schools at grades 4 and 8;
- writing computer-based assessment (WCBA) in public and private schools at grades 8 and 12;
- science assessments in public and private schools at grade 8.

The sample design aimed to achieve a nationally representative sample of students in the defined populations who were enrolled at the time of assessment.

The mathematics and reading samples were based on a two-stage sample design:

- selection of schools within strata, and
- selection of students within schools.

The computer-based writing and mathematics samples were based on a three-stage sample design:

- selection of primary sampling units (PSUs),
- selection of schools within strata, and
- selection of students within schools.

The samples of schools were selected with probability proportional to a measure of size based on the estimated grade-specific enrollment in the schools.

For the mathematics, reading, and science assessments in fourth- and eighth-grade public schools, the NAEP state student samples and assessments constitute the NAEP national student samples and assessments. Nationally representative samples were drawn for the remaining populations of private school students in fourth and eighth grades. By design, only Bureau of Indian Education (BIE) schools did not participate in the state science assessment, as it lacked the required number of students for the state science assessment. A small portion of students received the science assessment in BIE schools in science to supplement the national science sample.

Fourth- and Eighth-Grade Public School Assessments

[Fourth- and Eighth-Grade Private School National Assessment](#)

Writing Computer-Based Assessment (WCBA) and Mathematics Computer-Based Study (MCBS)

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_nat_main.aspx

2011 Fourth- and Eighth-Grade Private School National Assessment

The private school samples were designed to produce nationally representative samples of students enrolled in private schools in the United States. Fourth- and eighth-grade students were assessed in mathematics and reading.

Private school students were sampled for the eighth-grade national science assessment at a very low rate. The three operational subjects (reading, mathematics, science) were sampled in the ratio of 9:9:1. This ensured enough private school sample to report a national science result, but does not support breakdowns by type of private school.

Reading pilots and a special mathematics assessment in Puerto Rico were also conducted in the private school samples for fourth grade.

Oversampling of private schools at grades 4 and 8, last implemented in 2005, was reintroduced. Response rates permitting, allowed separate reporting for reading and mathematics, for Catholic, Lutheran, Conservative Christian, and other private schools.

The target sample sizes of assessed students for each grade and subject are shown in the table below. Prior to sampling, these target sample sizes were adjusted upward to offset expected rates of school and student attrition due to nonresponse and ineligibility.

Target Population

Sampling Frame

Stratification of Schools

School Sample Selection

Substitute Schools

Ineligible Schools

Student Sample Selection

School and Student Participation

Target sample sizes of assessed students, private school national assessment, by subject and grade: 2011

Grade	Total	Mathematics	Mathematics pilot	Reading	Reading pilot	Science	Special mathematics assessment
Total	25,240	12,000	200	12,000	220	670	150
4	12,570	6,000	200	6,000	220	†	150
8	12,670	6,000	†	6,000	†	670	†

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Samples were based on a two-stage design that involved selection of schools within strata and selection of students within schools. The first-stage samples of schools were selected with probability proportional to a measure of size based on the estimated grade-specific enrollment in the schools.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_priv_gr_4_8.aspx

Ineligible Schools for the 2011 Private School National Assessment

The Private School Universe Survey (PSS) school file, from which most of the sampled schools were drawn, corresponds to the 2007-2008 school year, 3 years prior to the assessment school year. During the intervening period, some of these schools either closed, no longer offered the grade of interest, or were ineligible for other reasons. In such cases, the sampled schools were coded as ineligible.

Eligibility Status of Sampled Schools by Grade and Private School Type

Ineligible Sampled Schools by Ineligibility Type

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_inelg.aspx

Eligibility Status of Sampled Schools for the 2011 Private School National Assessment

The following table presents a breakdown by private school type of ineligible and eligible schools in the fourth- and eighth-grade private school samples. There are considerable differences across private school types at grades 4 and 8. Schools whose private school type was unknown at the time of sampling subsequently had their affiliation determined during data collection. Therefore, such schools are not broken out separately.

Eligibility status of sampled private schools, national assessment, by grade and private school type: 2011

Private school type	Eligibility status	Fourth grade		Eighth grade	
		Count	Percentage	Count	Percentage
All private	Total	748	100.00	930	100.00
	Ineligible	102	13.64	126	13.55
	Eligible	646	86.36	804	86.45
Catholic	Total	264	100.00	332	100.00
	Ineligible	26	9.85	26	7.83
	Eligible	238	90.15	306	92.17
Non-Catholic	Total	484	100.00	598	100.00
	Ineligible	76	15.70	100	16.72
	Eligible	408	84.30	498	83.28
Lutheran	Total	107	100.00	141	100.00
	Ineligible	8	7.48	7	4.96
	Eligible	99	92.52	134	95.04
Conservative Christian	Total	123	100.00	150	100.00
	Ineligible	17	13.82	22	14.67
	Eligible	106	86.18	128	85.33
Other private	Total	254	100.00	307	100.00
	Ineligible	51	20.08	71	23.13
	Eligible	203	79.92	236	76.87

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_delta_inel_elig_status.aspx

Ineligible Sampled Private Schools for the 2011 National Assessment

The table below presents unweighted counts of sampled schools, by grade and eligibility status, for the private school samples.

NAEP sample private schools, national assessment, by grade and eligibility status: 2011

Grade and eligibility status	Unweighted count of schools	Unweighted percentage
All fourth-grade sampled private schools	748	100.00
Eligible	646	86.36
Has sampled grade, but no eligible students	14	1.87
Does not have sampled grade	22	2.94
Closed	55	7.35
Not a regular school	7	0.94
Duplicate on sampling frame	2	0.27
Other ineligible	2	0.27
All eighth-grade sampled private schools	930	100.00
Eligible	804	86.45
Has sampled grade, but no eligible students	19	2.04
Does not have sampled grade	26	2.80
Closed	52	5.59
Not a regular school	19	2.04
Duplicate on sampling frame	4	0.43
Other ineligible	6	0.65

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_delta_inelgtype.aspx

Sampling Frame for the 2011 Private School National Assessment

The frame of the private schools in all three grades was developed from the 2007-2008 Private School Universe Survey (PSS), a survey conducted by the U.S. Census Bureau for the National Center for Education Statistics (NCES). The PSS is a biennial mail survey of all private schools in the 50 states and the District of Columbia. The PSS frame of schools comprises both a list frame and an area frame. The 2007-2008 list frame is an assembly of the 2005-2006 PSS frame and more up-to-date lists from state education agencies, private school associations, and other easily accessible sources. To improve the coverage of the PSS list frame, the Census Bureau also conducted a survey to locate private schools in a random sample of geographic areas throughout the United States. The areas were single counties or groups of counties sampled from an area frame constructed from all counties in the nation. Within each selected area a complete list of private schools was gathered using information from the Yellow Pages, religious institutions, local education agencies, chambers of commerce, and local government offices. Schools not already on the list frame were identified and added to the frame of private schools. A weighting component was computed by the Census Bureau so that the additional area-frame schools would represent all schools absent from the list frame, not just those in the selected areas.

The sampling frame excluded schools that were ungraded, provided only special education, were part of hospital or treatment center programs, were juvenile correctional institutions, were home-school entities, or were for adult education.

Private school affiliation is unknown for nonrespondents to the PSS. Because oversampling was desired to report by affiliation, additional work was done to obtain affiliation for nonrespondents to the PSS. If a nonresponding school responded to a previous PSS (either two or four years prior), affiliation was obtained from the previous response. For those schools that were nonrespondents for the last three cycles of the PSS, in some cases Internet research was used to establish affiliation. There were still schools with unknown affiliation remaining after this process.

For quality control purposes, school and student counts from the 2011 sampling frame were compared to school and student counts from previous NAEP frames (2009 and 2007). No major issues were found.

[Fourth- and Eighth-Grade Schools and Enrollment in the Private School Sampling Frame](#)
[New-School Sampling Frame for the Private School Assessment](#)

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_delta_sampfrme.aspx

Fourth- and Eighth-Grade Schools and Enrollment in the 2011 Private School Sampling Frame

The following table displays, by grade and affiliation, the number of private schools in the sampling frame and their estimated enrollment. Enrollment was estimated for each school as the Private School Universe Survey (PSS)--reported enrollment averaged across grades 1 through 8.

The counts presented below are of schools with known affiliation. Schools with unknown affiliation do not appear in the table because their grade span, affiliation, and enrollment were unknown. Although PSS is a school universe survey, participation is voluntary and not all private schools respond. Since the NAEP sample must represent all private schools, not just PSS respondents, a small sample of PSS nonrespondents with unknown affiliation was selected for each of the targeted grades to improve NAEP coverage.

Number of schools and enrollment in private school sampling frame, national assessment, by affiliation and grade: 2011

Grade	Affiliation	Number of schools	Estimated enrollment
4	Total	20,110	383,849
	Catholic	5,974	171,054
	Non-Catholic private	14,136	212,795
	Lutheran	1,374	18,086
	Conservative Christian	4,080	61,504
	Other private	8,682	133,205
8	Total	17,968	369,381
	Catholic	5,465	170,509
	Non-Catholic private	12,503	198,872
	Lutheran	1,166	16,579
	Conservative Christian	3,636	57,363
	Other private	7,701	124,930

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_sampfme_gr_4_8.aspx

New-School Sampling Frame for the 2011 Private School Assessments

Whereas the Private School Universe Survey (PSS) file used for the frame corresponds to the 2007-2008 school year, the NAEP assessment year was the 2010-2011 school year. During this 3-year period, some schools closed, some changed their grade span, and still others came into existence.

To achieve as close to full coverage as possible, the private school frame was supplemented by a sample of new Catholic schools. The goal was to allow every such school a chance of selection, thereby fully covering the target population of Catholic schools in operation during the 2010-2011 school year. The first step in this process was the development of a new-school frame through the construction of a diocese-level file from the PSS school-level file. To develop the frame, the diocese-level file was divided into two files: one for small dioceses and the other for medium and large dioceses.

Small dioceses contained no more than three schools on the frame in total, with no more than one school at each grade (fourth, eighth, and twelfth). New schools in small dioceses were identified during school recruitment and added to the sample if the old school in the same diocese was sampled at the relevant grade. From a sampling perspective, the new school was viewed as an "annex" to the sampled school that had a well-defined probability of selection equal to that of the old school. The "frame" in this case was, in fact, the original frame; when the old school was sampled in a small diocese, the new school was automatically sampled as well.

To limit respondent burden and keep the level of effort within reasonable bounds, the new-school frame was created using information obtained from a sample of the remaining dioceses. The remaining dioceses were separated into two strata of large- and medium-size dioceses. These strata were defined by computing the percentage of the nation's total Catholic school enrollment each diocese represents, sorting the dioceses in descending order by that percentage, and cumulating the percentages across the sorted file. All dioceses up to and including the first diocese at or above the 80th cumulative percentage were defined as large dioceses. The remaining dioceses were defined as medium dioceses.

A simplified example is given below. Dioceses are ordered by percentage enrollment. The first six become large dioceses and the last six become medium dioceses.

Example showing assignment of Catholic dioceses to the large and medium strata, private school national assessment: 2011

Diocese	Percent enrollment	Cumulative percentage enrollment	Stratum
Diocese 1	20	20	L
Diocese 2	20	40	L
Diocese 3	15	55	L
Diocese 4	10	65	L
Diocese 5	10	75	L
Diocese 6	10	85	L
Diocese 7	5	90	M
Diocese 8	2	92	M
Diocese 9	2	94	M
Diocese 10	2	96	M
Diocese 11	2	98	M
Diocese 12	2	100	M

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

In actuality, there were 71 large and 103 medium dioceses in the sampling frame.

The target sample size was 10 dioceses total: 8 large and 2 medium. In the medium stratum, the dioceses were selected with equal probability. In the large stratum, dioceses were sampled with probability proportional to enrollment. These probabilities were retained and used in all later stages of sampling and weighting in order to represent all dioceses, whether or not they had been selected as new school samples for the assessment.

Each selected diocese was sent a listing of its schools extracted from the 2007-2008 PSS file and was asked to provide information about new schools and any changes to grade span in existing schools. This information provided by the selected dioceses was used to create sampling frames for the selection of new Catholic schools. The process of obtaining the information was conducted with the help of the National Catholic Educational Association (NCEA). NCEA was sent the school lists for the 10 sampled dioceses and was responsible for returning the completed updates.

The eligibility of a new school at a particular grade was determined by its grade span. A school already on PSS also was classified as "new" if a change of grade span had occurred such that the school status changed from ineligible to eligible at a particular grade.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_newschoofframe_4_8.aspx

Sampling of Schools for the 2011 Private School National Assessment

The private school samples were selected with probability proportional to size using systematic sampling from a sorted list. A school's measure of size was a complex function of the school's estimated grade enrollment. For the eighth grade sample, multiple "hits" were allowed per school, but this was not the case for the fourth grade sample.

Computation of
Measures of Size

Schools were ordered within each school type using a serpentine sort involving the following variables:

School Sample Sizes:
Frame and New
School

- census division,
- urbanicity classification (based on urban-centric locale),
- race/ethnicity status, and
- estimated grade enrollment.

A systematic sample was then drawn with probability proportional to size using this serpentine sorted list and the measures of size.

Schools with unknown affiliation were treated separately. A sample of about 30 schools with unknown affiliation was selected at each of the two grades.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_schlsamp.aspx

Computation of Measures of Size

There were five objectives underlying the process for determining the probability of selection for each school and for setting the number of students to be sampled within each selected school:

- to meet the target student sample size for each grade;
- to select an equal-probability sample of students;
- to limit the number of students selected from any one school;
- to ensure that the sample within a school does not include a very high percentage of the students in the school, unless all students are included; and
- to reduce the rate of sampling of small schools, in recognition of the greater cost and burden per student of conducting assessments in such schools.

The goal in determining the school's measure of size is to optimize across the last four objectives in terms of maintaining the precision of estimates and the cost effectiveness of the sample design. The following algorithm was used to assign a measure of size to each school based on its estimated grade-specific enrollment.

In the formulas below, x_{js} refers to the estimated grade enrollment for private school type j and school s and P_s is a primary sampling unit (PSU) weight associated with the private school universe (PSS) area sample.

The preliminary measures of size (MOS) were set as follows:

$$MOS_{js} = P_s \times \begin{cases} x_{js} & \text{if } 70 < x_{js} \\ 63 & \text{if } 20 < x_{js} \leq 70 \\ \left(\frac{63}{20}\right) \times x_{js} & \text{if } 5 < x_{js} \leq 20 \\ \frac{63}{4} & \text{if } x_{js} \leq 5 \end{cases}$$

The preliminary school measure of size was rescaled to create an expected number of hits by applying a multiplicative constant b_j , which varies by grade and school type. The private school sample design allowed multiple "hits." For example, a school with two hits will have twice as many students sampled as a single-hit school, etc. To limit respondent burden, constraints were placed on the number of hits allowed per school. For grade 4 it was one hit, and for grade 8 it was two.

It follows that the final measure of size, E_{js} , was defined as:

$$E_{js} = \min(b_j \times MOS_{js}, u_j)$$

where u_j is the maximum number of hits allowed.

The school's probability of selection π_{js} was given by:

$$\pi_{js} = \min(E_{js}, 1)$$

One can choose a value of b_j such that the expected overall student sample yield matches the desired targets specified by the design, where the expected yield is calculated by summing the product of an individual school's probability and its student sample yield across all schools in the frame.

In addition, new and newly eligible Catholic schools were sampled from the new-school frame. The assigned measures of size for these schools,

$$E_{js} = \min(b_j \times MOS_{js} \times \pi_{djs}^{-1}, u_j),$$

used the b_j and u_j values from the main school sample for the grade and school type (i.e., the same sampling rates as for the main school sample). The variable π_{djs} is the probability of selection of the diocese into the new-school diocese (d) sample.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_schlsamp_mos.aspx

School Sample Sizes: Frame and New School

The following table presents the number of schools selected from the private school sampling frame (constructed from the Private School Universe Survey file) and the new-school sampling frame, for grades 4 and 8, by school type.

NAEP private school frame-based and new school samples, by grade and school type: 2011

Grade	Private school type	Total school sample	Frame school sample	New school sample
4	All private			
	Catholic	264	260	4
	Non-Catholic	484	484	0
	Lutheran	109	109	0
	Conservative Christian	120	120	0
	Other private	230	230	0
	Unknown affiliation	25	25	0
8	All private			
	Catholic	330	323	7
	Non-Catholic	600	600	0
	Lutheran	141	141	0
	Conservative Christian	148	148	0
	Other private	285	285	0
	Unknown affiliation	26	26	0

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_delta_schlsamp_sampsize.aspx

School and Student Participation Rates for the 2011 Private School National Assessment

Private school participation in NAEP is not mandatory. The 2011 assessment holds true to the historic pattern of having higher rates of participation among Catholic and Lutheran schools than among Conservative Christian and other private schools. Although a portion of the participating school sample consisted of substitute schools, it is preferable to calculate school response rates on the basis of school participation before substitution.

In every NAEP survey, some of the sampled students are not assessed for the following reasons:

- withdrawn students,
- excluded students with disabilities (SD),
- excluded English language learner (ELL) students, or
- students absent from both the original session and the makeup session (not excluded but not assessed).

Withdrawn students are those who have left the school before the original assessment. Excluded students were determined by their school to be unable to meaningfully take the NAEP assessment in their assigned subject, even with an accommodation. Excluded students must also be classified as SD and/or ELL. Other students who were absent for the initial session are assessed in the makeup session. The last category includes students who were not excluded (i.e., "were to be assessed") but were not assessed either due to absence from both sessions or because of a refusal to participate. Assessed students are also classified as assessed without an accommodation or assessed with an accommodation. The latter group can be divided into SD students assessed with an accommodation, ELL students assessed with an accommodation, or students who are both SD and ELL and accommodated. Note that some SD and ELL students are assessed without accommodations, and students who are neither SD nor ELL can only be assessed without an accommodation.

The weighted response rates utilize the student base weights and indicate the weighted percentage of assessed students among all students to be assessed. The exclusion rates, in contrast, provide the weighted percentage of excluded SD or ELL students among all absent, assessed, and excluded students.

School Response Rates

Weighted Student Response and Exclusion Rates for Mathematics

Weighted Student Response and Exclusion Rates for Reading

Weighted Student Response and Exclusion Rates for Science

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_schl_and_stud_part.aspx

School Response Rates for the 2011 Private School National Assessment

The following table presents counts of eligible sampled schools and participating schools, as well as weighted school response rates, for the private school samples in which the mathematics and reading operational assessments were conducted. The weighted school response rates estimate the proportion of the student population that is represented by the participating school sample prior to substitution.

Private school response rates, national assessment, by school type and grade: 2011

Grade	Private school type	Eligible sampled schools	Participating schools, including substitutes	Weighted school response rate prior to substitution (percent)
4	All private	646	557	73.51
	Catholic	238	236	96.27
	Non-Catholic	408	321	55.34
	Lutheran	99	97	94.87
	Conservative Christian	106	94	73.13
	Other private	203	130	42.23
8	All private	804	696	74.40
	Catholic	306	299	93.23
	Non-Catholic	498	397	57.54
	Lutheran	134	129	92.73
	Conservative Christian	127	114	72.51
	Other private	237	154	45.71

NOTE: Detail may not sum to total due to rounding. Percentages are based on unrounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_schresp_rates.aspx

Weighted Student Response and Exclusion Rates for the 2011 Private School National Mathematics Assessment

The following table presents the weighted student response and exclusion rates for the mathematics assessment. The exclusion rates give the percentage of students excluded among all eligible students. Excluded students must be either students with disabilities (SD) or English language learner (ELL). The response rates indicate the percentage of students assessed among those who were intended to take the assessment from within the participating schools. Thus, students who were excluded are not included in the denominators of the response rates.

Weighted student response and exclusion rates for private schools, national mathematics assessment, by school type and grade: 2011

Grade	Private school type	Weighted student response rate	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded
4	All private	95.55	0.18	0.12
	Catholic	95.85	0.22	0.03
	Non-Catholic	95.19	0.15	0.20
	Lutheran	96.63	0.36	0.00
	Conservative Christian	93.90	0.30	0.00
	Other private	95.70	0.05	0.32
8	All private	94.77	0.44	0.06
	Catholic	95.05	0.46	0.07
	Non-Catholic	94.43	0.42	0.05
	Lutheran	96.30	0.33	0.00
	Conservative Christian	94.50	0.29	0.17
	Other private	94.02	0.49	0.00

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Mathematics Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_studresp_math.aspx

Weighted Student Response and Exclusion Rates for the 2011 Private School National Reading Assessment

The following table presents the weighted student response and exclusion rates for the reading assessment. The exclusion rates give the percentage of students excluded among all eligible students. Excluded students must necessarily be either students with disabilities (SD) or English language learners (ELL). The response rates indicate the percentage of students assessed among those who were intended to take the assessment from within the participating schools. Thus, students who were excluded are not included in the denominators of the response rates.

Weighted student response and exclusion rates for private schools, national reading assessment, by school type and grade: 2011

Grade	Private school type	Weighted student response rate	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded
4	All private	95.16	0.30	0.22
	Catholic	95.49	0.30	0.18
	Non-Catholic	94.77	0.30	0.26
	Lutheran	96.27	1.17	0.00
	Conservative Christian	95.48	0.00	0.00
	Other private	93.99	0.31	0.42
8	All private	94.80	0.40	0.07
	Catholic	95.36	0.30	0.09
	Non-Catholic	94.11	0.50	0.06
	Lutheran	95.19	0.48	0.00
	Conservative Christian	93.72	0.08	0.20
	Other private	94.13	0.69	0.00

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Reading Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_studresp_read.aspx

Weighted Student Response and Exclusion Rates for the 2011 Private School National Science Assessment

The following table presents the weighted student response and exclusion rates for the grade 8 national science assessment. The exclusion rates give the percentage of students excluded among all eligible students. Excluded students must necessarily be either students with disabilities (SD) or English language learners (ELL). The response rates indicate the percentage of students assessed among those who were intended to take the assessment from within the participating schools. Thus, students who were excluded are not included in the denominators of the response rates.

Weighted student response and exclusion rates for private schools, grade 8 science assessment, by school type: 2011

Private school type	Weighted student response rate	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded
All private	93.86	0.18	0.00
Catholic	94.27	0.35	0.00
Non-Catholic	93.34	0.00	0.00

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Science Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_nat_delta_studresp_science.aspx

Stratification of Schools in the 2011 Private School National Assessment

Explicit stratification for the NAEP 2011 private school samples was by private school type: Catholic, Lutheran, Conservative Christian, Other Private, and unknown affiliation. Private school affiliation was unknown for nonrespondents to the NCES Private School Universe Survey (PSS) for the past three cycles.

The implicit stratification of the schools involved four dimensions. Within each explicit stratum, the private schools were hierarchically sorted by census division, urbanicity status, race/ethnicity status, and estimated grade enrollment. The implicit stratification in this four-fold hierarchical stratification was achieved via a "serpentine sort."

Census division was used as the first level of implicit stratification for the NAEP 2011 private school sample.

Collapsing of census division varied by grade. For grade 4, all nine census divisions were used for stratifying Catholic and other private schools. However, due to small cell sizes, divisions in the Northeast and Midwest were collapsed within census regions for Conservative Christian schools. For Lutheran schools, a South Central stratum was created within the southern region and divisions were collapsed across regions to create an East Coast stratum. For grade 8, all census divisions were used to stratify Catholic and other private schools. Divisions in the Northeast were collapsed within region for both Conservative Christian and Lutheran schools. Additionally for Lutheran schools, two divisions were collapsed within the southern region to create a South Central stratum.

The next level of stratification was an urbanicity classification based on urban-centric locale, as specified on the PSS. Within a census division-based stratum, urban-centric locale cells that were too small were collapsed. The criterion for adequacy was that the cell had to have an expected school sample size of at least six.

The urbanicity variable was equal to the original urban-centric locale if no collapsing was necessary to cover an inadequate original cell. If collapsing was necessary, the scheme was to first collapse within the four major strata (city, suburbs, town, and rural). For example, if the expected number of large city schools sampled was less than six, large city was collapsed with midsize city. If the collapsed cell was still inadequate, they were further collapsed with small city. If a major stratum cell (all three cells collapsed together) was still deficient, it was collapsed with a neighboring major stratum cell. For example, city would be collapsed with suburbs.

The last stage of stratification was a division of the geographic/urbanicity strata into race/ethnicity strata if the expected number of schools sampled was large enough (i.e., at least equal to 12). This was done by deciding first on the number of race/ethnicity strata and then dividing the geography/urbanicity stratum into that many pieces. The school frame was sorted by the percentage of students in each school who were Black, Hispanic, or American Indian. The three race/ethnic groups defining the race/ethnicity strata were those that have historically performed substantially lower on NAEP assessments than White students. The sorted list was then divided into pieces, with roughly an equal expected number of sampled schools in each piece.

Finally, schools were sorted within stratification cells by estimated grade enrollment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_delta_strat.aspx

Student Sample Selection for the 2011 Private School National Assessment

The target student sample size within sampled schools for the fourth and eighth grades was 63 students. However, schools with 70 or fewer students automatically had all students sampled. In addition, at grade 4 only, a school that had more than 70 students but fewer than 121 could choose to have all students sampled.

There was only one spiral type for each grade. The percentage of booklets by subject within the spiral for each grade is given below.

Percentage of booklets, private school national assessment, by subject within the spiral and grade: 2011

Grade	Mathematics	Reading	Science	KaSA	Pilot
4	46.96	48.62	†	1.55	2.87
8	48.34	46.69	4.97	†	†

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

The process of student list submission, sampling students from year-round schools, sampling new enrollees, and determining student eligibility and exclusion status was the same as for the state NAEP student sample.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_delta_studsamp.aspx

Substitute Schools for the 2011 Private School National Assessment

Substitutes were preselected for the private school samples by sorting the school frame file according to the actual order used in the sampling process (the implicit stratification). Each sampled school had its two nearest neighbors on the school frame file identified as potential substitutes. As the last sort ordering was by grade enrollment, the nearest neighbors had grade enrollment values very close to that of the sampled school.

Schools were disqualified as potential substitutes if they were already selected in the private school sample or assigned as a substitute for another private school (earlier in the sort ordering). Schools assigned as substitutes for eighth-grade schools were disqualified as potential substitutes for fourth-grade schools.

If both nearest neighbors were still eligible to be substitutes, the one with the closer grade enrollment was chosen. If both nearest neighbors had the same grade enrollment (an uncommon occurrence), one of the two was randomly selected.

In the process described above, only schools with the same affiliation were selected as substitutes.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_nat_delta_subs.aspx

Target Population for the 2011 Private School National Assessment

The target population for the 2011 Private School National Assessment included all students enrolled in private schools in grades 4 and 8 within the 50 states and the District of Columbia.

Sampling Frame for
the Private School
National Assessment

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_targpop.aspx

2011 Fourth- and Eighth-Grade Public School National Assessment

For the mathematics, reading, and science assessments in fourth- and eighth-grade public schools, the national samples were the state assessment samples for each jurisdiction. All states participated in the mathematics, reading, and science assessments. By design, only Bureau of Indian Education (BIE) schools did not participate in the state science assessment, as it lacked the required number of students for the state science assessment. A small portion of students received the science assessment in BIE schools in science to supplement the national science sample.

Additional details of the national science sample are also described as part of the state assessment samples.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_pub_gr_4_8.aspx

2011 Writing Computer-Based Assessment (WCBA)

The sample design for the NAEP 2011 writing computer-based assessment (WCBA) provided a nationally representative sample of eighth- and twelfth-grade students.

This was accomplished by designing separate sample components for public and private schools. The selected samples were based on a three-stage sample design:

- selection of primary sampling units (PSUs),
- selection of schools within strata, and
- selection of students within schools.

The samples of schools were selected with probability proportional to a measure of size based on the estimated eighth- and twelfth-grade enrollment in the schools.

The target population respectively included all students in public and private schools, Bureau of Indian Education (BIE) schools, and Department of Defense Education Activity (DoDEA) schools in the 50 states and the District of Columbia, who were enrolled in the eighth and twelfth grade at the time of assessment.

The table below shows the target student sample sizes of assessed students for each sample.

Selection of Primary Sampling Units
Public School 2011 Writing
Computer-Based Assessment (WCBA)
Private School 2011 Writing
Computer-Based Assessment (WCBA)
School and Student Participation
Results for the 2011 Writing
Computer-Based Assessment

Target student sample sizes of assessed students for grades 8 and 12, writing computer-based assessment (WCBA), by school type: 2011

School type	Grade	Target student sample size
Public	8,12	19,800
Private	8,12	2,200

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Writing Computer-Based Assessment.

To reduce the burden on any particular school, efforts were made to minimize the overlap between the 2011 PSU sample and all other PSU samples selected for NAEP since 2006. The school samples were designed to have minimum overlap with both the United States school sample for the Trends in International Mathematics and Science Study (TIMSS), and the NAEP 2011 state sample.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_wcba.aspx

Private School 2011 Writing Computer-Based Assessment (WCBA)

The NAEP 2011 writing computer-based assessment (WCBA) sample design yielded nationally representative samples of private school students in grades 8 and 12 through a three-stage approach: selection of primary sampling units (PSUs), selection of schools within strata, and selection of students within schools. The sample of schools was selected with probability proportional to a measure of size based on the estimated grade enrollment in the schools.

The 2011 national WCBA sampling plan had a goal of assessing 2,200 eighth-graders and 2,200 twelfth-graders. Target sample sizes were adjusted to reflect expected private school and student response and eligibility.

Schools on the sampling frame were explicitly stratified prior to sampling by private school affiliation (Catholic, non-Catholic, and unknown affiliation). Within affiliation type, schools were implicitly stratified by PSU type (certainty/noncertainty). In certainty PSUs, further stratification was by census region, urban-centric locale, and estimated grade enrollment. In noncertainty PSUs, additional stratification was by PSU stratum, urban-centric locale, and estimated grade enrollment.

From the stratified frame of private schools, systematic random samples of eighth- and twelfth-grade schools were drawn with probability proportional to a measure of size based on the estimated grade enrollment of the school in the relevant grade.

Each selected school in the private school sample provided a list of eligible enrolled students from which a systematic, equal probability sample of students was drawn.

Target Population

Sampling Frame

Stratification of Schools

Sampling of Schools

Substitute Schools

Ineligible Schools

Student Sample Selection

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_priv_wbca.aspx

Ineligible Private Schools for the 2011 Writing Computer-Based Assessment (WCBA)

The Private School Universe Survey (PSS) school file from which most of the sampled schools were drawn corresponds to the 2007-2008 school year, 3 years prior to the assessment school year. During the intervening period, some of these schools either closed, no longer offered the grade of interest, or were ineligible for other reasons. In such cases, the sampled schools were coded as ineligible.

The table below presents unweighted counts of sampled private schools by eligibility status, including the reason for ineligibility.

Number of sampled private schools, writing computer-based assessment (WCBA), by eligibility status and grade: 2011

Eligibility status	Unweighted count of schools	Unweighted percentage
All eighth-grade sampled private schools	157	100.00
Eligible schools	140	89.17
No eligible students in grade	3	1.91
Does not have grade	4	2.55
School closed	8	5.10
Not a regular school	1	0.64
Other ineligible school	1	0.64
Duplicate on sampling frame	0	0.00
All twelfth-grade sampled private schools	177	100.00
Eligible schools	160	90.40
No eligible students in grade	4	2.26
Does not have grade	4	2.26
School closed	2	1.13
Not a regular school	4	2.26
Other ineligible school	2	1.13
Duplicate on sampling frame	1	0.56

NOTE: Detail may not add up to totals due to rounding. Percentages are based on rounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

The table below presents unweighted counts of sample private schools by collapsed private school type and eligibility status.

Number of sampled private schools, writing computer-based assessment (WCBA), by eligibility status and private school type: 2011

Private school type	Eligibility status	Unweighted count of schools	Unweighted percentage
All eighth-grade sampled private schools	Total	157	100.00
	Eligible	140	89.17
	Ineligible	17	10.83
Catholic	Total	50	100.00
	Eligible	42	84.00
	Ineligible	8	16.00
Non-Catholic	Total	107	100.00
	Eligible	98	91.59
	Ineligible	9	8.41
All twelfth-grade sampled private schools	Total	177	100.00
	Eligible	160	90.40
	Ineligible	17	9.60
Catholic	Total	55	100.00
	Eligible	55	100.00
	Ineligible	0	0.00
Non-Catholic	Total	122	100.00
	Eligible	105	86.07
	Ineligible	17	13.93

NOTE: Detail may not add up to totals due to rounding. Percentages are based on rounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

Sampling Frame for the Private School 2011 Writing Computer-Based Assessment (WCBA)

The sampling frame for private schools was developed from the 2007-2008 Private School Universe Survey (PSS), a survey conducted by the U.S. Census Bureau for the National Center for Education Statistics (NCES). The PSS is a biennial mail survey of all private schools in the 50 states and the District of Columbia. The PSS frame of schools comprises both a list frame and an area frame. The list frame is an assembly of the 2005-2006 PSS frame and more up-to-date lists from state education agencies, private school associations, and other easily accessible sources. To improve the coverage of the PSS list frame, the Census Bureau also conducted a survey to locate private schools in a random sample of geographic areas throughout the United States. The areas were single counties or groups of counties sampled from an area frame constructed from all counties in the nation. Within each selected area a complete list of private schools was gathered using information from telephone directories, religious institutions, local education agencies, chambers of commerce, and local government offices. Schools not already on the list frame were identified and added to the frame of private schools. A weighting component was computed by the Census Bureau so that the additional area-frame schools would represent all schools absent from the list frame, not just those in the selected areas.

The sampling frame was restricted to schools located in the primary sampling units (PSUs) selected for the NAEP 2011 writing computer-based assessment (WCBA). In addition, the sampling frame excluded ungraded schools, vocational schools with no enrollment, special-education-only schools, homeschool entities, prison and hospital schools, and juvenile correctional institutions.

For quality control purposes, school and student counts from the sampling frame were compared to school and student counts from previous private school frames by grade. No major discrepancies were found.

Eighth- and
Twelfth-Grade
Schools and
Enrollment in the
Private School WCBA
Sampling Frame

New-School Sampling
Frame for the Private
School Writing
Computer-Based
Assessment

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_priv_wcba_sampframe.aspx

Eighth- and Twelfth-Grade Schools and Enrollment in the Private School WCBA Sampling Frame

The following table presents the number of schools and estimated enrollment for the private school frame for grades 8 and 12. These enrollment numbers include only those schools with known affiliation. The unweighted estimated enrollment is restricted to the selected primary sampling units (PSUs). The weighted estimated enrollment incorporates the PSU weight (inverse of the probability of selecting the PSU), as well as the Private School Universe Survey (PSS) weight, and thus is a national estimate of the number of private school students in each grade.

Number of schools and enrollment in private school sampling frame for the writing computer-based assessment (WCBA), by school affiliation and grade: 2011

Grade	Affiliation	Number of schools	Estimated enrollment (unweighted)	Estimated enrollment (weighted)
8	Total	9,366	234,221	374,445
	Catholic	3,438	112,863	169,638
	Non-Catholic	5,928	121,358	204,807
12	Total	4,539	213,828	338,291
	Catholic	780	111,164	158,660
	Non-Catholic	3,759	102,664	179,631

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_sampfme_gr_8_12_wcba.aspx

New-School Sampling Frame for the Private School Writing Computer-Based Assessment (WCBA)

Whereas the Private School Universe Survey (PSS) file used for the frame corresponds to the 2007-2008 school year, the NAEP assessment year was the 2010-2011 school year. During this 3-year period, some schools closed, some changed their grade span, and still others came into existence.

To achieve as close to full coverage as possible, the private school frame for the writing computer-based assessment (WCBA) was supplemented by a sample of new Catholic schools. The goal was to allow every such school a chance of selection, thereby fully covering the target population of Catholic schools in operation during the 2010-2011 school year. The first step in this process was the development of a new-school frame through the construction of a diocese-level file from the PSS school-level file. To develop the frame, the diocese-level file was divided into two files: one for small dioceses and a second for medium and large dioceses.

Small dioceses contained no more than three schools on the frame in total, with no more than one school at each grade (fourth, eighth, and twelfth). New schools in small dioceses were identified during school recruitment and added to the sample if the old school in the same diocese was sampled at the relevant grade. From a sampling perspective, the new school was viewed as an "annex" to the sampled school that had a well-defined probability of selection equal to that of the old school. The "frame" in this case was, in fact, the original frame; when the old school was sampled in a small diocese, the new school was automatically sampled as well.

To limit respondent burden and keep the level of effort within reasonable bounds, the new-school frame was created using information obtained from a sample of the remaining dioceses. The remaining dioceses were separated into two strata of large- and medium-size dioceses. These strata were defined by computing the percentage of the nation's total Catholic school enrollment each diocese represents, sorting the dioceses in descending order by that percentage, and cumulating the percentages across the sorted file. All dioceses up to and including the first diocese at or above the 80th cumulative percentage were defined as large dioceses. The remaining dioceses were defined as medium dioceses.

A simplified example is given below. Dioceses are ordered by percentage enrollment. The first six become large dioceses and the last six become medium dioceses.

Example showing assignment of Catholic dioceses to the large and medium strata, private school assessment: 2011

Diocese	Percent enrollment	Cumulative percentage enrollment	Stratum
Diocese 1	20	20	L
Diocese 2	20	40	L
Diocese 3	15	55	L
Diocese 4	10	65	L
Diocese 5	10	75	L
Diocese 6	10	85	L
Diocese 7	5	90	M
Diocese 8	2	92	M
Diocese 9	2	94	M
Diocese 10	2	96	M
Diocese 11	2	98	M
Diocese 12	2	100	M

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

In actuality there were 71 large and 103 medium dioceses in the sampling frame.

The target sample size was 10 dioceses total: 8 large and 2 medium. In the medium stratum, the dioceses were selected with equal probability. In the large stratum, dioceses were sampled with probability proportional to enrollment. These probabilities were retained and used in all later stages of sampling and weighting in order to represent all dioceses, whether or not they had been sampled to be surveyed for new schools.

Each selected diocese was sent a listing of its schools extracted from the 2007-2008 PSS file and was asked to provide information about new schools and any changes to grade span in existing schools. This information provided by the selected dioceses was used to create sampling frames for the selection of new Catholic schools. The process of obtaining the information was conducted with the help of the National Catholic Educational Association (NCEA). NCEA was sent the school lists for the 10 sampled dioceses and was responsible for returning the completed updates.

The eligibility of a new school at a particular grade was determined by its grade span. A school already on PSS also was classified as "new" if a change of grade span had occurred such that the school status changed from ineligible to eligible at a particular grade.

As was done for the original sampling frame, the new-school sampling frame was restricted to schools located in the primary sampling units (PSUs) selected for the NAEP 2011 WCBA. Weights for schools in the new-school sample were adjusted to account for the PSU selection probability.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_newschoolframe_wcba.aspx

Sampling of Private Schools for the 2011 Writing Computer-Based Assessment (WCBA)

The writing computer-based assessment (WCBA) private school sample was selected with probability proportional to size using systematic sampling from a sorted list. A school's measure of size was a complex function of the school's estimated grade enrollment.

Schools were ordered within each grade using the serpentine sort described under the stratification of private schools. A systematic sample was then drawn using this serpentine sorted list and the measures of size.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_priv_wcba_schlsamp.aspx

Computation of
Measures of Size for
the 2011 Private
School Writing
Computer-Based
Assessment

School Sample Sizes
for 2011 Private
School Writing
Computer-Based
Assessment: Frame
and New School

Computation of Measures of Size for the 2011 Private School Writing Computer-Based Assessment (WCBA)

In the design of each school sample, five objectives underlie the process of determining the probability of selection for each school and how many students are to be sampled from each selected school containing grade-eligible students:

- to meet the target student sample size;
- to select an equal-probability sample of students;
- to limit the number of students who are selected from a school;
- to ensure that the sample within a school does not include a very high percentage of the students in the school, unless all students are included; and
- to reduce the rate of sampling of small schools, in recognition of the greater cost and burden per student of conducting assessments in such schools.

The goal in determining the school's measure of size is to optimize across the last four objectives in terms of maintaining the accuracy of estimates and the cost-effectiveness of the sample design. The following algorithm was used to assign a measure of size to each school based on its estimated grade enrollment as indicated on the sampling frame.

The measures of size vary by enrollment size. The initial measures of size (MOS) were set as follows, for both eighth and twelfth grades:

$$MOS_{js} = PSU_{WT} \times \begin{cases} x_{js} & \text{if } 30 < x_{js} \\ 30 & \text{if } 20 < x_{js} \leq 30 \\ 1.5 \times x_{js} & \text{if } 5 < x_{js} \leq 20 \\ 7.5 & \text{if } x_{js} \leq 5 \end{cases}$$

where x_{js} is the estimated grade enrollment for grade j ($j = 8, 12$) in school s , PSU_{WT_s} = the Private School Universe Survey (PSS) area frame weight for school s , computed by the U.S. Census Bureau, and PSU_{WT_j} = the primary sampling unit (PSU) weight for school s .

An adjustment to the initial measure of size was made for some schools. Schools in the PSU containing Honolulu County had their measure of size increased by a factor of two in order to double their probability of selection.

The school measure of size was then rescaled to create an expected number of hits by applying a multiplicative constant b_j , which varies by grade and school type. For the national WCBA sample, by design, a school could not be selected or "hit" in the sampling process more than once.

The rescaled measure of size, E_{js} , was defined as:

$$E_{js} = \min(b_j \times MOS_{js}, 1)$$

For grade 8 only, a final adjustment was made to the measures of size (E_{js}) in the national sample to attempt to reduce school burden by minimizing the number of schools that were selected for simultaneous administration of the WCBA, the operational private school assessments (mathematics, reading, and science), and the Trends in International Mathematics and Science Study (TIMSS). The NAEP 2011 studies for grade 8 used an adaptation of the Keyfitz process to compute conditional measures of size that, by their design, minimized the overlap of schools selected for the three types of assessment. Grade 12 did not have any operational assessments or a TIMSS sample in 2011.

The school's probability of selection π_{js} was given by:

$$\pi_{js} = \min(E_{js}, 1)$$

One can choose a value of b_j such that the expected overall student sample yield matches the desired targets specified by the design, where the expected yield is calculated by summing the product of an individual school's probability and its student sample yield across all schools in the frame.

In addition, new and newly eligible schools were sampled from the new-school frame. The assigned measures of size for these schools,

$$E_{js} = \min(b_j \times MOS_{js} \times \pi_{djs}^{-1}, 1)$$

used the b_j value from the main school sample for the grade and school type (i.e., the same sampling rates as for the main school sample). The variable π_{djs} is the probability of selection of the diocese into the new-school diocese (d) sample.

School Sample Sizes for 2011 Private School Writing Computer-Based Assessment (WCBA): Frame and New School

The following table presents the number of schools selected from the private school WCBA sampling frame (constructed from the Private School Universe Survey file) and the new-school sampling frame, for eighth and twelfth grade, by school type.

NAEP private school frame-based and new school writing computer-based assessment (WCBA) samples, by grade and school type: 2011

Grade and private school type	Total school sample	Frame school sample	New school sample
Eighth grade			
All private	157	155	2
Catholic	50	48	2
Non-Catholic	106	106	0
Unknown affiliation	1	1	0
Twelfth grade			
All private	177	177	0
Catholic	55	55	0
Non-Catholic	120	120	0
Unknown affiliation	2	2	0

NOTE: Details may not sum to totals due to rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_delta_schlsamp_sampsize_wcba.aspx

Stratification of Private Schools for the 2011 Writing Computer-Based Assessment (WCBA)

Prior to stratification, the private school sampling frame was divided into grade-specific files, one each for eighth and twelfth grade. For each such grade-specific file, schools were explicitly stratified by private school affiliation (Catholic, non-Catholic, and unknown affiliation). Private school affiliation was unknown for nonrespondents to the NCES Private School Universe Survey (PSS). Within private school type, separate implicit stratification schemes were used to sort schools in certainty primary sampling units (PSUs) and noncertainty PSUs. In all cases, the implicit stratification was achieved via a serpentine sort.

Within each certainty PSU, the schools were hierarchically sorted by

- census region,
- urbanization classification (urban-centric locale), and
- estimated grade enrollment.

Schools in noncertainty PSUs were hierarchically sorted by

- PSU stratum,
- urbanization classification (urban-centric locale), and
- estimated grade enrollment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_priv_wcba_strat.aspx

Student Sample Selection for the Private School 2011 Writing Computer-Based Assessment (WCBA)

For the NAEP 2011 writing computer-based assessment (WCBA), the target student sample sizes within sampled schools were the same for both eighth and twelfth grades. All students were sampled if the school had 30 or fewer students in that grade. Otherwise, a sample of 30 students was selected without replacement.

The process of list submission, sampling new enrollees, and determining student eligibility and exclusion status was the same as the process used for the NAEP 2011 state student samples.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_priv_science_wcba_studsamp.aspx

Substitute Private Schools for the 2011 Writing Computer-Based Assessment (WCBA)

Substitutes were preselected for the private school samples by sorting the school frame file according to the actual order used in the sampling process (the implicit stratification). For operational reasons, the original selection order was embedded within the sampled primary sampling unit (PSU). Each sampled school had each of its nearest neighbors within the same sampling stratum on the school frame file identified as a potential substitute. When grade enrollment was used as the last sort ordering variable, the nearest neighbors had grade enrollment values very close to that of the sampled school. This was done to facilitate the selection of about the same number of students within the substitute as would have been selected from the original sampled school.

Schools were disqualified as potential substitutes if they were already selected in any of the original private school samples or assigned as a substitute for another private school (earlier in the sort ordering), or if they were already selected in the original 2011 Trends in International Mathematics and Science Study (TIMSS) sample. TIMSS substitutes were eligible to be used as substitutes for the writing computer-based assessment (WCBA). Schools assigned as substitutes for twelfth-grade schools were disqualified as potential substitutes for eighth-grade schools.

If both nearest neighbors were still eligible to be substitutes, the one with a closer grade enrollment was chosen. If both nearest neighbors were equally distant from the sampled school in their grade enrollment (an uncommon occurrence), one of the two was randomly selected. If the grade enrollment of the nearest neighbor school was less than half of the expected student sample size of the original sampled school, then it was considered ineligible as a substitute for that school.

Of the approximately 330 originally sampled private schools for the WCBA, about 100 had a substitute activated because the original school, although eligible, did not participate. Ultimately, about 40 substitute private schools participated in the WCBA.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_priv_wcba_subs.aspx

Target Population of the Private School 2011 Writing Computer-Based Assessment (WCBA)

The target population for the private school 2011 writing computer-based assessment (WCBA) included all students who were enrolled in eighth and twelfth grades in private schools. The sample frame included private schools in the 50 states and the District of Columbia.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_priv_wcba_targpop.aspx

Public School 2011 Writing Computer-Based Assessment (WCBA)

The NAEP 2011 writing computer-based assessment (WCBA) sample design yielded nationally representative samples of public school students in each grade (grades 8 and 12) through a three-stage approach: selection of primary sampling units (PSUs), selection of schools within strata, and selection of students within schools. The sample of schools was selected with probability proportional to a measure of size based on the estimated grade enrollment in the schools.

The 2011 WCBA was administered in both grades 8 and 12, with the goal of assessing 19,800 students in each grade. The target sample size was adjusted to reflect expected public school and student response and eligibility.

Schools on the sampling frame were explicitly stratified prior to sampling by PSU type (certainty/noncertainty). Within certainty PSUs, schools were implicitly stratified by census region, urban-centric locale and median household income in the zip code area where the school is located. Within noncertainty PSUs, schools were implicitly stratified by PSU stratum, urban-centric locale, and median income in the zip code area where the school is located.

From the stratified frame of public schools, systematic random samples of eighth- and twelfth-grade schools were drawn with probability proportional to a measure of size based on the estimated grade enrollment of the school, in the relevant grade.

Each selected school in the public school samples provided a list of eligible enrolled students from which a systematic, equal probability sample of students was drawn.

Target Population

Sampling Frame

Stratification of Schools

Sampling of Schools

Substitute Schools

Ineligible Schools

Student Sample Selection

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_science_pub_wcba.aspx

Ineligible Public Schools for the 2011 Writing Computer-Based Assessment (WCBA)

The Common Core of Data (CCD) public school file from which most of the sampled schools were drawn corresponds to the 2007-2008 school year, 3 years prior to the assessment school year. During the intervening period, some of these schools either closed, no longer offered the grade of interest, or became ineligible for other reasons. In such cases, the sampled schools were considered to be ineligible.

The table below presents unweighted counts of sampled public schools by grade and eligibility status, including the reason for ineligibility.

Number of sampled public schools, writing computer-based assessment (WCBA), by eligibility status and grade: 2011

Eligibility status	Unweighted count of schools	Unweighted percentage
All eighth-grade sampled public schools	890	100.00
Eligible schools	841	94.49
No eligible students in grade	1	0.11
Does not have grade	12	1.35
School closed	28	3.15
Not a regular school	8	0.90
Other ineligible school	0	0.00
Duplicate on sampling frame	0	0.00
All twelfth-grade sampled public schools	1,200	100.00
Eligible schools	1,100	94.57
No eligible students in grade	3	0.25
Does not have grade	13	1.09
School closed	16	1.34
Not a regular school	30	2.51
Other ineligible school	3	0.25
Duplicate on sampling frame	0	0.00

NOTE: Detail may not add up to totals due to rounding. Percentages are based on rounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_pub_wcba_inelg.aspx

Sampling Frame for the Public School 2011 Writing Computer-Based Assessment (WCBA)

The sampling frame for public schools was derived from the Common Core of Data (CCD) file corresponding to the 2007-2008 school year. The CCD files provided the frame for all regular public, state-operated public, Bureau of Indian Education (BIE), and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) open during the 2007-2008 school year.

The sampling frame was restricted to schools located in the primary sampling units (PSUs) selected for the NAEP 2011 writing computer-based assessment (WCBA). The sampling frame also excluded ungraded schools, vocational schools with no enrollment, special-education-only schools, homeschool entities, prison or hospital schools, and juvenile correctional institutions.

For quality control purposes, school and student counts from the sampling frame were compared to school and student counts from previous public school frames by grade. No major discrepancies were found.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_pub_science_wcba_sampfrme.aspx

Eighth- and
Twelfth-Grade
Schools and
Enrollment in the
Public School WCBA
Sampling Frame

New-School Sampling
Frame for the Public
School Writing
Computer-Based
Assessment

Eighth- and Twelfth-Grade Schools and Enrollment in the Public School WCBA Sampling Frame

The following table presents the number of schools and estimated enrollment for the public school frame for grades 8 and 12. The unweighted estimated enrollment is restricted to the selected primary sampling units (PSUs). The weighted estimated enrollment incorporates the PSU weight (inverse of the probability of selecting the PSU), and thus is a national estimate of the number of public school students in each grade.

Number of schools and enrollment in public school sampling frame, writing computer-based assessment (WCBA), by grade: 2011

Grade	School count in sampled PSUs	Estimated enrollment (unweighted)	Estimated enrollment (weighted)
8	11,379	1,952,079	3,635,336
12	9,068	1,833,707	3,423,860

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_pub_sampfrme_gr_8_12.aspx

New-School Sampling Frame for the Public School Writing Computer-Based Assessment (WCBA)

The Common Core of Data (CCD) file used for the frame corresponds to the 2007-2008 school year, whereas the assessment year is the 2010-2011 school year. During this 3-year period, some schools closed, some changed structure (one school becoming two schools, for example), and others came into existence.

To achieve as close to full coverage as possible, the writing computer-based assessment (WCBA) school frame was supplemented by a sample of new schools obtained from a sample of districts. Each sampled district was sent a list of the CCD schools and asked to add in any new schools or old schools that had become newly eligible for eighth or twelfth grades.

Since asking every school district to list new- and newly-eligible schools would have generated too much of a burden, a sample of districts was contacted to obtain a list of new schools. To represent the unsampled districts in the full sample of schools, weights for schools included in the new-school sample were adjusted to reflect the district selection probability.

As was done for the original sampling frame, the new-school sampling frame was restricted to schools located in the primary sampling units (PSUs) selected for the NAEP 2011 WCBA. Weights for schools in the new-school sample were further adjusted to account for the PSU selection probability.

The goal was to allow every new school a chance of selection, thereby fully covering the target population of schools in operation during the 2008-2009 school year. The first step in this process was the development of a new-school frame through the construction of a district-level file from the CCD school-level file. To develop the frame, the district-level file was divided into two files: one for small districts and a second for medium and large districts.

Small districts contained no more than three schools on the frame in total, with no more than one school at each targeted grade (fourth, eighth, and twelfth). New schools in small districts were identified during school recruitment and added to the sample if the old school was sampled. From a sampling perspective, the new school was viewed as an "annex" to the sampled school that had a well-defined probability of selection equal to that of the old school. The "frame" in this case was, in fact, the original frame; when the old school was sampled in a small district, the new school was automatically sampled as well.

The remaining districts were defined as medium and large districts. In these districts, a frame of new schools was developed based on information provided by the district. To limit the required effort, the new-school frame was created through developing information on a sample of medium and large public school districts in each jurisdiction. All districts were selected in the following classes of districts:

- jurisdictions where all schools were sampled with certainty at either grade 8 or 12 (so that all new schools would be selected with certainty, as well),
- state-operated districts,
- districts in states with fewer than 10 districts,
- districts containing no schools other than charter schools, and
- TUDA districts.

The remaining districts in each jurisdiction (excepting the certainty jurisdictions) were separated into two strata of large- and medium-size districts. These strata were defined by computing an aggregate percentage of enrollment for each district within the state (removing districts in the certainty strata defined above) and sorting in descending order by percentage of jurisdiction enrollment represented by the district. All districts up to and including the first district at or above the 80th cumulative percentage were defined as large districts. The remaining districts were defined as medium districts.

An example is given below. A state's districts are ordered by percentage enrollment. The first six become large districts and the last six become medium districts.

Large and medium districts example, state assessment, by enrollment, stratum, and district: 2011

District	Percentage enrollment	Cumulative percentage enrollment	Stratum
1	20	20	L
2	20	40	L
3	15	55	L
4	10	65	L
5	10	75	L
6	10	85	L
7	5	90	M
8	2	92	M
9	2	94	M
10	2	96	M
11	2	98	M
12	2	100	M

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011.

The target sample size for each jurisdiction was 10 districts. Where possible, we selected 8 large and 2 medium districts. However, in the example above, since there are only 6 large districts, all of the large districts and 4 of the medium districts were selected for the new-school inquiry.

If sampling was needed in the medium stratum (i.e., it was not a certainty jurisdiction), the medium districts were selected with equal probability. If sampling was needed in the large stratum, the large districts were sampled with probability proportional to enrollment. These probabilities were retained and used in all later stages of sampling and weighting, as the district probability then represented the number of other districts that were not sampled to be surveyed for new schools.

The selected districts in each jurisdiction were then sent a listing of all their schools that appeared on the 2007-2008 CCD file and were asked to provide information about the new schools not included in the file and grade span changes of existing schools. These listings provided by the selected districts were used as sampling frames for selection of new public schools and updates of existing schools. This process was conducted through the NAEP State Coordinator in each jurisdiction. The coordinators were sent the information for all sampled districts in their respective states and were responsible for returning the completed updates.

The eligibility of a school was determined based on the grade span. A school also was classified as "new" if a change of grade span had occurred such that the school status changed from ineligible to eligible in a particular grade.

Sampling of Public Schools for the 2011 Writing Computer-Based Assessment (WCBA)

The writing computer-based assessment (WCBA) public school sample was selected with probability proportional to size, using systematic sampling from a sorted list. A school's measure of size was a complex function of the school's estimated grade enrollment.

Schools were ordered within each grade, using the serpentine sort described under the stratification of public schools. A systematic sample was then drawn using this serpentine-sorted list and the measures of size.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_pub_wcba_schlsamp.aspx

Computation of
Measures of Size for
the 2011 Public
School Writing
Computer-Based
Assessment

School Sample Sizes
for 2011 Public
School WCBA: Frame
and New School

Computation of Measures of Size for the 2011 Public School Writing Computer-Based Assessment (WCBA)

In the design of each school sample, five objectives underlie the process of determining the probability of selection for each school and how many students are to be sampled from each selected school containing grade-eligible students:

- to meet the target student sample size;
- to select an equal-probability sample of students;
- to limit the number of students who are selected from a school;
- to ensure that the sample within a school does not include a very high percentage of the students in the school, unless all students are included; and
- to reduce the rate of sampling of small schools, in recognition of the greater cost and burden per student of conducting assessments in such schools.

The goal in determining the school's measure of size is to optimize across the last four objectives in terms of maintaining the accuracy of estimates and the cost-effectiveness of the sample design. The following algorithm was used to assign a measure of size to each school based on its estimated grade enrollment as indicated on the sampling frame.

The measures of size vary by enrollment size. The initial measures of size (MOS) were set as follows, for both eighth and twelfth grades:

$$MOS_{js} = PSU_{WT} \times \begin{cases} x_{js}, & \text{if } 30 < x_{js} \\ 30, & \text{if } 20 < x_{js} \leq 30 \\ 1.5 \times x_{js}, & \text{if } 5 < x_{js} \leq 20 \\ 7.5, & \text{if } x_{js} \leq 5 \end{cases}$$

where x_{js} is the estimated grade enrollment for grade j ($j = 8, 12$) in school s , and PSU_{WT} is the primary sampling unit (PSU) weight for school j .

An adjustment to the initial measure of size was made for some schools. Schools with a high percentage of Black or Hispanic students, and schools in the PSU containing Honolulu County, had their measure of size increased by a factor of two, in order to double their probability of selection.

The school measure of size was then rescaled to create an expected number of hits by applying a multiplicative constant b_j , which varies by grade. For the national writing computer-based assessment (WCBA) sample, by design, a school could not be selected or "hit" in the sampling process more than once.

The rescaled measure of size, E_{js} , was defined as:

$$E_{js} = \min(b_j \times MOS_{js}, 1)$$

A final adjustment was made to the measures of size (E_{js}) in the national sample to attempt to reduce school burden by minimizing the number of schools selected for simultaneous administration of both the state and national studies. The NAEP 2011 studies used an adaptation of the Keyfitz process to compute conditional measures of size that, by their design, minimized the number of schools selected for the national study (WCBA) that were also selected for the state assessment or the Trends in International Mathematics and Science Study (TIMSS).

The school's probability of selection π_{js} was given by:

$$\pi_{js} = \min(E_{js}, 1)$$

One can choose a value of b_j such that the expected overall student sample yield matches the desired targets specified by the design, where the expected yield is calculated by summing the product of an individual school's probability and its student sample yield across all schools in the frame.

In addition, new and newly eligible schools were sampled from the new-school frame. The assigned measures of size for these schools,

$$E_{js} = \min(b_j \times MOS_{js} \times \pi_{djs}^{-1}, 1)$$

used the b_j value from the main school sample for the grade (i.e., the same sampling rates as for the main school sample). The variable π_{djs} is the probability of selection of the district into the new-school district (d) sample.

School Sample Sizes for 2011 Public School WCBA: Frame and New School

The following table presents the number of schools selected for the 2011 public school writing computer-based assessment from the public school sampling frame and the new school sampling frame, for grades 8 and 12.

NAEP public school WCBA frame-based and new school samples, by grade: 2011

Grade	Total school sample	Frame school sample	New school sample
8	890	866	24
12	1,200	1,200	12

NOTE: Details may not sum to totals due to rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_pub_schlsamp_sampsizes_wcba.aspx

Stratification of Public Schools for the 2011 Writing Computer-Based Assessment (WCBA)

Prior to stratification, the public school sampling frame was divided into grade-specific files, one each for eighth and twelfth grade. For each grade-specific frame file, separate implicit stratification schemes were used to sort schools into certainty primary sampling units (PSUs) and noncertainty PSUs. In all cases, the implicit stratification was achieved via a "serpentine sort."

For certainty PSUs, the schools were hierarchically sorted by

- census region,
- urbanization classification (urban-centric locale), and
- median household income in the zip code area where the school is located.

Schools in noncertainty PSUs were hierarchically sorted by

- PSU stratum,
- urbanization classification (urban-centric locale), and
- median household income in the zip code area where the school is located.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_pub_wcba_strat.aspx

Student Sample Selection for the Public School 2011 Writing Computer-Based Assessment (WCBA)

For the NAEP 2011 writing computer-based assessment (WCBA), the target student sample sizes within sampled schools were the same for both eighth and twelfth grades. All students were sampled if the school had 30 or fewer students in that grade. Otherwise, a sample of 30 students was selected without replacement.

The process of list submission, sampling students from year-round schools, sampling new enrollees, and determining student eligibility and exclusion status was the same as the process used for the NAEP 2011 state student samples.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_pub_science_wcba_studsamp.aspx

Substitute Public Schools for the 2011 Writing Computer-Based Assessment (WCBA)

Substitutes were preselected for the public school samples by sorting the school frame file according to the actual order used in the sampling process (the implicit stratification). For operational reasons, the original selection order was embedded within the sampled primary sampling unit (PSU) and state. Each sampled school had each of its nearest neighbors within the same sampling stratum on the school frame file identified as a potential substitute. When grade enrollment was used as the last sort ordering variable, the nearest neighbors had grade enrollment values very close to that of the sampled school. This was done to facilitate the selection of about the same number of students within the substitute as would have been selected from the original sampled school.

Schools were disqualified as potential substitutes if they were already selected in any of the original public school samples or assigned as a substitute for another public school (earlier in the sort ordering), or if they were already selected in the original 2011 Trends in International Mathematics and Science Study (TIMSS) sample. TIMSS substitutes could be used as substitutes for the writing computer-based assessment (WCBA). Schools assigned as substitutes for twelfth-grade schools were disqualified as potential substitutes for eighth-grade schools.

If both nearest neighbors were still eligible to be substitutes, the one with a closer grade enrollment was chosen. If both nearest neighbors were equally distant from the sampled school in their grade enrollment (an uncommon occurrence), one of the two was randomly selected. If the grade enrollment of the nearest neighbor school was less than half of the expected student sample size of the original sampled school, then it was considered ineligible as a substitute for that school.

Of the approximately 2,090 originally sampled public schools for the WCBA assessment, about 30 schools had a substitute activated, because the original school, although eligible, did not participate. Ultimately, about 20 of the activated substitute public schools, all in twelfth-grade, participated in the computer-based assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_pub_wcba_subs.aspx

Target Population of the Public School 2011 Writing Computer-Based Assessment (WCBA)

The target population for the 2011 public school writing computer-based assessment (WCBA) included all students who were enrolled in eighth and twelfth grades, in public schools, Bureau of Indian Education (BIE) schools, and Department of Defense Education Activity Schools (DoDEA) in the 50 states and the District of Columbia.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_pub_science_wcba_targpop.aspx

School and Student Participation Results for the 2011 Writing Computer-Based Assessment (WCBA)

Writing computer-based assessment (WCBA) participation in NAEP is not mandatory. Although a portion of the participating school sample consisted of substitute schools, it is preferable to calculate school response rates on the basis of school participation before substitution.

School Response Rates for the 2011
Writing Computer-Based
Assessment

In every NAEP survey, some of the sampled students are not assessed for the following reasons:

Weighted Student Response and
Exclusion Rates for the 2011 Writing
Computer-Based Assessment

- withdrawn students,
- excluded students with disabilities (SD),
- excluded English language learner (ELL) students, or
- students absent from both the original session and the makeup session (not excluded but not assessed).

Withdrawn students are those who have left the school before the original assessment. Excluded students were determined by their school to be unable to meaningfully take the NAEP assessment in their assigned subject, even with an accommodation. Excluded students must also be classified as SD and/or ELL. Other students who were absent for the initial session are assessed in the makeup session. The last category includes students who were not excluded (i.e., "were to be assessed") but were not assessed either due to absence from both sessions or because of a refusal to participate. Assessed students are also classified as assessed without an accommodation or assessed with an accommodation. The latter group can be divided into SD students assessed with an accommodation, ELL students assessed with an accommodation, or students who are both SD and ELL and accommodated. Note that some SD and ELL students are assessed without accommodations, and students who are neither SD nor ELL can only be assessed without an accommodation.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_science_results_wcba.aspx

School Response Rates for 2011 Writing Computer-Based Assessment (WCBA)

The following table presents counts of eligible sampled schools and participating schools, as well as weighted school response rates, for the writing computer-based assessment (WCBA) school sample. The weighted school response rates estimate the proportion of the student population that is represented by the participating school sample prior to substitution.

School response counts and rates for public and private schools, writing computer-based assessment (WCBA), by school type, geographic region, and grade: 2011

Grade	School type and geographic region	Number of sample eligible schools	Number of participating schools, including substitutes	Weighted school response rate prior to substitution (percent)
8	National all	981	947	97.27
	Northeast all	167	156	95.36
	Midwest all	189	186	98.83
	South all	377	365	97.15
	West all	248	240	97.43
	National public	841	839	99.73
	National private	140	108	71.21
	Catholic	42	42	95.53
	Non-Catholic private	98	66	52.06
	National all	1,300	1,200	93.52
	12	Northeast all	233	213
Midwest all		249	245	96.93
South all		468	441	94.66
West all		341	318	89.70
National public		1,100	1,100	96.04
National private		160	122	67.23
Catholic		55	50	76.60
Non-Catholic private		105	72	58.35

NOTE: Detail may not sum to totals because of rounding. Percentages are based on unrounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_schresp_rates_wcba.aspx

Weighted Student Response and Exclusion Rates for the 2011 Writing Computer-Based Assessment (WCBA)

The following table presents the weighted student response and exclusion rates for the writing computer-based assessment (WCBA). The exclusion rates give the percentage of students excluded among all eligible students. Excluded students must be either students with disabilities (SD) or English language learners (ELL). The response rates indicate the percentage of students assessed among those who were intended to take the assessment in participating schools. Thus, students who were excluded are not included in the denominators of the response rates.

Weighted student response and exclusion rates for public and private schools, writing computer-based assessment (WCBA), by school type and geographic region and grade: 2011

Grade	School type and geographic region	Weighted student response rates (percent)	Weighted percent of all students who are SD and excluded	Weighted percent of all students who are ELL and excluded	
8	National all	94.00	1.40	0.49	
	Northeast all	93.18	1.52	0.80	
	Midwest all	94.18	1.59	0.33	
	South all	94.63	1.32	0.44	
	West all	93.47	1.25	0.46	
	National public	93.99	1.51	0.52	
	National private	94.09	0.24	0.05	
	Catholic	94.72	0.53	0.00	
	Non-Catholic private	93.29	0.00	0.08	
	12	National all	86.98	2.11	0.35
		Northeast all	84.18	1.75	0.43
Midwest all		86.39	2.11	0.15	
South all		88.38	2.48	0.32	
West all		87.71	1.84	0.51	
National public		86.98	2.29	0.38	
National private		87.01	0.24	0.03	
Catholic		85.95	0.11	0.00	
Non-Catholic private		88.36	0.36	0.06	

NOTE: SD = students with disabilities; ELL = English language learners.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Writing Computer-Based Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_science_wcba_stud_resp.aspx

Selection of Primary Sampling Units for the 2011 WCBA and MCBS Assessments

For the writing computer-based assessment (WCBA) and mathematics computer-based study (MCBS), a sample of 105 primary sampling units (PSUs) was drawn from a frame of PSUs based on current Census information.

After the PSU frame was created, 29 certainty PSUs (those with measures of size large enough that it is efficient to take them with probability of selection equal to 1) were identified and set aside.

Stratification of the noncertainty PSUs (the remaining PSUs with probabilities of selection strictly less than 1) was carried out after analysis of Census 2000 data and NAEP 2000 achievement scores identified the stratification variables. This analysis identified the set of PSU-level, Census-based variables that had as much association with NAEP assessment scores as possible. The intent was that the results of this analysis and stratification would be used for multiple design years and subject matter. The results were used previously in 2006, 2008, 2009, and 2010. Periodically, this analysis and stratification will be conducted according to the availability of Census data and key assessment scores. Measures of size and probabilities of selection were defined for each PSU, and a stratified systematic sample of PSUs was drawn. For WCBA and MCBS, 76 noncertainty PSUs were selected.

The PSUs on the frame satisfied the following criteria:

- The PSU sampling frame included all U.S. states and the District of Columbia, but excluded the U.S. territories and Puerto Rico;
- PSUs consisted of one county or contiguous multiple counties;
- Metropolitan Statistical Areas (MSAs) were designated as separate PSUs even with their large size, as they were sufficiently compact in terms of their travel costs (due to higher levels of transportation infrastructure);
- PSUs did not cross Census region boundaries;
- PSUs did not cross state boundaries, in general;
- Non-MSA PSUs in the Northeast and South Census regions had a minimum population of 15,000 youths (age 0 to 17 inclusive), and in the Midwest and West Census regions had a minimum population of 10,000 youths, in general, according to the 2003 U.S. Census Bureau's Population Estimates Program; and
- Non-MSA PSUs were to be of minimum size (defined in terms of maximum distance between points—a rough proxy for travel time) while still satisfying the minimum population constraints.

PSU Generation: Metropolitan
Statistical Areas

PSU Generation: Certainty PSUs

PSU Generation:
Non-Metropolitan Statistical
Areas

PSU Frame: Stratification

Final PSU Sample

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_wcba_mcbps_psu_selection.aspx

Final Primary Sampling Unit (PSU) Sample for the 2011 Assessments

For the writing computer-based assessment (WCBA) and mathematics computer-based study (MCBS), a primary sampling unit (PSU) sample was drawn independently from each of the 76 noncertainty strata defined in Final Primary Sampling Unit Strata. One PSU was selected with probability proportionate to size (with size equal to estimated number of youths) within each stratum. The selection of the noncertainty PSUs was designed to minimize the overlap with the 2008 LTT sample, the 2009 science sample, and the 2010 sample.

Also, 29 PSUs were included in the sample of PSUs with certainty.

Distribution of sampled PSUs, computer-based writing and mathematics assessments, by PSU type: 2011

PSU type	Number of sampled PSUs
Total	105
Census region	
Northeast	15
Midwest	23
South	42
West	25
Certainty/metropolitan status	
Certainty metropolitan	29
Noncertainty metropolitan	54
Noncertainty non-metropolitan	22

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_nat_psu_finalsample.aspx

Primary Sampling Unit (PSU) Generation: Certainty PSUs for the 2011 Assessments

Any primary sampling unit (PSU) was defined as a certainty PSU if it had 500,000 or more youths. The estimated number of youths is the number of persons age 17 or under from the 2008 U.S. Census Bureau's Population Estimates Program.¹ These PSUs were so large that a sample of schools was taken from all of them (rather than from only a subsample of them, as with noncertainty PSUs). There were two exceptions to the 500,000 cutoff. The Honolulu, Hawaii, and Washington, D.C., PSUs were included as certainties by design: Honolulu, Hawaii in order to reduce the variability of including Native Hawaiian students, and Washington, D.C., as it is essentially a part of the larger MD-VA-DC Washington area PSU. A total of 29 PSUs were classified as certainties in the 2011 frame. The table below provides a listing of the certainty PSUs by census region.

Metropolitan statistical area (MSA) definition for certainty PSUs, by primary sampling unit (PSU): 2011

Primary sampling unit (PSU)	Metropolitan statistical area (MSA)	State	Number of counties	Number of youths
Grand total			203	30,407,927
Total Northeast			40	6,753,238
1--1	Boston-Cambridge-Quincy	MA	5	903,391
1--2	New York-Northern New Jersey-Long Island	NJ-PA	13	1,518,504
1--3	New York-Northern New Jersey-Long Island	NY	10	2,915,787
1--4	Pittsburgh	PA	7	481,884
1--5	Philadelphia-Camden-Wilmington	PA	5	933,672
Total Midwest			40	5,113,204
2--1	Chicago-Naperville-Joliet	IL	9	2,231,409
2--2	Detroit-Warren-Livonia	MI	6	1,089,901
2--3	Minneapolis-St. Paul-Bloomington	MN	11	782,054
2--4	St. Louis	MO	9	519,876
2--5	Cleveland-Elyria-Mentor	OH	5	489,964
Total South			93	9,089,075
3--1	Washington-Arlington-Alexandria	DC	1	112,016
3--2	Tampa-St. Petersburg-Clearwater	FL	4	592,372
3--3	Miami-Fort Lauderdale-Miami Beach	FL	3	1,204,361
3--4	Atlanta-Sandy Springs-Marietta	GA	28	1,443,448
3--5	Washington-Arlington-Alexandria	MD	5	546,557
3--6	Baltimore-Towson	MD	7	629,656
3--7	San Antonio	TX	8	561,126
3--8	Houston-Sugar Land-Baytown	TX	10	1,615,543
3--9	Dallas-Fort Worth-Arlington	TX	12	1,755,255
3--10	Washington-Arlington-Alexandria	VA	15	628,741
Total West			30	9,452,410
4--1	Phoenix-Mesa-Scottsdale	AZ	2	1,168,524
4--2	Sacramento--Arden-Arcade--Roseville	CA	4	519,855
4--3	San Diego-Carlsbad-San Marcos	CA	1	744,470
4--4	San Francisco-Oakland-Fremont	CA	5	923,680
4--5	Riverside-San Bernardino-Ontario	CA	2	1,174,107
4--6	Los Angeles-Long Beach-Santa Ana	CA	2	3,314,817
4--7	Denver-Aurora	CO	10	637,268
4--8	Honolulu	HI	1	199,268
4--9	Seattle-Tacoma-Bellevue	WA	3	770,421

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

¹ The U.S. Census Bureau's Population Estimates Program (<http://www.census.gov/popest/>) yearly publishes total resident population estimates by demographics such as age, sex, race, and Hispanic origin for the nation, states, and counties.

Primary Sampling Unit Frame: Stratification for the 2011 Assessments

The noncertainty primary sampling unit (PSU) strata were initially determined by census region and metropolitan status (metropolitan or non-metropolitan)—a total of eight primary strata. Measures of size were defined for each of these strata, determined by the relative share of the eventual PSU sample (the sample size is designed to be proportional to the number of youths). The PSU stratum measure of size then is the total number of youths in the stratum. The table below presents these counts for each of the eight primary strata. The relative share of the PSU sample size for each stratum is the number of youths in the stratum divided by the total number of youths, multiplied by 76 (the total noncertainty PSU strata for the writing computer-based assessment [WCBA] and mathematics computer-based study [MCBS]). The results of these calculations are given in the table below.

Stepwise Regression Analysis
Results for PSU Stratification
Final PSU Strata

Noncertainty primary sampling unit (PSU) frame size statistics, by primary stratum: 2011

Primary stratum	PSUs	Counties	Youths	Target number of PSU strata	Set number of PSU strata	Youths per PSU stratum
Total noncertainty PSUs	1,040	2,937	43,533,921	76.0	76	572,815
Northeast Region Metropolitan	46	83	4,531,012	7.9	8	566,377
Northeast Region Non-Metropolitan	50	94	1,098,293	1.9	2	549,147
Midwest Region Metropolitan	100	246	7,458,159	13.0	12	621,513
Midwest Region Non-Metropolitan	249	769	3,505,128	6.1	6	584,188
South Region Metropolitan	153	458	13,269,054	23.2	22	603,139
South Region Non-Metropolitan	269	872	5,190,589	9.1	10	519,059
West Region Metropolitan	71	101	6,803,588	11.9	12	566,966
West Region Non-Metropolitan	102	314	1,678,098	2.9	4	419,525

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

The division of the primary strata into the final strata was done on a stratum-by-stratum basis. The criteria for good PSU strata were: (1) the strata should have as equal measures of size as possible, which reduces sampling variance, and (2) the strata should be as heterogeneous in measured achievement as possible (i.e., there should be strata with low mean achievement, strata with mid-level mean achievement, and strata with high mean achievement). This second criterion also ultimately reduces the variance of the assessment estimates since the final PSU sample will be balanced in terms of assessment means.

PSU assessment means from the current year cannot be used, as assessments are only conducted after sampling is completed. Information is available about PSU sociodemographic characteristics in advance, however. An analysis was done within each primary stratum to find sociodemographic variables that were good predictors of the NAEP 2000 mathematics and science assessment results. Using these sociodemographic variables to define strata should increase the chance of having efficient strata definitions. The page Stepwise Regression Analysis Results for PSU Stratification describes this analysis for each primary stratum.

The final step in stratification was to define the desired number of strata using the selected stratifiers while constructing strata that were as close to equal size as possible (with size defined by number of youths). The objective was to establish strata that had a high between-stratum variance for the stratifiers (i.e., which "spread out" the stratifiers as much as possible). These strata are given on the page Final PSU Strata.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_psu_framestrat.aspx

Final Primary Sampling Unit (PSU) Strata for the 2011 Assessments

The strata were defined using the selected stratifiers from the stepwise regression analysis (see Stepwise Regression Analysis Results for PSU Stratification). The cutoffs were selected so that roughly equal measures of size were represented by each stratum.

Stratification for Northeast metropolitan noncertainty primary sampling units (PSUs), national assessment, by stratum: 2011

Stratum	Primary stratifier	Secondary stratifier	PSUs	Measure of size
Total	†	†	46	4,531,012
1	Percent child poverty <=10.1%	Percent Black <=15.9%	8	572,628
2	Percent child poverty <=10.1%	15.9%< Percent Black <=27.7%	2	533,970
3	10.1%< Percent child poverty <=12.5%	Percent Black <=14.9%	7	578,198
4	10.1%< Percent child poverty <=12.5%	14.9%< Percent Black <=38.2%	4	624,044
5	12.5%< Percent child poverty <=13.4%	†	5	543,994
6	13.4%< Percent child poverty <=15.1%	†	7	574,735
7	15.1%< Percent child poverty <=17%	†	5	516,879
8	17%< Percent child poverty <=20.7%	†	8	586,564
Mean	†	†	†	566,377

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratification for Northeast non-metropolitan noncertainty primary sampling units (PSUs), national assessment, by stratum: 2011

Stratum	Primary stratifier	PSUs	Measure of size
Total	†	50	1,098,293
1	Percent child poverty <=15.7%	22	544,762
2	15.7%< Percent child poverty <=22.8%	28	553,531
Mean	†	†	549,147

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratification for Midwest metropolitan noncertainty primary sampling units (PSUs), national assessments, by stratum: 2011

Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of size
Total	†	†	†	100	7,458,159
1	Percent child poverty <=12.5%	†	Pct Asian <=1.1%	17	623,684
2	Percent child poverty <=12.5%	†	1.1%< Pct Asian <=1.4%	4	668,332
3	Percent child poverty <=12.5%	†	1.4%< Pct Asian <=2.4%	8	598,589
4	Percent child poverty <=12.5%	†	2.4%< Pct Asian <=2.6%	3	706,518
5	Percent child poverty <=12.5%	†	2.6%< Pct Asian <=3.4%	7	602,499
6	Percent child poverty <=12.5%	†	3.4%< Pct Asian <=10.3%	13	618,908
7	12.5%< Percent child poverty <=12.9%	†	†	6	619,810
8	12.9%< Percent child poverty <=14.5%	†	Pct Asian <=1.3%	7	623,599
9	12.9%< Percent child poverty <=14.5%	†	1.3%< Pct Asian <=2.7%	7	602,409
10	14.5%< Percent child poverty <=27.6%	Med HH Income <=\$38,291	†	17	603,071
11	14.5%< Percent child poverty <=27.6%	\$38,291< Med HH Income <=\$46,460	Pct Asian <=0.9%	7	569,605
12	14.5%< Percent child poverty <=27.6%	\$38,291< Med HH Income <=\$46,460	0.9%< Pct Asian <=3.1%	4	621,135
Mean	†	†	†	†	621,513

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratification for Midwest non-metropolitan noncertainty primary sampling units (PSUs), national assessment, by stratum: 2011

Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of size
Total	†	†	†	249	3,505,128
1	Percent child poverty <=15.7%	Percent college grd <=12.5%	†	41	577,244
2	Percent child poverty <=15.7%	12.5%< Percent college grd <=36.0%	Pct BHI <=4.2%	42	577,144
3	Percent child poverty <=15.7%	12.5%< Percent college grd <=36.0%	4.2%< Pct BHI <=8.5%	42	582,552
4	Percent child poverty <=15.7%	12.5%< Percent college grd <=36.0%	8.5%< Pct BHI <=41.4%	38	591,909
5	15.7%< Percent child poverty <=45.5%	Percent college grd <=13.2%	†	41	584,830
6	15.7%< Percent child poverty <=45.5%	13.2%< Percent college grd <=23.0%	†	45	591,449
Mean	†	†	†	†	584,188

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratification for South metropolitan noncertainty primary sampling units (PSUs), national assessment, by stratum: 2011

Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of size
Total	†	†	†	153	13,269,054
1	Percent child poverty <=22.7%	Percent Black <=39.7%	Percent Hispanic <=1.7%	17	596,069
2	Percent child poverty <=22.7%	Percent Black <=39.7%	1.7%< Percent Hispanic <=2.6%	11	630,434
3	Percent child poverty <=22.7%	Percent Black <=39.7%	2.6%< Percent Hispanic <=2.7%	3	578,311
4	Percent child poverty <=22.7%	Percent Black <=39.7%	2.7%< Percent Hispanic <=3.0%	8	571,617
5	Percent child poverty <=22.7%	Percent Black <=39.7%	3.0%< Percent Hispanic <=3.5%	8	663,600
6	Percent child poverty <=22.7%	Percent Black <=39.7%	3.5%< Percent Hispanic <=4.2%	5	655,658
7	Percent child poverty <=22.7%	Percent Black <=39.7%	4.2%< Percent Hispanic <=4.8%	4	626,966
8	Percent child poverty <=22.7%	Percent Black <=39.7%	4.8%< Percent Hispanic <=5.5%	6	518,112
9	Percent child poverty <=22.7%	Percent Black <=39.7%	5.5%< Percent Hispanic <=7.3%	8	589,272
10	Percent child poverty <=22.7%	Percent Black <=39.7%	7.3%< Percent Hispanic <=8.5%	5	531,498
11	Percent child poverty <=22.7%	Percent Black <=39.7%	8.5%< Percent Hispanic <=9.1%	3	701,272
12	Percent child poverty <=22.7%	Percent Black <=39.7%	9.1%< Percent Hispanic <=11.2%	7	700,785
13	Percent child poverty <=22.7%	Percent Black <=39.7%	11.2%< Percent Hispanic <=14.6%	8	571,531
14	Percent child poverty <=22.7%	Percent Black <=39.7%	14.6%< Percent Hispanic <=21.1%	6	548,529
15	Percent child poverty <=22.7%	Percent Black <=39.7%	21.1%< Percent Hispanic <=30.8%	5	691,494

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of size
Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of size
16	Percent child poverty <=22.7%	Percent Black <=39.7%	30.8%< Percent Hispanic <=51.2%	7	710,494
17	Percent child poverty <=22.7%	39.7%< Percent Black <=56.6%	Percent Hispanic <=2.7%	6	547,257
18	Percent child poverty <=22.7%	39.7%< Percent Black <=56.6%	2.7%< Percent Hispanic <=7.8%	6	655,387
19	22.7%< Percent child poverty <=24.3%	†	†	11	498,732
20	24.3%< Percent child poverty <=45.7%	Percent Black <=2.0%	†	4	712,425
21	24.3%< Percent child poverty <=45.7%	2.0%< Percent Black <=60.8%	Percent Hispanic <=3.8%	10	494,418
22	24.3%< Percent child poverty <=45.7%	2.0%< Percent Black <=60.8%	3.8%< Percent Hispanic <=64.1%	5	475,193
Mean	†	†	†	†	603,139

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratification for South non-metropolitan noncertainty primary sampling units (PSUs), national assessment, by stratum: 2011

Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of size
Total	†	†	†	269	5,190,589
1	Percent Black <=31.2%	Median HH income <=\$36,049	Percent Asian <=0.2%	32	511,172
2	Percent Black <=31.2%	Median HH income <=\$36,049	0.2%< Percent Asian <0.3%	29	515,004
3	Percent Black <=31.2%	Median HH income <=\$36,049	Percent Asian =0.3%	32	520,184
4	Percent Black <=31.2%	Median HH income <=\$36,049	0.3%< Percent Asian <=0.4%	29	510,176
5	Percent Black <=31.2%	Median HH income <=\$36,049	0.4%< Percent Asian <=0.7%	26	538,940
6	Percent Black <=31.2%	Median HH income <=\$36,049	0.7%< Percent Asian <=3.0%	26	542,335
7	Percent Black <=31.2%	\$36,049< Median HH income <=\$54,721	†	21	551,780
8	31.2%< Percent Black <=51.3%	Median HH income <=\$29,555	†	25	513,918
9	31.2%< Percent Black <=51.3%	\$29,555< Median HH income <=\$44,421	†	22	499,209
10	51.3%< Percent Black <=79.4%	†	†	27	487,871
Mean	†	†	†	†	519,059

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratification for West metropolitan noncertainty primary sampling units (PSUs), national assessment, by stratum: 2011

Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of size
Total	†	†	†	71	6,803,588
1	Percent high school graduates <=70%	Percent college graduates <=13.5%	†	7	635,482
2	Percent high school graduates <=70%	13.5%< Percent college graduates <=22.5%	†	4	508,977
3	70%< Percent high school graduates <=78.9%	†	†	6	513,746
4	78.9%< Percent high school graduates <=79.6%	†	†	3	656,056
5	79.6%< Percent high school graduates <=88.3%	Percent college graduates <=21.8%	†	14	559,117
6	79.6%< Percent high school graduates <=88.3%	21.8%< Percent college graduates <=25.5%	†	8	587,808
7	79.6%< Percent high school graduates <=88.3%	25.5%< Percent college graduates <=26.9%	†	4	525,628
8	79.6%< Percent high school graduates <=88.3%	26.9%< Percent college graduates <=27.8%	†	3	583,877
9	79.6%< Percent high school graduates <=88.3%	27.8%< Percent college graduates <=30.3%	†	3	557,408
10	79.6%< Percent high school graduates <=88.3%	30.3%< Percent college graduates <=39.5%	†	3	527,923
11	88.3%< Percent high school graduates <=90.1%	†	†	8	556,141
12	90.1%< Percent high school graduates <=93.1%	†	†	8	591,425
Mean	†	†	†	†	555,966

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stratification for West non-metropolitan noncertainty primary sampling units (PSUs), national assessment, by stratum: 2011

Stratum	Primary stratifier	Secondary stratifier	Tertiary stratifier	PSUs	Measure of Size
Total	†	†	†	102	1,678,098
1	Percent college graduates <=21%	Percent child poverty <=17.3%	†	26	425,522
2	Percent college graduates <=21%	17.3%< Percent child poverty <=43%	Percent HS graduates <=77.3%	23	409,676
3	Percent college graduates <=21%	17.3%< Percent child poverty <=43%	77.3%< Percent HS graduates <=86.4%	28	409,881
4	21%< Percent college graduates <=42.6%	†	†	25	433,019
Mean	†	†	†	†	419,525

† Not applicable.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Stepwise Regression Analysis Results for Primary Sampling Unit (PSU) Stratification for the 2011 Assessments

The objective was to find the optimum set of primary sampling unit (PSU)-level sociodemographic characteristics in terms of strength of relationship to achievement. The PSU-level values of these characteristics were derived from the 2000 Census Summary Files and the 2003 county population estimates, computed by combining the county-level data (using county youth estimates as the relative weighting factor for each county within the PSU). The characteristics used, and their abbreviations as used in the tables, were as follows:

- race/ethnicity percentages in schools (percent Black, Hispanic, or American Indian/Alaska Native – "Pct BHI" below; percent Black; percent Hispanic – "Hsp" below; percent Asian; percent American Indian/Alaska Native; percent two or more races);
- income levels (median household income – "Med Inc" below; percent children below the poverty line – "Cld pov" below);
- education levels in population (i.e., percent of persons age 25 and over who completed high school but have no college degree – "HS grd," percent of persons age 25 and over with college degrees – "CG grd" below);
- percent of renters (i.e., percent of householders who rent rather than own their place of residence); and
- percent of female householders living alone.

These PSU-level census characteristics were examined within each of the four NAEP 2000 assessment values: fourth-grade mathematics achievement, fourth-grade science achievement, eighth-grade mathematics achievement, and eighth-grade science achievement. These PSU-level values for achievement were computed using the 2000 state NAEP database. The criterion was that good strata should be heterogeneous for each of the four characteristics (i.e., within-stratum variance for each assessment value should be low and between-stratum variance high), so that strata are defined that do a good job for both mathematics and science, in both grades, not just the best possible job for one subject and one grade. This prevents overfitting to some extent.

The analysis was done separately within each of the eight primary strata (census region by metro status), using a forward stepwise regression approach, with a p-value cutoff of 20 percent. The results are given in the tables below. The order of the regressors is the order of entry into the stepwise procedure. The p-value is for an F-test for entry of the regressor into the forward stepwise model. The minus or plus sign indicates the direction of effect (negative indicates that increase in the regressor is related to lower achievement; positive indicates that increase in the regressor is related to higher achievement). The regressor is in italics if the direction of the effect is unexpected (i.e., negative when we generally expect a positive effect, or vice versa). The stratifiers chosen to generate the final PSU strata are indicated in a note below the regression analysis result tables.

Northeast metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	Cld pov - (p=0.084)	Cld pov - (p=0.174)	Black - (p=0.068)	HS grd + (p=0.026)
Second variable	Black - (p=0.159)	†	†	Black - (p=0.193)

† Not applicable.

NOTE: Stratifiers chosen were percent child poverty (Cld pov) and percent Black. HS grd = high school graduate with no college degree. Black includes African American.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Northeast non-metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	<i>Renters</i> + (p=0.092)	CG grd + (p=0.010)	Cld pov - (p=0.085)	Black + (p=0.005)
Second variable	†	Black + (p=0.176)	Med Inc - (p=0.002)	HS grd + (p=0.030)
Third variable	†	†	Renters - (p=0.085)	†

† Not applicable.

NOTE: Stratifier chosen was percent child poverty (Cld pov). Renters = householders who rent rather than own their place of residence; CG grd = college graduate; Med Inc = median household income; HS grd = high school graduate with no college degree. Black includes African American.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Midwest metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	Cld pov - (p=0.003)	Asian + (p=0.004)	Cld pov - (p<0.001)	†
Second variable	Med Inc - (p=0.200)	Med Inc - (p=0.055)	Med Inc - (p=0.001)	†
Third variable	Pct BHI + (p=0.100)	†	Black + (p=0.006)	†
Fourth variable	†	†	HS grd - (p=0.050)	†

† Not applicable.

NOTE: Stratifiers chosen were percent child poverty (Cld pov), median household income (Med Inc), and percent Asian. Pct BHI = percent Black, Hispanic, or American Indian/Alaska Native; HS grd = high school graduate with no college degree. Black includes African American and Hispanic includes Latino; Asian includes Native Hawaiian and other Pacific Islander.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

Midwest non-metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	Cld pov - (p=0.012)	Cld pov - (p=0.002)	†	CG grd + (p=0.005)
Second variable	Pct BHI + (p=0.128)	Asian + (p=0.124)	†	Pct BHI - (p=0.079)

† Not applicable.

NOTE: Stratifiers chosen were percent child poverty (Cld pov), percent college graduates (CG grd), and percent Black, Hispanic, or American Indian/Alaska Native (BHI). Black includes African American and Hispanic includes Latino; Asian includes Native Hawaiian and other Pacific Islander.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

South metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	Hsp + (p=0.001)	Asian + (p=0.014)	Black - Hsp - (p=0.005)	Cld pov - (p=0.011)
Second variable	Cld pov - (p=0.001)	Black - (p=0.038)	†	Black - (p=0.127)

† Not applicable.

NOTE: Stratifiers chosen were percent child poverty (Cld pov), percent Black, and percent Hispanic (Hsp). Black includes African American and Hispanic includes Latino; Asian includes Native Hawaiian and other Pacific Islander.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

South non-metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	Black - (p<0.001)	Black - (p=0.005)	Black - (p=0.014)	Black - (p=0.001)
Second variable	Asian + (p=0.037)	Med Inc + (p=0.037)	Asian + (p=0.036)	Med Inc + (p=0.045)
Third variable	†	Black-Hsp + (p=0.176)	Cld Pov - (p=0.068)	†
Fourth Variable	†	†	CG grd - (p=0.127)	†

† Not applicable.

NOTE: Stratifiers chosen were percent Black, median household income (Med Inc), and percent Asian. Cld Pov = children below the poverty line; CG grd = college graduate. Black includes African American and Hispanic includes Latino; Asian includes Native Hawaiian and other Pacific Islander.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

West metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	CG grd + (p=0.094)	Pct BHI - (p=0.049)	HS grd + (p<0.001)	HS grd - (p=0.160)
Second variable	HS grd + (p=0.191)	†	Asian + (p=0.007)	Med Inc - (p=0.001)
Third variable	†	†	Black - (p=0.080)	CG grd + (p=0.003)
Fourth variable	†	†	†	Asian + (p=0.009)
Fifth variable	†	†	†	Cld pov - (p=0.037)
Sixth variable	†	†	†	Renters - (p=0.087)

† Not applicable.

NOTE: Stratifiers chosen were percent high school graduates (HS grd) and percent college graduates (CG grd). Pct BHI = percent Black, Hispanic, or American Indian/Alaska Native; Med Inc = median household income; Cld pov = children below the poverty line; Renters = householders who rent rather than own their place of residence. Black includes African American and Hispanic includes Latino; Asian includes Native Hawaiian and other Pacific Islander.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

West non-metropolitan stepwise regression analysis on NAEP 2000 achievement scores, national assessment, by subject, grade, and variable: 2011

Variable	Mathematics 4	Mathematics 8	Science 4	Science 8
First variable	Renter - (p=0.013)	CG grd + (p=0.006)	HS grd + (p<0.001)	CG grd + (p=0.220)
Second variable	Black + (p=0.040)	†	Cld pov + (p=0.008)	Med Inc - (p=0.038)
Third variable	Cld pov - (p=0.005)	†	Asian - (p=0.017)	Cld pov - (p=0.135)
Fourth variable	HS grd - (p=0.092)	†	†	†

† Not applicable.

NOTE: Stratifiers chosen were percent college graduates (CG grd), percent child poverty (Cld pov), and percent high school graduates (HS grd). Renter = householders who rent rather than own their place of residence; Med Inc = median household income. Black includes African American; Asian includes Native Hawaiian and other Pacific Islander.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_samp_natl_psu_stepregr.aspx

Primary Sampling Unit Generation: Metropolitan Statistical Areas for the 2011 Assessments

The 2004 definitions of Core Based Statistical Areas (CBSAs) and Combined Statistical Areas (CSAs), which are also referred to as Metropolitan Statistical Areas (MSAs), were used to define primary sampling units (PSUs). These definitions were the most recently available definitions from the U.S. Office of Management and Budget (OMB) at the time of 2011 PSU frame creation. The new CBSA areas consisted of clusters of one or more counties classified as metropolitan and micropolitan statistical areas.

The metropolitan PSUs were manually created by grouping counties in MSAs. Each MSA constituted a PSU, except for those areas that crossed state boundaries. These areas were split into "proto-PSUs" along state boundaries. Proto-PSUs consisted of portions of MSAs within individual states¹. For example, the New York-Newark-Bridgeport, NY-NJ-CT-PA MSA was partitioned into four proto-PSUs, by state.

If the proto-PSU did not violate the size constraints, it was defined as a PSU. In some cases, these proto-PSUs violated the minimum size constraint of 15,000 youths for the Northeast or South census regions, and 10,000 youths for the Midwest and West census regions. There were 14 of these proto-PSUs violating size constraints. In one of these 14 cases where the size was close to the constraint, the proto-PSU was defined as a PSU. In the remaining 13 smaller cases, these proto-PSUs were combined with the adjacent MSA proto-PSUs to form the final PSUs. In these cases, the combined PSUs crossed state boundaries.

A total of 29 of the PSUs were defined as certainty PSUs. The remaining 370 PSUs comprised the MSA frame for PSU sampling, covering a total of 888 counties. The table below presents estimates for the number of youths by census region. These estimates come from the county-level estimates of numbers of persons aged 0 to 17 from the 2008 U.S. Census Bureau's Population Estimates Program². The 2008 estimates were the most recent demographic data at the time of the PSU selection.

Noncertainty Metropolitan Statistical Area (MSA) primary sampling unit (PSU) frame, by census region: 2011

Census region	PSUs	Counties	Youths	Average number of youths per PSU
Total	370	888	32,061,813	86,654
Northeast	46	83	4,531,012	98,500
Midwest	100	246	7,458,159	74,582
South	153	458	13,269,054	86,726
West	71	101	6,803,588	95,825

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

¹ Note that starting in 2006, this is a change from earlier NAEP cycles. Field personnel had indicated that contacts with state officials were very important in the process of recruiting schools. Because of this, it was decided that making single-state rather than multi-state PSUs was a better approach. In a few cases, small size proto-PSUs were combined across state lines if it was necessary to satisfy other criteria.

² The U.S. Census Bureau's Population Estimates Program (<http://www.census.gov/popest/>) yearly publishes total resident population estimates by demographics such as age, sex, race, and Hispanic origin for the nation, states, and counties.

Primary Sampling Unit Generation: Non-Metropolitan Statistical Areas for the 2011 Assessments

A software algorithm was utilized to define a preliminary set of primary sampling units (PSUs) satisfying the design constraints. The input set consisted of all of the non-metropolitan counties. The software formed PSUs that satisfied the minimum size constraints while not crossing state boundaries. The software also minimized the maximum point-to-point distance for the candidate PSUs, while still satisfying the minimum size constraints (15,000 youths in the Northeast and South census regions, and 10,000 youths in the Midwest and West census regions). "Worst first" was the general approach: the county that had the PSU with the largest maximum point-to-point distance was fitted first, with those counties that best fit within a PSU containing the "worst-first" county put together to form the first PSU. The algorithm was then run on the remaining counties not yet assigned to a PSU finding the next "worst-first" county.

Initially, there were 22 counties that could not be combined into PSUs to satisfy the minimum size constraints while still remaining within a single state. Nine of the PSUs (formed from 12 of these counties) that were below the minimum size requirement were allowed to stand, since satisfying the minimum size requirement was not reasonably possible. The remaining counties were in Alaska, for which PSUs were manually drawn to better respect interstate highways (being drawn along the axis of these highways) and mountain ranges (avoiding crossing of ranges with poor road access). For Alaska, the proto-PSUs created by the program were replaced by the PSUs created for the NAEP 2004 assessment. The end result of this procedure was that all non-metropolitan PSUs were contained within state boundaries. There were a total of 670 final non-metropolitan PSUs.

The table below presents the number of PSUs, the number of counties represented, and the estimated number of youths (total and mean per PSU) by census region. The estimated number of youths (persons age 0 to 17) for each county comes from the 2008 U.S. Census Bureau's Population Estimates Program.

Non-metropolitan statistical area primary sampling unit (PSU) frame, by census region: 2011

Census region	PSUs	Counties	Youths	Average number of youths per PSU
Total	670	2,049	11,472,108	17,123
Northeast	50	94	1,098,293	21,966
Midwest	249	769	3,505,128	14,077
South	269	872	5,190,589	19,296
West	102	314	1,678,098	16,452

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 National Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_samp_natl_psu_nonmsa.aspx

Sample Design for the 2011 State Assessment

Each assessment cycle, a sample of students in designated grades within both public and private schools throughout the United States is selected for assessment. In state assessment years, of which 2011 is an example, the samples of public schools and their students in each state are large enough to support state-level estimates.

Target Population

The NAEP 2011 state assessments covered fourth- and eighth-grade students in public schools for operational mathematics and reading. It also covered operational science but only for students at grade 8, not grade 4. A representative sample of students was drawn in each participating jurisdiction, including the 50 states, the District of Columbia, Bureau of Indian Education (BIE) schools, Department of Defense Education Activity (DoDEA) schools, and in school districts chosen for the Trial Urban District Assessment (TUDA) study.

Sampling Frame

Stratification of Schools

All jurisdictions, including the TUDA districts, were included in the mathematics and reading assessments at grades 4 and 8. Only the 50 states, the District of Columbia, and DoDEA schools were included in the state science assessment. TUDA districts were not part of the state science assessment because they lacked the required number of students for this third subject.

School Sample Selection

The District of Columbia, which generally does not have enough students for an assessment in a third subject, participated in the grade 8 science assessment. This was accomplished by testing each student in two of the three assessment subjects.

Ineligible Schools

Student Sample Selection

Generally for the state assessments, each non-TUDA jurisdiction sample is designed to produce aggregate estimates with approximately equal precision for all the participating jurisdictions, as well as estimates for various subpopulations of interest. The target sample size for these jurisdictions is 3,150 for each operational subject. In 2011, the samples for operational mathematics and reading at grades 4 and 8 were designed in this fashion. However, the grade 8 sample for operational science used a sample size that was 20 percent smaller (2,520) in order to ensure that there were enough students available for the various pilot test and special studies that were also being conducted in eighth grade.

School and Student Participation

In 2011, the overall target student sample size for the operational samples in each non-TUDA jurisdiction was 6,300 at grade 4 and 8,820 at grade 8 (except for BIE schools). Since BIE schools did not have enough students for a state-level assessment for science, its target sample size at grade 8 was 6,600—3,150 each for mathematics and reading and 300 for science, enough samples so that it was adequately represented at the national level. Details can be found in the school sample selection.

The target population for the NAEP 2011 state assessment included students in public schools who were enrolled in the grades 4 and 8 at the time of assessment. The sampling frame included public schools having the relevant grade in each jurisdiction. The samples were selected based on a two-stage sample design:

1. selection of schools within participating strata, and
2. selection of students within schools.

From the stratified frame of public schools for each grade within each jurisdiction, a systematic random sample of grade-eligible schools was drawn with probability proportional to a measure of size based on the estimated grade-specific enrollment of the school.

For the TUDA study, schools were sampled from the 21 participating TUDA districts at the same time schools were selected for the jurisdiction samples. The TUDA districts are listed below. The ones in bold are those introduced in 2011.

- **Albuquerque Public Schools, New Mexico;**
- Atlanta Public Schools, Georgia;
- Austin Independent School District, Texas;
- Baltimore City Public Schools, Maryland;
- Boston Public Schools, Massachusetts;
- Charlotte-Mecklenburg Schools, North Carolina;
- Chicago Public Schools, Illinois;
- Cleveland Metropolitan School District, Ohio;
- **Dallas Independent School District, Texas;**
- Detroit Public Schools, Michigan;
- District of Columbia Public Schools, District of Columbia;
- Fresno Unified School District, California;
- **Hillsborough County Public Schools, Florida;**
- Houston Independent School District, Texas;
- Jefferson County Public Schools (Louisville), Kentucky;
- Los Angeles Unified School District, California;
- Miami-Dade County Public Schools, Florida;
- Milwaukee Public Schools, Wisconsin;
- New York City Department of Education, New York;
- School District of Philadelphia, Pennsylvania; and
- San Diego Unified School District, California.

These subsamples affected the design of the state samples in those states where TUDA districts were oversampled. In each of these states, there were distinct sampling rates for each TUDA district and for the balance of the state (i.e., the rest of the state not in a TUDA district).

Each selected school provided a list of eligible enrolled students from which a systematic sample of students was drawn. In fourth-grade schools, 63 students, if possible, were selected from each school: 30 for mathematics, 30 for reading, and 3 for the pilot tests. In eighth-grade schools, the within-school target sample size ranged from 63 to 114. The target sample sizes depended upon the size of the state. Very small states did not have any pilot test/special study sample and larger states had somewhat more pilot test/special study sample than other states. This is to ensure that the samples for the pilot tests and special studies would be reasonably nationally representative. Details can be found in the student sample selection.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdgn_state.aspx

Ineligible Schools for the 2011 State Assessment

The Common Core of Data (CCD) public school file from which most of the sampled schools were drawn corresponds to the 2007-2008 school year, some 3 years prior to the assessment school year. During the intervening period, some of these schools either closed, no longer offered the grade of interest, or were ineligible for other reasons. In such cases, the sampled school was coded as ineligible.

Eligible Schools Sampled by Jurisdiction

Ineligible Sampled Schools by
Ineligibility Type

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsgn_state_inelg.aspx

Eligible Schools Sampled for the 2011 State Assessment

The following table shows the number of eligible fourth- and eighth-grade schools sampled for each NAEP 2011 state assessment jurisdiction.

Eligible sampled schools, state assessment, by grade and jurisdiction: 2011

Jurisdiction	Grade 4		Grade 8	
	Total school sample	Eligible school sample	Total school sample	Eligible school sample
Total	8,500	8,000	7,700	7,000
Alabama	117	108	125	113
Alaska	202	180	167	136
Arizona	127	122	136	126
Arkansas	123	122	126	123
California–Fresno	55	54	28	21
California–Los Angeles	82	81	78	75
California–San Diego	58	56	39	33
California–Balance	97	90	112	100
Colorado	124	122	130	122
Connecticut	116	109	119	110
Delaware	109	95	68	49
Florida–Hillsborough County	56	54	50	46
Florida–Miami	88	83	86	76
Florida–Balance	91	87	97	89
Georgia–Atlanta	66	62	26	21
Georgia–Balance	108	107	106	102
Hawaii	118	116	81	78
Idaho	137	131	113	106
Illinois–Chicago	95	94	117	113
Illinois–Balance	98	97	106	101
Indiana	117	110	113	106
Iowa	141	137	138	134
Kansas	148	140	148	137
Kentucky–Jefferson County	56	54	46	35
Kentucky–Balance	102	99	110	104
Louisiana	134	121	163	119
Maine	166	157	143	132
Maryland–Baltimore	70	69	75	55
Maryland–Balance	103	102	101	97
Massachusetts–Boston	86	75	45	36
Massachusetts–Balance	116	106	109	106
Michigan–Detroit	58	47	63	48
Michigan–Balance	114	104	117	108
Minnesota	148	136	168	143
Mississippi	117	106	121	110
Missouri	131	130	136	123
Montana	206	194	200	189
Nebraska	181	158	169	143
Nevada	118	116	100	90
New Hampshire	133	131	96	94
New Jersey	118	113	116	111
New Mexico–Albuquerque	57	57	48	43
New Mexico–Balance	98	94	88	83
New York–New York City	82	81	91	89
New York–Balance	76	75	83	81
North Carolina–Charlotte	57	57	38	35
North Carolina–Balance	116	110	121	116
North Dakota	272	254	209	187
Ohio–Cleveland	86	73	71	57
Ohio–Balance	119	106	123	111
Oklahoma	140	135	149	149
Oregon	149	142	144	137
Pennsylvania–Philadelphia	109	104	61	54
Pennsylvania–Balance	58	57	107	104
Rhode Island	125	114	61	53
South Carolina	114	108	116	109
South Dakota	209	194	261	225
Tennessee	119	116	123	117
Texas–Austin	55	55	24	19
Texas–Dallas	55	54	41	36
Texas–Houston	86	82	50	45
Texas–Balance	110	105	121	113
Utah	130	124	125	115
Vermont	226	219	124	123
Virginia	115	112	108	105
Washington	141	133	140	135
West Virginia	152	145	117	110
Wisconsin–Milwaukee	69	65	60	47
Wisconsin–Balance	121	117	119	112
Wyoming	202	184	108	88
Other jurisdictions				
Bureau of Indian Education (BIE)	135	132	116	111
Department of Defense Education Activity (DoDEA)	120	112	72	64
District of Columbia (TUDA)	106	83	50	36
District of Columbia–Balance	47	41	52	40

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), 2011 State Assessment.

Ineligible Sampled Schools by Ineligibility Type for the 2011 State Assessment

The following table shows the unweighted counts and percentages of NAEP 2011 state assessment fourth- and eighth-grade schools that were eligible and ineligible, by reason for ineligibility.

School eligibility status, state assessment, by grade and eligibility status: 2011

Eligibility status	Grade 4		Grade 8	
	Unweighted count of schools	Unweighted percentage	Unweighted count of schools	Unweighted percentage
All sampled public schools	8,500	100.00	7,700	100.00
Eligible	8,000	94.45	7,000	90.93
No eligible students in grade	50	0.59	86	1.12
Does not have sampled grade	94	1.11	173	2.24
School closed	256	3.02	233	3.02
Not a regular school	57	0.67	182	2.36
Other ineligible school	13	0.15	22	0.29
Duplicate on sampling frame	1	0.01	3	0.04

NOTE: Detail may not sum to totals because of rounding. Percentages are based on unrounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Assessment.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_inelg_by_type.aspx

Sampling Frame for the 2011 State Assessment

Drawing the school samples for the 2011 assessments required a comprehensive list of public schools in each jurisdiction containing information for stratification purposes. As in previous NAEP assessments, the Common Core of Data (CCD) file developed by NCES was used to construct the sampling frame. The CCD file corresponding to the 2007-2008 school year provided the frame for all regular and state-operated public, Bureau of Indian Education (BIE), and Department of Defense Education Activity (DoDEA) schools.

The sampling frame excluded ungraded schools, vocational schools with no enrollment, special education only schools, prison and hospital schools, virtual or online schools, home-school entities, and juvenile correctional institutions.

For quality control purposes, school and student counts from the NAEP 2011 sampling frame were compared to school and student counts from the previous frames (2009 and 2010). No revisions to the frame were needed as a result of this check.

Fourth- and Eighth-Grade Schools
and Enrollment in Public School
Sampling Frame

New-School Sampling Frame

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_sampframe1.aspx

Fourth- and Eighth-Grade Schools and Enrollment in the 2011 Public School Sampling Frame

The following table displays, by jurisdiction, the number of fourth- and eighth-grade public schools and their estimated enrollment, as contained in the Common Core of Data (CCD) sampling frame. Grade-specific enrollment was estimated for each school as the average grade enrollment for grades 1 through 8.

Number of schools and enrollment in public school sampling frame, state assessment, by grade and jurisdiction: 2011

Jurisdiction	Grade 4		Grade 8	
	Schools	Enrollment	Schools	Enrollment
Total	51,519	3,673,587	27,641	3,641,680
Alabama	763	59,054	505	58,239
Alaska	363	9,579	284	9,835
Arizona	1,143	83,916	744	81,376
Arkansas	502	36,449	314	35,808
California–Fresno	70	5,873	26	5,611
California–Los Angeles	504	54,180	133	51,100
California–San Diego	137	10,065	50	9,574
California–Balance	5,133	384,702	2,499	392,507
Colorado	1011	60,527	492	57,972
Connecticut	598	42,270	278	42,542
Delaware	115	9,504	61	9,491
Florida–Hillsborough County	161	15,343	75	15,058
Florida–Miami	245	26,643	134	24,024
Florida–Balance	1,643	162,122	913	157,461
Georgia–Atlanta	63	4,312	25	3,452
Georgia–Balance	1,148	123,371	503	119,823
Hawaii	200	13,851	73	12,941
Idaho	364	20,976	197	20,471
Illinois–Chicago	466	30,920	454	30,801
Illinois–Balance	1,837	125,105	1,121	128,360
Indiana	1,117	80,673	474	80,568
Iowa	681	34,940	397	35,366
Kansas	742	35,050	415	34,433
Kentucky–Jefferson County	99	7,476	42	6,775
Kentucky–Balance	643	42,911	351	42,632
Louisiana	784	53,559	536	51,495
Maine	358	14,206	218	14,713
Maryland–Baltimore	123	6,361	63	5,078
Maryland–Balance	758	54,205	270	57,227
Massachusetts–Boston	78	3,897	35	4,177
Massachusetts–Balance	925	67,111	444	68,760
Michigan–Detroit	110	7,527	60	4,795
Michigan–Balance	1,782	111,920	974	117,678
Minnesota	954	60,507	678	61,854
Mississippi	446	38,785	294	37,858
Missouri	1,146	67,781	676	67,957
Montana	401	10,681	281	11,063
Nebraska	629	21,892	366	21,728
Nevada	364	34,060	153	33,909
New Hampshire	267	14,874	137	15,606
New Jersey	1,363	99,858	731	99,748
New Mexico–Albuquerque	96	7,504	46	7,167
New Mexico	328	17,604	157	17,549
New York–New York City	705	64,599	439	63,113
New York–Balance	1,652	128,764	862	135,640
North Carolina–Charlotte	96	10,868	36	9,800
North Carolina–Balance	1,289	105,420	663	100,588
North Dakota	268	6,972	190	7,357
Ohio–Cleveland	79	3,806	82	4,038
Ohio–Balance	1,874	130,150	1,023	132,090
Oklahoma	900	47,863	599	45,837
Oregon	763	42,709	386	42,267
Pennsylvania–Philadelphia	178	12,792	137	11,927
Pennsylvania–Balance	1,589	116,985	760	125,353
Rhode Island	184	10,656	57	11,416
South Carolina	607	54,359	291	52,743
South Dakota	329	9,196	254	9,270
Tennessee	980	74,086	562	69,972
Texas–Austin	80	6,579	21	5,181
Texas–Dallas	147	12,840	38	10,233
Texas–Houston	185	16,058	58	13,087
Texas–Balance	3,716	323,325	1,998	307,294
Utah	554	45,248	217	41,261
Vermont	225	6,566	123	6,674
Virginia	1,127	91,508	389	92,680
Washington	1,222	77,223	619	77,318
West Virginia	434	20,632	201	21,015
Wisconsin–Milwaukee	118	6,029	92	6,013
Wisconsin–Balance	996	54,566	536	56,481
Wyoming	191	6,568	95	6,375
Other jurisdictions				
Bureau of Indian Education (BIE)	131	2,971	111	2,736
Department of Defense Education Activity (DoDEA)	110	7,164	60	5,093
District of Columbia (TUDA)	102	3,861	34	2,601
District of Columbia–Balance	28	1,080	29	1,645

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 NAEP Assessments.

New-School Sampling Frame for the 2011 State Assessment

The Common Core of Data (CCD) file used for the frame corresponds to the 2007-2008 school year, whereas the assessment year is the 2010-2011 school year. During this 3-year period, some schools closed, some changed structure (one school becoming two schools, for example), and others came into existence.

As was done in previous years, to achieve as close to full coverage as possible, the school frame was supplemented by a sample of new schools obtained from a sample of districts. Each sampled district was sent a list of the CCD schools and asked to add in any new schools or old schools that had become newly eligible for grades 4, 8, and 12.

Since asking every school district to list new and newly-eligible schools would have generated too much of a burden, a sample of districts was contacted to obtain a list of new schools. To represent the unsampled districts in the full sample of schools, weights for schools included in the new-school sample were adjusted to reflect the district selection probability.

The goal was to allow every new school a chance of selection, thereby fully covering the target population of schools in operation during the 2010-2011 school year. The first step in this process was the development of a new-school frame through the construction of a district-level file from the CCD school-level file. To develop the frame, the district-level file was divided into two files: one for small districts and a second for medium and large districts.

Small districts contained no more than three schools on the frame in total, with no more than one school at each targeted grade (4, 8, and 12). New schools in small districts were identified during school recruitment and added to the sample if the old school was sampled. From a sampling perspective, the new school was viewed as an "annex" to the sampled school that had a well-defined probability of selection equal to that of the old school. The "frame" in this case was, in fact, the original frame, when the old school was sampled in a small district, the new school was automatically sampled as well.

The remaining districts were defined as medium and large districts. In these districts, a frame of new schools was developed based on information provided by the district. To limit the required effort, the new-school frame was created through developing information on a sample of medium and large public school districts in each jurisdiction.

Prior to district sampling, specific districts were in sample with certainty. They included the following districts:

- districts in jurisdictions where all schools were selected for sample at either grade 4 or 8,
- state-operated districts,
- districts in states with fewer than 10 districts,
- charter only districts (that is, districts containing no schools other than charter schools), and
- TUDA districts.

The remaining districts in each jurisdiction (excepting the certainty jurisdictions) were separated into two strata of large- and medium-size districts. These strata were defined by computing an aggregate percentage of enrollment for each district within the state (removing districts in the certainty strata defined above) and sorting in descending order by percentage of jurisdiction enrollment represented by the district. All districts up to and including the first district at or above the 80th cumulative percentage were defined as large districts. The remaining districts were defined as medium districts.

An example is given below. A state's districts are ordered by percentage enrollment. The first six become large districts and the last six become medium districts.

Large and medium districts example, state assessment, by enrollment, stratum, and district: 2007

District	Percentage enrollment	Cumulative percentage enrollment	Stratum
1	20	20	L
2	20	40	L
3	15	55	L
4	10	65	L
5	10	75	L
6	10	85	L
7	5	90	M
8	2	92	M
9	2	94	M
10	2	96	M
11	2	98	M
12	2	100	M

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 State Assessment.

The target sample size for each jurisdiction was 10 districts. Where possible, eight large and two medium districts were selected. However, in the example above, since there are only six large districts, all of the large districts and four of the medium districts were selected for the new-school inquiry.

If sampling was needed in the medium stratum, the medium districts were selected with equal probability. If sampling was needed in the large stratum, the large districts were sampled with probability proportional to enrollment. These probabilities were retained and used in later stages of sampling and weighting, as the district probability then represented the number of other districts that were not sampled to be surveyed for new schools.

The selected districts in each jurisdiction were then sent a listing of all their schools that appeared on the 2007-2008 CCD file and were asked to provide information about the new schools not included in the file and grade span changes of existing schools. These listings provided by the selected districts were used as sampling frames for selection of new public schools and updates of existing schools. This process was conducted through the NAEP State Coordinator in each jurisdiction. The coordinators were sent the information for all sampled districts in their respective states and were responsible for returning the completed updates.

The eligibility of a school was determined based on the grade span. A school also was classified as "newly-eligible" if a change of grade span had occurred such that the school status changed from ineligible to eligible in a particular grade.

School and Student Participation in the 2011 State Assessment

In all cases in the 2011 state assessment for grades 4 and 8, the weighted response rates for schools in each jurisdiction exceeded the 85 percent standard established by the National Center for Education Statistics (NCES). As participation is effectively mandatory, substitute schools for nonresponding schools were not provided.

In every NAEP survey, some of the sampled students are not assessed for the following reasons:

- withdrawn students,
- excluded students with disabilities (SD),
- excluded English language learner (ELL) students, or
- students absent from both the original session and the make-up session (not excluded but not assessed).

Withdrawn students are those who have left the school before the original assessment. Excluded students were determined by their school to be unable to meaningfully take the NAEP assessment in their assigned subject, even with an accommodation. Excluded students must also be classified as SD and/or ELL. Other students who were absent for the initial session are assessed in the makeup session. The last category includes students who were not excluded (i.e., "were to be assessed") but were not assessed either due to absence from both sessions or because of a refusal to participate. Assessed students are also classified as assessed without an accommodation or assessed with an accommodation. The latter group can be divided into SD students assessed with an accommodation, ELL students assessed with an accommodation, or students who are both SD and ELL and accommodated. Note that some SD and ELL students are assessed without accommodations, and students neither SD nor ELL can only be assessed without an accommodation.

The weighted response rates utilize the student base weights and indicate the weighted percentage of assessed students among all students to be assessed. The exclusion rates, in contrast, provide the weighted percentage of excluded SD or ELL students among all absent, assessed, and excluded students.

Response Rates of Fourth-Grade School Sample by Participating Jurisdiction

Response Rates of Eighth-Grade School Sample by Participating Jurisdiction

Weighted Student Response and Exclusion Rates, Mathematics Assessment

Weighted Student Response and Exclusion Rates, Reading Assessment

Weighted Student Response and Exclusion Rates, Science Assessment

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsgn_state_schl_and_stud_part.aspx

Response Rates of Eighth-Grade School Sample by Participating Jurisdiction for the 2011 State Assessment

The following table presents unweighted counts for sampled eligible and participating schools and weighted response rates for grade 8. States with Trial Urban District Assessment (TUDA) districts are shown in multiple rows: for the TUDA district(s) and for the state as a whole (the TUDA district[s] plus the rest of the state). The weighted response rates use the student enrollment from the sample frame divided by the school probability of selection. The weighted aggregation for the eligible schools for each jurisdiction is an estimate of the total population of students in the grade within each jurisdiction.

Participation is effectively mandatory for all states and districts, but not for Bureau of Indian Education (BIE) or Department of Defense Education Activity (DoDEA) schools.

School counts and response rates of sampled eligible schools, grade 8 state assessment, by jurisdiction: 2011

Jurisdiction	Number of sampled eligible schools	Number of participating schools	Weighted school response rates (percent)
Total	7,000	7,000	99.78
Alabama	113	113	100.00
Alaska	136	134	99.90
Arizona	126	125	99.02
Arkansas	123	123	100.00
California–Fresno	18	18	100.00
California–Los Angeles	69	69	100.00
California–San Diego	28	28	100.00
California	229	229	100.00
Colorado	122	121	99.87
Connecticut	110	110	100.00
Delaware	49	49	100.00
Florida–Hillsborough County	46	46	100.00
Florida–Miami	76	76	100.00
Florida	211	211	100.00
Georgia–Atlanta	21	21	100.00
Georgia	123	123	100.00
Hawaii	78	78	100.00
Idaho	106	106	100.00
Illinois–Chicago	113	113	100.00
Illinois	214	214	100.00
Indiana	106	106	100.00
Iowa	134	134	100.00
Kansas	137	137	100.00
Kentucky–Jefferson County	35	35	100.00
Kentucky	139	139	100.00
Louisiana	119	119	100.00
Maine	132	132	100.00
Maryland–Baltimore	55	55	100.00
Maryland	152	151	99.05
Massachusetts–Boston	36	36	100.00
Massachusetts	142	141	99.46
Michigan–Detroit	48	48	100.00
Michigan	156	156	100.00
Minnesota	143	143	100.00
Mississippi	110	110	100.00
Missouri	123	123	100.00
Montana	189	187	99.86
Nebraska	143	143	100.00
Nevada	90	89	99.70
New Hampshire	94	94	100.00
New Jersey	111	111	100.00
New Mexico–Albuquerque	29	29	100.00
New Mexico	126	125	99.09
New York–New York City	89	89	100.00
New York	170	169	99.08
North Carolina–Charlotte	35	35	100.00
North Carolina	151	151	100.00
North Dakota	187	186	99.99
Ohio–Cleveland	57	57	100.00
Ohio	168	168	100.00
Oklahoma	149	149	100.00
Oregon	137	136	99.10
Pennsylvania–Philadelphia	53	53	100.00
Pennsylvania	158	158	100.00
Rhode Island	53	53	100.00
South Carolina	109	109	100.00
South Dakota	225	225	100.00
Tennessee	117	117	100.00
Texas–Austin	19	19	100.00
Texas–Dallas	36	36	100.00
Texas–Houston	45	45	100.00
Texas	213	212	99.09
Utah	115	115	100.00
Vermont	123	123	100.00
Virginia	105	105	100.00
Washington	135	135	100.00
West Virginia	110	110	100.00
Wisconsin–Milwaukee	46	46	100.00
Wisconsin	159	159	100.00
Wyoming	88	88	100.00
Other jurisdictions			
Bureau of Indian Education (BIE)	111	94	83.16
Department of Defense Education Activity (DoDEA)	64	61	98.56
District of Columbia (TUDA)	36	36	100.00
District of Columbia	76	76	100.00

NOTE: Detail may not sum to totals because of rounding. Percentages are based on unrounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Assessment.

Response Rates of Fourth-Grade School Sample by Participating Jurisdiction for the 2011 State Assessment

The following table presents unweighted counts and weighted response rates at grade 4 for sampled eligible and participating schools. States with Trial Urban District Assessment (TUDA) districts are shown in multiple rows: for the TUDA district(s) and for the state as a whole (the TUDA district[s] plus the rest of the state). The weighted school response rates estimate the proportion of the student population that is represented by the participating school sample prior to substitution.

Participation is effectively mandatory for all states and districts, but not for Bureau of Indian Education (BIE) or Department of Defense Education Activity (DoDEA) schools.

School counts and response rates of sampled eligible schools, grade 4 state assessment, by jurisdiction: 2011

Jurisdiction	Number of sampled eligible schools	Number of participating schools	Weighted school response rates (percent)
Total	8,000	8,000	99.81
Alabama	108	107	98.95
Alaska	180	180	100.00
Arizona	122	121	99.03
Arkansas	122	122	100.00
California–Fresno	52	52	100.00
California–Los Angeles	75	75	100.00
California–San Diego	52	52	100.00
California	281	281	100.00
Colorado	122	122	100.00
Connecticut	109	109	100.00
Delaware	95	95	100.00
Florida–Hillsborough County	54	54	100.00
Florida–Miami	83	83	100.00
Florida	224	224	100.00
Georgia–Atlanta	62	62	100.00
Georgia	169	169	100.00
Hawaii	116	116	100.00
Idaho	131	131	100.00
Illinois–Chicago	94	94	100.00
Illinois	191	191	100.00
Indiana	110	110	100.00
Iowa	137	137	100.00
Kansas	140	139	99.18
Kentucky–Jefferson County	54	54	100.00
Kentucky	153	153	100.00
Louisiana	121	121	100.00
Maine	157	157	100.00
Maryland–Baltimore	69	69	100.00
Maryland	171	171	100.00
Massachusetts–Boston	75	75	100.00
Massachusetts	181	181	100.00
Michigan–Detroit	47	47	100.00
Michigan	151	151	100.00
Minnesota	136	136	100.00
Mississippi	106	106	100.00
Missouri	130	130	100.00
Montana	194	194	100.00
Nebraska	158	158	100.00
Nevada	116	116	100.00
New Hampshire	131	131	100.00
New Jersey	113	112	99.17
New Mexico–Albuquerque	52	52	100.00
New Mexico	151	151	100.00
New York–New York City	81	81	100.00
New York	156	156	100.00
North Carolina–Charlotte	57	57	100.00
North Carolina	167	167	100.00
North Dakota	254	253	99.97
Ohio–Cleveland	73	73	100.00
Ohio	179	179	100.00
Oklahoma	135	135	100.00
Oregon	142	141	99.08
Pennsylvania–Philadelphia	56	56	100.00
Pennsylvania	161	161	100.00
Rhode Island	114	114	100.00
South Carolina	108	108	100.00
South Dakota	194	194	100.00
Tennessee	116	116	100.00
Texas–Austin	55	55	100.00
Texas–Dallas	54	54	100.00
Texas–Houston	82	82	100.00
Texas	296	295	99.08
Utah	124	124	100.00
Vermont	219	219	100.00
Virginia	112	112	100.00
Washington	133	133	100.00
West Virginia	145	145	100.00
Wisconsin–Milwaukee	64	64	100.00
Wisconsin	182	182	100.00
Wyoming	184	184	100.00
Other jurisdictions			
Bureau of Indian Education (BIE)	132	112	83.26
Department of Defense Education Activity (DoDEA)	112	109	98.91
District of Columbia (TUDA)	83	83	100.00
District of Columbia	124	124	100.00

NOTE: Detail may not sum to totals because of rounding. Percentages are based on unrounded counts.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Assessment.

Weighted Student Response and Exclusion Rates for the 2011 State Mathematics Assessment

The following table presents the weighted student response and exclusion rates for the mathematics assessment. States with Trial Urban District Assessment (TUDA) districts are shown in multiple rows: for the TUDA district(s) and for the state as a whole (the TUDA district(s) plus the rest of the state). The weighted response rates utilize the student base weights and indicate the weighted percentage of assessed students as a percentage of all students to be assessed. The exclusion rates give the weighted percentage of excluded students with disabilities (SD) or students who are English language learners (ELL) among all absent, assessed, and excluded students.

Weighted student response and exclusion rates, state mathematics assessment, by grade and jurisdiction: 2011

Jurisdiction	Fourth grade			Eighth grade		
	Weighted student response rates (percent)	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded	Weighted student response rates (percent)	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded
Total	94.44	1.98	0.47	92.54	2.39	0.41
Alabama	95.01	1.12	0.07	93.72	1.16	0.05
Alaska	92.61	2.23	1.05	89.49	2.98	0.55
Arizona	94.29	1.02	0.10	92.96	1.14	0.03
Arkansas	94.87	0.91	0.13	92.53	1.26	0.20
California-Fresno	94.15	1.22	0.33	91.56	1.15	0.21
California-Los Angeles	94.87	1.73	0.86	92.39	1.26	0.73
California-San Diego	94.87	2.46	1.20	94.84	2.80	0.83
California	95.27	1.39	0.78	91.81	0.91	0.52
Colorado	92.20	1.07	0.15	92.71	0.82	0.21
Connecticut	93.37	1.20	0.20	92.90	1.19	0.29
Delaware	94.12	3.28	0.45	93.15	2.97	0.22
Florida-Hillsborough County	95.05	0.94	0.77	93.23	1.70	0.33
Florida-Miami	96.27	1.83	1.34	92.99	1.08	0.91
Florida	94.51	1.30	0.38	92.56	1.62	0.24
Georgia-Atlanta	96.13	0.94	0.10	92.80	2.38	0.19
Georgia	94.48	1.47	0.27	92.94	2.60	0.16
Hawaii	93.35	1.55	0.32	91.67	1.03	0.92
Idaho	95.34	1.01	0.31	94.32	1.19	0.20
Illinois-Chicago	94.44	2.09	0.85	95.57	2.95	1.06
Illinois	93.33	1.95	0.50	93.35	2.13	0.38
Indiana	94.69	2.10	0.11	92.93	2.47	0.17
Iowa	94.94	1.16	0.31	93.13	1.33	0.11
Kansas	94.26	1.50	0.20	93.02	1.27	0.06
Kentucky-Jefferson County	95.22	2.81	2.50	91.83	2.44	0.72
Kentucky	94.48	2.67	0.52	93.50	3.11	0.22
Louisiana	93.65	1.71	0.03	92.63	1.40	0.04
Maine	94.48	1.54	0.05	91.83	1.46	0.09
Maryland-Baltimore	93.05	11.11	0.19	87.44	12.14	0.62
Maryland	94.66	5.28	0.85	92.28	5.73	0.75
Massachusetts-Boston	93.74	3.28	2.76	92.02	4.42	2.91
Massachusetts	94.16	2.59	0.90	91.65	3.41	0.92
Michigan-Detroit	88.79	5.71	0.08	84.39	8.07	0.12
Michigan	94.11	1.92	0.23	92.84	3.26	0.41
Minnesota	94.02	1.37	0.17	93.04	1.79	0.44
Mississippi	94.99	0.76	0.10	93.56	1.04	0.04
Missouri	93.58	1.63	0.02	93.71	1.32	0.02
Montana	94.24	1.50	0.16	89.95	1.58	0.15
Nebraska	95.57	1.39	0.20	93.50	3.32	0.29
Nevada	94.93	2.19	0.48	93.73	2.74	1.08
New Hampshire	93.95	1.65	0.18	90.87	1.64	0.17
New Jersey	94.51	2.99	0.35	92.20	4.13	0.08
New Mexico-Albuquerque	93.15	2.20	1.22	89.29	2.64	1.63
New Mexico	93.89	2.19	1.10	91.32	1.72	0.74
New York-New York City	94.24	0.86	1.11	91.35	0.53	0.59
New York	94.13	0.91	0.58	91.03	1.03	0.37
North Carolina-Charlotte	94.44	1.01	0.17	92.39	1.11	0.38
North Carolina	94.01	1.65	0.37	91.83	1.68	0.21
North Dakota	95.26	3.26	0.47	94.66	4.22	0.13
Ohio-Cleveland	94.44	4.79	1.26	91.35	5.32	1.19
Ohio	94.03	2.18	0.22	92.54	5.01	0.06
Oklahoma	95.42	7.84	0.86	92.29	9.41	0.72
Oregon	93.47	2.30	0.88	93.11	1.39	0.11
Pennsylvania-Philadelphia	94.52	3.63	0.46	90.86	6.21	0.74
Pennsylvania	94.13	1.31	0.16	91.60	2.34	0.17
Rhode Island	94.41	0.84	0.09	92.06	0.99	0.30
South Carolina	94.15	1.23	0.07	93.57	3.52	0.30
South Dakota	95.31	1.76	0.15	94.31	1.33	0.44
Tennessee	93.79	3.24	0.29	91.46	3.66	0.17
Texas-Austin	94.12	3.42	1.87	91.43	3.71	1.83
Texas-Dallas	96.63	2.29	1.20	93.88	3.94	2.31
Texas-Houston	95.17	3.03	2.07	92.79	4.83	2.00
Texas	95.25	3.74	1.05	93.64	4.52	1.21
Utah	93.69	1.84	0.42	91.14	2.58	0.76
Vermont	94.12	1.46	0.12	93.96	1.12	0.00
Virginia	94.76	1.98	0.35	93.47	2.42	0.71
Washington	94.31	1.61	0.44	91.82	1.44	0.28
West Virginia	94.58	1.48	0.02	93.34	1.51	0.00
Wisconsin-Milwaukee	94.32	2.70	0.23	91.92	4.63	0.88
Wisconsin	94.98	1.65	0.25	92.92	1.90	0.18
Wyoming	93.72	1.55	0.13	92.42	1.25	0.00
Other jurisdictions						
Bureau of Indian Education (BIE)	91.89	1.38	0.88	91.20	1.85	0.87
Department of Defense Education Activity (DoDEA)	94.06	1.63	1.46	95.39	1.62	1.49
District of Columbia (TUDA)	93.97	5.47	1.13	88.34	5.44	1.32
District of Columbia	94.54	4.51	0.85	89.99	3.62	0.85

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Mathematics Assessment.

Weighted Student Response and Exclusion Rates for the 2011 State Reading Assessment

The following table presents the overall weighted student response and exclusion rates for the reading assessment. States with Trial Urban District Assessment (TUDA) districts are shown in multiple rows: for the TUDA district(s) and for the state as a whole (the TUDA district[s] plus the rest of the state). The weighted response rates utilize the student base weights and indicate the weighted percentage of assessed students as a percentage of all students to be assessed. The exclusion rates give the weighted percentage of excluded students with disabilities (SD) or students who are English language learners (ELL) among all absent, assessed, and excluded students.

Weighted student response and exclusion rates, state reading assessment, by grade and jurisdiction: 2011

Jurisdiction	Fourth grade			Eighth grade		
	Weighted student response rates (percent)	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded	Weighted student response rates (percent)	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded
Total	94.54	3.03	1.26	92.84	2.99	0.82
Alabama	95.29	2.22	0.12	94.17	1.86	0.21
Alaska	92.56	1.26	1.10	91.25	1.46	0.47
Arizona	94.43	1.37	0.16	93.73	1.13	0.09
Arkansas	94.70	1.19	0.14	93.81	1.40	0.17
California-Fresno	93.71	2.30	0.68	92.21	1.85	0.59
California-Los Angeles	95.11	1.83	1.05	91.67	1.86	1.30
California-San Diego	95.08	3.43	1.46	95.63	1.26	0.50
California	95.22	1.88	1.20	93.19	2.08	0.93
Colorado	92.67	1.15	0.39	92.12	1.16	0.56
Connecticut	93.99	1.42	0.98	92.28	1.37	1.03
Delaware	95.06	6.11	1.38	93.01	4.57	0.87
Florida-Hillsborough County	94.61	2.07	0.87	94.45	1.32	0.67
Florida-Miami	95.68	1.97	2.46	92.91	1.42	2.66
Florida	94.55	1.67	0.76	91.62	1.61	0.84
Georgia-Atlanta	96.23	3.16	0.94	92.38	3.07	0.48
Georgia	94.42	5.10	1.59	93.50	3.83	0.80
Hawaii	93.39	1.21	1.27	92.40	0.83	1.47
Idaho	95.46	1.61	0.27	94.08	1.45	0.47
Illinois-Chicago	95.27	1.42	1.41	94.92	1.78	0.90
Illinois	93.82	1.26	0.61	93.67	1.36	0.33
Indiana	95.23	1.12	0.11	92.91	1.89	0.32
Iowa	95.68	0.95	0.12	92.53	0.75	0.03
Kansas	95.15	1.81	0.62	93.46	1.81	0.16
Kentucky-Jefferson County	94.80	6.47	3.44	91.58	5.07	1.94
Kentucky	94.41	7.72	1.26	94.27	6.70	0.57
Louisiana	93.88	1.33	0.00	92.69	0.95	0.09
Maine	93.86	1.51	0.88	92.31	1.73	0.03
Maryland-Baltimore	92.70	15.39	1.83	88.94	15.86	1.09
Maryland	94.44	8.16	2.94	91.82	7.00	1.55
Massachusetts-Boston	94.48	5.84	4.44	89.97	5.33	6.43
Massachusetts	94.48	4.84	1.43	92.17	5.44	1.27
Michigan-Detroit	88.99	6.75	0.72	85.41	7.68	0.65
Michigan	94.40	3.33	0.25	93.15	4.48	0.50
Minnesota	94.46	1.47	0.19	92.58	2.67	0.32
Mississippi	93.78	0.94	0.11	92.33	0.86	0.10
Missouri	94.56	1.55	0.10	94.09	1.35	0.03
Montana	93.94	4.11	0.32	92.04	3.89	0.38
Nebraska	95.31	3.31	1.22	93.78	4.12	0.65
Nevada	95.59	1.06	0.22	92.86	1.57	0.67
New Hampshire	93.93	2.62	0.26	92.22	3.64	0.67
New Jersey	94.75	7.86	1.36	92.32	6.07	1.08
New Mexico-Albuquerque	92.87	4.04	2.21	88.93	4.35	4.11
New Mexico	93.43	3.79	3.05	91.25	4.20	2.41
New York-New York City	93.01	1.24	1.73	91.54	1.28	1.73
New York	93.75	1.53	1.29	91.32	2.15	1.23
North Carolina-Charlotte	94.57	1.35	0.55	92.97	1.57	0.97
North Carolina	93.81	2.07	0.29	92.09	1.78	0.43
North Dakota	96.01	5.84	1.14	93.50	6.91	1.46
Ohio-Cleveland	93.03	4.59	1.24	91.23	4.85	0.60
Ohio	94.22	5.40	0.58	93.25	5.48	0.37
Oklahoma	95.14	4.03	1.24	92.52	3.84	0.68
Oregon	94.59	2.47	0.67	92.32	1.97	0.34
Pennsylvania-Philadelphia	94.40	3.06	0.49	91.11	2.87	2.21
Pennsylvania	94.28	2.24	0.85	91.91	2.48	0.79
Rhode Island	95.01	1.57	0.54	92.66	0.76	0.41
South Carolina	94.30	2.41	0.47	93.72	4.65	0.96
South Dakota	95.71	2.88	0.57	94.70	2.70	0.59
Tennessee	94.71	6.66	0.62	91.99	5.94	0.46
Texas-Austin	94.26	9.42	10.15	93.15	6.78	3.60
Texas-Dallas	95.49	4.62	15.23	92.60	4.40	3.32
Texas-Houston	95.27	3.83	12.08	94.12	4.95	2.44
Texas	94.83	5.52	5.34	93.75	5.06	1.81
Utah	94.15	3.68	0.99	92.06	3.05	1.24
Vermont	93.54	2.21	0.18	93.05	2.44	0.49
Virginia	94.73	2.41	0.59	93.66	2.81	1.27
Washington	95.44	2.41	0.70	92.07	1.63	0.65
West Virginia	95.16	1.70	0.00	92.44	1.45	0.00
Wisconsin-Milwaukee	94.75	2.46	0.25	90.89	3.19	0.80
Wisconsin	94.53	1.65	0.39	93.81	1.96	0.39
Wyoming	94.66	1.74	0.33	92.67	1.55	0.60
Other jurisdictions						
Bureau of Indian Education (BIE)	91.04	1.62	0.75	89.52	2.01	0.84
Department of Defense Education Activity (DoDEA)	94.09	5.38	2.04	91.76	2.56	1.17
District of Columbia (TUDA)	94.99	3.02	0.97	87.69	2.64	1.49
District of Columbia	94.66	2.52	0.87	89.51	2.09	0.99

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Reading Assessment.

Weighted Student Response and Exclusion Rates for the 2011 State Science Assessment

The following table presents the overall weighted student response and exclusion rates for the science assessment. States with Trial Urban District Assessment (TUDA) districts are shown in multiple rows: for the TUDA district(s) and for the state as a whole (the TUDA district[s] plus the rest of the state). The weighted response rates utilize the student base weights and indicate the weighted percentage of assessed students as a percentage of all students to be assessed. The exclusion rates give the weighted percentage of excluded students with disabilities (SD) or English language learners (ELL) among all absent, assessed, and excluded students. Note that the table only includes those jurisdictions participating in the science assessment.

Weighted student response and exclusion rates, grade 8 state science assessment, by jurisdiction: 2011

Jurisdiction	Weighted student response rates (percent)	Weighted percentage of all students who are SD and excluded	Weighted percentage of all students who are ELL and excluded
Total	92.81	1.51	0.36
Alabama	93.22	1.02	0.10
Alaska	89.90	0.95	0.31
Arizona	93.19	0.85	0.03
Arkansas	94.08	0.95	0.10
California	92.96	1.76	0.62
Colorado	92.56	0.87	0.11
Connecticut	91.41	1.22	0.30
Delaware	92.10	1.53	0.20
Florida	93.10	0.92	0.32
Georgia	92.81	1.51	0.08
Hawaii	92.59	0.99	1.00
Idaho	93.06	1.32	0.25
Illinois	94.03	0.94	0.24
Indiana	93.83	1.25	0.05
Iowa	92.75	0.91	0.07
Kansas	94.45	1.28	0.14
Kentucky	93.04	2.27	0.47
Louisiana	93.36	1.17	0.05
Maine	92.66	1.74	0.13
Maryland	92.56	1.59	0.38
Massachusetts	92.20	2.77	0.72
Michigan	92.28	2.40	0.38
Minnesota	92.13	1.76	0.24
Mississippi	92.49	0.82	0.13
Missouri	93.44	1.23	0.05
Montana	91.02	1.53	0.15
Nebraska	94.57	1.27	0.17
Nevada	93.06	1.06	0.40
New Hampshire	90.78	1.95	0.24
New Jersey	91.77	1.12	0.21
New Mexico	91.94	1.58	0.78
New York	91.25	1.11	0.50
North Carolina	92.21	1.50	0.21
North Dakota	94.63	3.17	0.22
Ohio	92.62	2.13	0.02
Oklahoma	92.26	2.66	0.28
Oregon	92.65	1.55	0.24
Pennsylvania	93.28	1.03	0.03
Rhode Island	92.15	0.46	0.27
South Carolina	94.22	1.13	0.06
South Dakota	95.08	1.03	0.26
Tennessee	92.35	1.39	0.08
Texas	93.05	1.91	0.80
Utah	91.80	1.67	0.24
Vermont	93.95	1.29	0.11
Virginia	94.01	1.93	0.99
Washington	91.86	1.71	0.20
West Virginia	93.48	1.60	0.00
Wisconsin	93.21	1.78	0.13
Wyoming	92.27	1.26	0.13
Other jurisdictions			
Bureau of Indian Education (BIE)	88.42	0.00	0.00
Department of Defense Education Activity (DoDEA)	94.31	0.55	0.84
District of Columbia	87.56	1.13	0.50

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Science Assessment.

School Sample Selection for the 2011 State Assessment

The target student sample size per jurisdiction for a reading, mathematics, and science operational assessment was 3,150, 3,150 and 2,520 students, respectively. In the grade 4, jurisdictions had a target sample size of 6,600 which included the reading and mathematics assessments and 300 students for the pilot study. In grade 8, jurisdictions (except for Bureau of Indian Education [BIE] schools) had a target sample size of 8,820 which included the reading, mathematics, and science assessments plus additional students for the pilot and special study samples where the number of students sampled varied by the enrollment of the jurisdiction. By design, BIE schools did not participate in the science assessment, as it lacked the required number of students for the state science assessment. Thus, BIE schools had a target sample size of 6,600 which included the reading and mathematics assessments and 300 students for national science.

Computation of Measures of Size

School Sample Sizes: Frame and New School

The District of Columbia, which generally does not have enough students for an assessment in a third subject, also participated in the grade 8 science assessment. To accomplish this, each student in the District of Columbia was assigned to two of the three assessment subjects and thus tested twice.

Evaluation of State Achievement Data in the Sampling Frame

The general goal is to achieve a "self-weighting" sample at the student level; that is, as much as is possible, every eligible student should have the same probability of selection. Differences in the probability of selection among students introduce unwanted design effects, which increase the variance (reducing the marginal benefit of each added student).

When all students in a grade are taken in each sampled school, a self-weighting sample results from setting a fixed probability of selection across schools (as each student in the grade then has a probability of selection equal to the school probability of selection, which is equal across schools). When a fixed sample size of students (e.g., six) is taken in a selected grade in each sampled school, a self-weighting sample is achieved by taking a probability proportional-to-size (PPS) sample of schools, with size equal to the number of grade-eligible students in schools divided by a constant, such that the sum of the measures of size is the sample size. Each student then has a conditional probability of selection, which, when multiplied by the school's probability of selection, again gives equal unconditional probabilities of selection for students across schools.

There is also an added need to lower the expected number of very small schools in the sample, as the marginal cost for each assessed student in these schools is higher. These very small schools are sampled at half the rate of the larger schools, and their weights are doubled to account for the half sampling.

Schools were ordered within each jurisdiction using a serpentine sort (by urbanicity status, race/ethnicity status, and achievement score or zip code area median income). Next, a systematic sample was drawn with probability proportional to the measures of size, using a sampling interval of one. We refer to sampled schools as being "hit" in the sampling process.

Some larger schools had size measures larger than one. These schools may have been sampled more than once (i.e., they had multiple "hits"), meaning that a larger sample of students was selected from these schools.

The goal of deeply stratifying the school sample in each jurisdiction was to reflect the population distribution as closely as possible, thus minimizing the sampling error. The success of this approach was shown by comparing the proportion of minorities enrolled in schools (based on Common Core of Data values for each school), median income, and urban-centric locale (viewed as an interval variable) reported in the original frame against the school sample.

In addition, the distribution of state assessment achievement scores for the original frame can be compared with that of the school sample for those jurisdictions for which state assessment achievement data are available, as was done in the evaluation of state achievement data in the sampling frame. The number of significant differences found in this analysis is smaller than what would be expected to occur by chance, given the large number of comparisons that were made. The number of significant differences remained small even with the use of a finite population correction factor in the calculation of the sampling variances. However, the close adherence of sample values to frame values suggests that there is little evidence that the school sample for NAEP 2011 is not representative of the frame from which it was selected. The achievement/median income variable is used as the third-level sort order variable in the school systematic selection procedure. While it may be a rather low-level sort variable, it still helps control how representative the sampled schools are in terms of achievement. The close agreement between frame and sample values of these achievement/median income variables provides assurance that the selected sample is representative of the frame with respect to achievement status.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_schlsamp.aspx

Computation of Measures of Size for the 2011 Assessment

In designing each school sample, five objectives underlie the process of determining the probability of selection for each school and how many students are to be sampled from each selected school containing the respective grade:

- to meet the target student sample size for each grade;
- to select an equal-probability sample of students;
- to limit the number of students selected from any one school;
- to ensure that the sample within a school does not include a very high percentage of the students in the school, unless all students are included; and
- to reduce the rate of sampling of small schools, in recognition of the greater cost and burden per student of conducting assessments in such schools.

The goal in determining the school's measure of size is to optimize across the last four objectives in terms of maintaining the accuracy of estimates and the cost effectiveness of the sample design. In certain jurisdictions, a census of students was taken so as to meet, as nearly as possible, the target student sample size. Elsewhere, to meet the target student sample and achieve a reasonable compromise among the other four objectives above, the following algorithm was used to assign a measure of size to each school based on its enrollment per grade as indicated on the sampling frame.

The preliminary measures of size (MOS) were set as follows:

$$MOS_{js} = \begin{cases} x_{js} & \text{if } z_{js} < x_{js} \\ y_j & \text{if } 20 < x_{js} \leq z_{js} \\ \left(\frac{y_j}{20}\right) \times x_{js} & \text{if } 10 < x_{js} \leq 20 \\ \frac{y_j}{2} & x_{js} \leq 10 \end{cases}$$

where x_{js} is the estimated grade enrollment for jurisdiction j and school s , y_j the target within-school student sample size for jurisdiction j , and z_j the within-school take-all student cutoff for jurisdiction j .

For grade 4, the target sample sizes and take-all cutoffs were 63 and 70, respectively. For grade 8, the target sample sizes and take-all cutoffs varied by jurisdiction due to the pilot and special studies samples. The target sample sizes and take-all cutoffs ranged from 63 to 114 and 70 to 125, respectively. For the majority of the states and TUDAs, the target sample sizes were 89 and 63 and take-all cutoffs were 98 and 70, respectively.

The next task in this development is to describe b_j , the constant of proportionality for a specified jurisdiction. It is a sampling parameter that, when multiplied by a school's preliminary measure of size (MOS_{js}), yields the school's final measure of size. It is computed in such a way that, when used with the systematic sampling procedure, the target student sample size is achieved.

The final measure of size, E_{js} , is defined as:

$$E_{js} = \min\left(b_j \times MOS_{js}, u_j\right)$$

The quantity u_j (the maximum number of "hits" allowed) in this formula is designed to put an upper bound on the burden for the sampled schools. In most jurisdictions, u_j was set to 3. In Alaska, u_j was set to 8.

In addition, new and newly-eligible schools were sampled from the new-school frame. The assigned measures of size for these schools,

$$E_{js} = \min\left(b_j \times MOS_{js} \times \pi_{djs}^{-1}, u_j\right)$$

used the b_j and u_j values from the main school sample for the jurisdiction (i.e., the same sampling rates as for the main school sample within each jurisdiction). The variable π_{djs} is the probability of selection of the district into the new-school district (d) sample.

Evaluation of State Achievement Data in the Sampling Frame for the 2011 State Assessment

The purpose of this analysis was to determine whether public schools selected for the 2011 samples were representative of the schools on the NAEP sampling frames in terms of student achievement. Percentiles of the achievement distributions were compared between the frame and sample schools for each public school jurisdiction in grades 4, 8, and 12.

Achievement Data

The achievement variable used in the analysis was the same variable used in the NAEP sample design to stratify the public school frame. For most jurisdictions, the variable was an achievement score provided by the jurisdiction. However, for some jurisdictions where achievement data were not available, median household income from the 2000 Census was used. (In 2000, the Census determined median household income based on the five-digit zip code area in which the school was located.) The achievement data consisted of various types of school-specific achievement measures from state assessment programs. The type of achievement data available varied by jurisdiction. For instance, in some states, the measure was the average score for a given state assessment. In other states, the measure was a percentile rank or percentage of students above a specific score.

During frame development, not every record on the Common Core of Data (CCD) file matched to the achievement data files created for the National Center for Education Statistics (NCES), even in jurisdictions where those data were generally available. For schools that did not match, their achievement score was imputed by a mean matching imputation approach using the mean achievement score for schools with complete achievement data within the same jurisdiction/urbanicity/race/ethnicity stratum combination.

Methodology

To determine whether the distributions between the frame and sample schools were different, comparisons of percentile estimates were made for the 10th, 25th, 50th, 75th, and 90th percentile levels as well as the mean for each public school jurisdiction by grade. Frame and sample school estimates were considered statistically different if the frame value fell outside the 95 percent confidence interval of the corresponding sample estimate. The percentile values for the frame schools were calculated by weighting each school by the estimated number of students in the given grade. The percentile estimates for the sample schools were calculated using school weights and weighted by the school measure of size (estimated number of students in the given grade). The 95 percent confidence intervals for the school sample estimates were calculated in WesVar—software for computing estimates of sampling variance from complex sample survey (Westat, 2000b)—using the Woodruff method (Samdal, Swensson, and Wretman 1992) and without the use of a finite population correction factor.

Results

As mentioned above, sample and frame achievement distributions were determined to be different if at least one of the percentile estimates or the mean differed significantly at the 95 percent confidence level. Out of all the jurisdiction and grade comparisons, only 14 of the 738 distributions compared were found to be significantly different. They are shown in the table below.

Summary of significant differences in achievement measures between the sample and the frame, state assessment, by jurisdiction and grade: 2011

Grade	Jurisdiction	Achievement data / median income	Estimate	Frame	Sample	Confidence interval
4	Maryland	Achievement data	10th percentile	14.98	13.30	(13.26, 13.95)
	Oklahoma	Achievement data	mean	83.46	84.27	(83.65, 84.89)
	Detroit TUDA	Achievement data	90th percentile	28.41	28.27	(28.14, 28.40)
	Fresno TUDA	Achievement data	50th percentile	38.88	38.18	(36.94, 38.87)
	Fresno TUDA	Achievement data	75th percentile	47.07	45.10	(44.62, 46.18)
	Miami TUDA	Achievement data	90th percentile	85.73	84.80	(84.48, 85.22)
8	Alaska	Median Income	10th percentile	36,350.99	35,919.49	(35,837.17, 36,046.46)
	Illinois	Achievement data	50th percentile	83.87	82.91	(82.57, 83.76)
	Maine	Achievement data	75th percentile	60.06	58.20	(56.59, 58.98)
	Maine	Achievement data	90th percentile	69.10	67.43	(66.68, 69.01)
	Maine	Achievement data	mean	51.58	51.04	(50.61, 51.46)
	Mississippi	Achievement data	mean	54.31	55.21	(54.51, 55.92)
	New Jersey	Achievement data	10th percentile	36.35	35.40	(33.25, 36.22)
	New York City TUDA	Achievement data	50th percentile	47.33	45.96	(45.52, 46.43)

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Assessment.

The number of significant differences found in this analysis is smaller than what would be expected to occur by chance, given the large number of comparisons that were made. The number of significant differences remained small even with the use of a finite population correction factor in the calculation of the sampling variances. However, the close adherence of sample values to frame values suggests that there is little evidence that the school sample for NAEP 2011 is not representative of the frame from which it was selected. The achievement/median income variable is used as the fourth-level sort order variable in the school systematic selection procedure. While this variable was low in the sorting hierarchy, it still helps control how representative the sampled schools are in terms of achievement. The close agreement between frame and sample values of these achievement/median income variables provides assurance that the selected sample is representative of the frame with respect to achievement status.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_schlsamp_samp_eval.aspx

School Sample Sizes: Frame and New School for the 2011 State Assessment

The following table lists the number of sampled schools taken from the public school sampling frame (as constructed from the Common Core of Data) and the new-school sampling frame, for fourth and eighth grade, by participating jurisdiction. The school counts shown are at the time of sampling. After school sampling, it was determined that in some Trial Urban District Assessments (TUDAs) a few schools did not contribute to the TUDA's Adequate Yearly Progress (AYP). These schools were then classified as out of scope for the TUDA but in scope for the state.

NAEP state frame-based and new public school samples, state assessment, by grade and jurisdiction: 2011

Jurisdiction	Grade 4			Grade 8		
	Total school sample	Frame school sample	New school sample	Total school sample	Frame school sample	New school sample
Total	8,500	8,300	228	7,700	7,300	366
Alabama	117	117	0	125	121	4
Alaska	202	199	3	167	164	3
Arizona	127	126	1	136	134	2
Arkansas	123	121	2	126	123	3
California-Fresno	55	53	2	28	26	2
California-Los Angeles	82	81	1	78	73	5
California-San Diego	58	58	0	39	35	4
California-Balance	97	97	0	112	109	3
Colorado	124	124	0	130	126	4
Connecticut	116	115	1	119	112	7
Delaware	109	97	12	68	61	7
Florida-Hillsborough County	56	55	1	50	48	2
Florida-Miami	88	81	7	86	73	13
Florida-Balance	91	88	3	97	95	2
Georgia-Atlanta	66	63	3	26	25	1
Georgia-Balance	108	103	5	106	104	2
Hawaii	118	117	1	81	73	8
Idaho	137	131	6	113	107	6
Illinois-Chicago	95	92	3	117	111	6
Illinois-Balance	98	98	0	106	105	1
Indiana	117	117	0	113	112	1
Iowa	141	138	3	138	138	0
Kansas	148	145	3	148	147	1
Kentucky-Jefferson County	56	55	1	46	42	4
Kentucky-Balance	102	102	0	110	107	3
Louisiana	134	122	12	163	136	27
Maine	166	164	2	143	140	3
Maryland-Baltimore	70	68	2	75	48	27
Maryland-Balance	103	101	2	101	101	0
Massachusetts-Boston	86	78	8	45	35	10
Massachusetts-Balance	116	111	5	109	107	2
Michigan-Detroit	58	58	0	63	60	3
Michigan-Balance	114	114	0	117	117	0
Minnesota	148	146	2	168	164	4
Mississippi	117	115	2	121	118	3
Missouri	131	130	1	136	135	1
Montana	206	205	1	200	200	0
Nebraska	181	179	2	169	168	1
Nevada	118	113	5	100	93	7
New Hampshire	133	133	0	96	95	1
New Jersey	118	117	1	116	116	0
New Mexico-Albuquerque	57	55	2	48	46	2
New Mexico-Balance	98	94	4	88	81	7
New York-New York City	82	81	1	91	86	5
New York-Balance	76	76	0	83	80	3
North Carolina-Charlotte	57	54	3	38	36	2
North Carolina-Balance	116	115	1	121	120	1
North Dakota	272	268	4	209	190	19
Ohio-Cleveland	86	79	7	71	69	2
Ohio-Balance	119	116	3	123	118	5
Oklahoma	140	139	1	149	149	0
Oregon	149	142	7	144	135	9
Pennsylvania-Philadelphia	58	58	0	61	58	3
Pennsylvania-Balance	109	105	4	107	103	4
Rhode Island	125	124	1	61	57	4
South Carolina	114	112	2	116	113	3
South Dakota	209	202	7	261	254	7
Tennessee	119	117	2	123	120	3
Texas-Austin	55	54	1	24	21	3
Texas-Dallas	55	54	1	41	38	3
Texas-Houston	86	83	3	50	46	4
Texas-Balance	110	101	9	121	113	8
Utah	130	126	4	125	122	3
Vermont	226	225	1	124	123	1
Virginia	115	112	3	108	107	1
Washington	141	141	0	140	137	3
West Virginia	152	150	2	117	116	1
Wisconsin-Milwaukee	69	68	1	60	55	5
Wisconsin-Balance	121	118	3	119	116	3
Wyoming	202	191	11	108	95	13
Other jurisdictions						
Bureau of Indian Education (BIE)	135	131	4	116	111	5
Department of Defense Education Activity (DoDEA)	120	110	10	72	60	12
District of Columbia (TUDA)	106	102	4	50	34	16
District of Columbia-Balance	47	28	19	52	29	23

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), 2011 State Assessment.

Stratification of Schools in the 2011 State Assessment

The purpose of school stratification is to increase the efficiency and ensure the representativeness of the school samples in terms of important school-level characteristics, such as geography (e.g., states and TUDA districts), urbanicity, and race/ethnicity classification. NAEP school sampling utilizes two types of stratification: explicit and implicit.

Stratification
Variables

Explicit stratification partitions the sampling frame into mutually exclusive groupings called strata. The systematic samples selected from these strata are independent, meaning that each is selected with its own unique random start. The explicit school strata for the 2011 NAEP state assessments were usually states. If a state contained Trial Urban District Assessment (TUDA) districts, the explicit strata were each individual TUDA district and the balance of the state. In 2011, there were 21 participating TUDA districts in the NAEP state assessment program. They are listed below. The ones in bold are those introduced in 2011.

- **Albuquerque Public Schools, New Mexico;**
- Atlanta Public Schools, Georgia;
- Austin Independent School District, Texas;
- Baltimore City Public Schools, Maryland;
- Boston Public Schools, Massachusetts;
- Charlotte-Mecklenburg Schools, North Carolina;
- Chicago Public Schools, Illinois;
- Cleveland Metropolitan School District, Ohio;
- **Dallas Independent School District, Texas;**
- Detroit Public Schools, Michigan;
- District of Columbia Public Schools, District of Columbia;
- Fresno Unified School District, California;
- **Hillsborough County Public Schools, Florida;**
- Houston Independent School District, Texas;
- Jefferson County Public Schools (Louisville) , Kentucky;
- Los Angeles Unified School District, California;
- Miami-Dade County Public Schools, Florida;
- Milwaukee Public Schools, Wisconsin;
- New York City Department of Education, New York;
- School District of Philadelphia, Pennsylvania; and
- San Diego Unified School District, California.

Implicit stratification involves sorting the sampling frame, as opposed to grouping the frame. For NAEP, schools are sorted by key school characteristics within explicit strata and sampled systematically using this ordering. This type of stratification ensures the representativeness of the school samples with respect to the key school characteristics. The implicit school stratification variables for the 2011 state assessments included urbanicity, race/ethnicity classification, and achievement score/median income. Further details about these variables can be found here.

http://nces.ed.gov/nationsreportcard/tdw/sample_design/2011/2011_sampdsgn_state_strat.aspx

Stratification Variables for the 2011 State Assessment

The implicit stratification of public schools for the NAEP 2011 state assessments involved three dimensions:

- urbanicity classification (urban-centric locale);
- race/ethnicity classification; and
- achievement level or median income.

Stratification by Urbanicity Classification

Stratification by Race/ethnicity Classification

Stratification by Achievement Data and Median Income

Missing Stratification Variables

The urbanicity stratum is the top-level implicit stratification variable and is assigned within each explicit stratum. It is derived from the NCES urban-centric locale variable and classifies schools based on location (city, suburb, town, rural) and proximity to urbanized areas. It has 12 possible values.

The race/ethnicity stratum classifies schools by the relative magnitude of non-Hispanic White, non-Hispanic Black, Hispanic, Asian and Pacific Islander, and American Indian and Alaska Native enrollments represented in schools. The source of the race/ethnicity data is the Common Core of Data (CCD). The race/ethnicity stratum is the second-level variable in the stratification hierarchy and is nested within the urbanicity stratum.

The last stratification dimension is a classification of schools based on either achievement data or median household income. For most states, it is based on achievement data. However, not all states provide achievement data. In these cases, median household income is used instead. Median income comes from the 2000 Census and it corresponds to the zip code area where the school is located.

Missing values for stratification variables were imputed.

The implicit stratification in this three-fold hierarchical procedure was achieved via a "serpentine sort" within a given explicit stratum. This sort was accomplished by alternating between ascending and descending sort order on each variable successively through the sort hierarchy. Within this sorted list the schools were arranged in serpentine order by achievement data (or median household income) within each cell determined by the two higher stratification variables (urbanicity and race/ethnicity classifications), with ascending order for achievement data/median household income used in every other cell, and descending order for achievement data/median household income used in the remaining cells, giving an ascending-descending-ascending-descending pattern. Schools in these urbanicity and race/ethnicity classification cells were also sorted in serpentine order. Within each urbanicity and race/ethnicity classification cells, schools were sorted in ascending order within one urbanicity stratum, by descending order within the next urbanicity stratum, and so on. The following table shows an oversimplified example to illustrate the ascending-descending-ascending-descending pattern of the serpentine sort.

Stratification variables sorted by serpentine sort: 2011

TUDA	Urbanicity	Race/ethnicity level	Achievement score	
Yes	Large City	High minority	20	
			22	
			27	
		30		
		Low minority	29	
			26	
	20			
	Mid-size City	Low minority	18	
			15	
			25	
		High minority	27	
			31	
			35	
	No	Mid-size City	High minority	32
				30
				28
			Low minority	20
				22
27				
Large City		Low minority	30	
			29	
			26	
		High minority	20	
			18	
			15	
Large City		Low minority	25	
			27	
			31	
		High minority	35	
			32	
			30	
			28	

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Assessment.

The third dimension of stratification differed for schools in the National Indian Education Study (NIES) oversample. These schools were implicitly stratified by the percentage American Indian/Alaskan Native students within the school instead of achievement scores or median household income.

Missing Stratification Variables for the 2011 State Assessment

Schools with missing stratification variables had their data imputed as follows:

Schools with missing estimated grade enrollment had their estimated grade enrollment set to 20. Schools missing the urbanicity (urban-centric locale) variable were assigned the modal value of urbanicity for schools in the same five-digit zip code or the same city. The mean ethnicity percentage was imputed at the five-digit zip code level only if all schools were missing ethnicity at the district level, and only at the three-digit zip code prefix if the five-digit zip code mean was missing as well.

Schools with missing or questionable values in race/ethnicity enrollment data—those in which the summation of the ethnicity percentages did not fall in the range 97 through 103, indicating a gross error—were assigned the average race/ethnicity enrollment within their school district, five-digit zip code, or three-digit zip code prefix.

Schools with missing achievement data in jurisdictions and grades for which achievement data were used in stratification were assigned the mean achievement data value within their urbanization and race/ethnicity classification. The achievement data were imputed only for those schools in jurisdictions and grades in which achievement data were used for stratification.

Schools missing median household income were assigned the mean value of median household income for the three-digit zip code prefix in which they were located. In some cases, imputation was not possible at the three-digit zip code level, and needed to be done at the city and state level.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsgn_state_strat_var_missing.aspx

Stratification by Achievement Data and Median Income for the 2011 State Assessment

The achievement data are derived from the results of state assessment programs that were obtained from each jurisdiction. The contents of the achievement data files varied by jurisdiction and included achievement measures for a variety of subjects, grades, and multiple assessment programs. One achievement measure was selected for each responding jurisdiction to be used in the stratification process. Where available, the achievement data were used for implicit stratification by grade. Since the achievement data are more current than the median household income data, as well as more likely to be well-correlated to NAEP assessment scores, they were judged to be a more effective stratification variable. The achievement measures were selected according to the following criteria:

Jurisdictions Using Achievement Data or Median Household Income in Stratification

- Achievement measures from state assessments conducted in mathematics and reading (in that order of priority) were utilized, if available. For grade 4, data from fourth-grade assessments were used, if available; otherwise, data from third-grade assessments. For grade 8, data from eighth-grade assessments were used, if available; otherwise, data from seventh-grade assessments. For both grades, data from 2007 assessments (the latest available) were used.
- Achievement measures should match to at least 70 percent of the schools on the sampling frames.
- Achievement measures should differentiate schools from one another. For example, district-level measures, those with high missing rates or pass/fail indicators, were judged not to be useful for differentiating schools. In addition, achievement measures that did not have good dispersion were not used for stratification.
- All other things being equal, the possibilities for score types were average scale score, median scale score, percentile rank, median percentile rank, normal curve equivalent, raw score, index score, and percentage above a particular cut score or quartile. In general, the availability varied for any given state/grade/subject/year.

Achievement data useful for implicit stratification were obtained from 47 of 51 jurisdictions for both fourth- and eighth-grade assessments. Where achievement data were not used, median household income from the 2000 Census was used.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsgn_state_strat_var_achieve.aspx

Jurisdictions Using Achievement Data or Median Household Income in Stratification for the 2011 State Assessment

This table shows whether achievement data or median household income was used as a stratification variable for participating jurisdictions. Neither achievement nor median income data was available for stratification of Bureau of Indian Education (BIE) and Department of Defense Education Activity (DoDEA) schools. The estimated grade enrollment was used in these two jurisdictions.

Type of data, achievement or median household income, used for stratification, state assessment, by grade and jurisdiction: 2011

Jurisdiction	Grade 4		Grade 8	
	Achievement	Income	Achievement	Income
Alabama	YES	NO	YES	NO
Alaska	NO	YES	NO	YES
Arizona	YES	NO	YES	NO
Arkansas	YES	NO	YES	NO
California	YES	NO	YES	NO
Colorado	YES	NO	YES	NO
Connecticut	YES	NO	YES	NO
Delaware	YES	NO	YES	NO
Florida	YES	NO	YES	NO
Georgia	YES	NO	YES	NO
Hawaii	YES	NO	YES	NO
Idaho	YES	NO	YES	NO
Illinois	YES	NO	YES	NO
Indiana	YES	NO	YES	NO
Iowa	YES	NO	YES	NO
Kansas	YES	NO	YES	NO
Kentucky	YES	NO	YES	NO
Louisiana	YES	NO	YES	NO
Maine	YES	NO	YES	NO
Maryland	YES	NO	YES	NO
Massachusetts	YES	NO	YES	NO
Michigan	YES	NO	YES	NO
Minnesota	YES	NO	YES	NO
Mississippi	YES	NO	YES	NO
Missouri	YES	NO	YES	NO
Montana	NO	YES	NO	YES
Nebraska	NO	YES	NO	YES
Nevada	YES	NO	YES	NO
New Hampshire	YES	NO	YES	NO
New Jersey	YES	NO	YES	NO
New Mexico	YES	NO	YES	NO
New York	YES	NO	YES	NO
North Carolina	YES	NO	YES	NO
North Dakota	YES	NO	YES	NO
Ohio	YES	NO	YES	NO
Oklahoma	YES	NO	YES	NO
Oregon	YES	NO	YES	NO
Pennsylvania	YES	NO	YES	NO
Rhode Island	YES	NO	YES	NO
South Carolina	YES	NO	YES	NO
South Dakota	YES	NO	YES	NO
Tennessee	YES	NO	YES	NO
Texas	YES	NO	YES	NO
Utah	YES	NO	YES	NO
Vermont	YES	NO	YES	NO
Virginia	YES	NO	YES	NO
Washington	YES	NO	YES	NO
West Virginia	YES	NO	YES	NO
Wisconsin	YES	NO	YES	NO
Wyoming	YES	NO	YES	NO
Other jurisdictions				
Bureau of Indian Education (BIE)	—	—	—	—
Department of Defense Education Activity (DoDEA)	—	—	—	—
District of Columbia	NO	YES	NO	YES

— Not available.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 State Assessment.

Stratification by Race/Ethnicity Classification for the 2011 State Assessment

Race/ethnicity classification was based on the second and third largest race/ethnicity percentages (among non-Hispanic White, non-Hispanic Black, Hispanic, Asian, and American Indian and Alaska Native students) within each urbanicity classification stratum. The race/ethnicity strata were formed using one of three classification schemes as follows:

Case 1: Urbanicity cells where both the second and third largest race/ethnicity groups contained less than 7 percent of students in the urbanicity cell were not stratified by race/ethnicity enrollment (race/ethnicity stratification value was set to 0). There were no race/ethnicity strata formed within these urbanicity cells.

Case 2: Urbanicity cells where the second largest race/ethnicity group contained at least 7 percent but no more than 15 percent of students in the urbanicity cell were stratified into three race/ethnicity cells. Schools were ordered by the sum of the percentage of race/ethnicity enrollment for the second and third largest groups within the urbanicity cell and then divided into three approximately equal size groups in terms of students.

Case 3: Urbanicity cells where both the second and third largest race/ethnicity groups contained more than 15 percent of students in the urbanicity cell were stratified into four race/ethnicity cells. The second largest group provided the primary stratification variable; the third largest group provided the secondary stratification variable. Within an urbanicity cell, schools were first sorted based on the primary stratification variable. Then they were divided into two strata of schools containing approximately equal numbers of students. Within each of these two strata, the schools were sorted by the secondary stratification variable and subdivided into two substrata of schools containing approximately equal numbers of students. The four race/ethnicity classifications consisted of the following values; low primary variable/low secondary variable, low primary variable/high secondary variable, high primary variable/low secondary variable, and high primary variable/high secondary variable.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_strat_var_race_eth.aspx

Stratification by Urbanization Classification for the 2011 State Assessment

The creation of the urbanicity classification variable was based on the NCES urban-centric locale and was defined within each explicit stratum. The NCES urban-centric locale contains the following categories:

1. Large City: Territory inside an urbanized area and inside a principal city with population of 250,000 or more.
2. Mid-size City: Territory inside an urbanized area and inside a principal city with population less than 250,000 and greater than or equal to 100,000.
3. Small City: Territory inside an urbanized area and inside a principal city with population less than 100,000.
4. Large Suburb: Territory outside a principal city and inside an urbanized area with population of 250,000 or more.
5. Mid-size Suburb: Territory outside a principal city and inside an urbanized area with population less than 250,000 and greater than or equal to 100,000.
6. Small Suburb: Territory outside a principal city and inside an urbanized area with population less than 100,000.
7. Fringe Town: Territory inside an urban cluster that is less than or equal to 10 miles from an urbanized area.
8. Distant Town: Territory inside an urban cluster that is more than 10 miles and less than or equal to 35 miles from an urbanized area.
9. Remote Town: Territory inside an urban cluster that is more than 35 miles of an urbanized area.
10. Fringe Rural: Census-defined rural territory that is less than or equal to 5 miles from an urbanized area, as well as rural territory that is less than or equal to 2.5 miles from an urban cluster.
11. Distant Rural: Census-defined rural territory that is more than 5 miles but less than or equal to 25 miles from an urbanized area, as well as rural territory that is more than 2.5 miles but less than or equal to 10 miles from an urban cluster.
12. Remote Rural: Census-defined rural territory that is more than 25 miles from an urbanized area and is also more than 10 miles from an urban cluster.
13. Outside of the United States: Department of Defense Education Activity (DoDEA) overseas schools.

For the definitions of the geographic terms used in these descriptions, please refer to the Census Bureau's website (for example, www.census.gov/geo/www/ua/ua_2k.html; www.census.gov/population/www/metroareas/aboutmetro.html)

The urbanicity classification cells were created by starting with the original NCES urban-centric locale categories. Urbanicity strata were collapsed with neighboring strata until a minimum cell size criterion, in terms of the percentage of students, was met. The minimum cell size criterion varied by type of explicit stratum. The criterion for explicit strata comprising the largest TUDA districts (Los Angeles, New York City, Chicago, Miami, and Houston) was 12 percent; for the other TUDA districts, it was 18 percent; and for all other explicit strata, it was 9 percent.

The urbanicity classification variable was equal to the original NCES urban-centric locale if no collapsing was necessary. If collapsing was necessary, the collapsing scheme first collapsed within the four major strata (city, suburbs, town, rural), in this order. For example, urbanicity categories 1, 2, and 3 within City were collapsed in order of the list (1 with 2, 2 with 3) if cells 1 or 3 were deficient. If the middle cell (e.g., 2) was deficient, then it was collapsed with the smaller of the two end cells. If a collapsed pair was still deficient, it was collapsed with the remaining unit within the major stratum. That is, a single category would be created by combining large city, mid-size city and small city categories. If these collapsed cells were still inadequate, they were further collapsed with all three types of suburb cells to form a single cell made up of large, mid-size and small cities and large, mid-size and small suburbs. The values of the urbanicity classification variable were set equal to the cell value of the final level of collapsing.

Prior experience with this type of stratification has shown that the greatest efficiency of stratification results when cities and suburb fringe areas are always kept separate from towns and rural areas, even if the enrollment criterion is violated.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_strat_var_urban.aspx

Student Sample Selection for the 2011 State Assessment

Within each sampled school, a sample of students was selected from a listing of the students in the grade such that every student had an equal chance of selection. The student lists were submitted in multiple ways. E-filing is an electronic submission system. Excel files are submitted for sampled schools by school coordinators or NAEP State Coordinators. Files can be submitted for one school at a time or for an entire jurisdiction at once. This method allows schools to easily submit student demographic data electronically with the student lists, easing the burden on NAEP field supervisors and school coordinators. Schools that are unable to submit their student lists using the e-filing system provide hard copy lists via the student listing form to NAEP field supervisors. In 2011, there were 18,023 schools that E-filed their student lists, while 781 lists were submitted using the student listing form.

In year-round, multi-track schools, students who were not scheduled to be in school on the assessment day were removed from the student lists prior to sampling. Student base weights were adjusted to account for these students.

The sampling process was the same, regardless of list submission type. The sampling process was systematic (e.g., if the sampling rate was one-half, a random starting point of one or two was chosen, and every other student on the list was selected). For E-filed schools only, where demographic data was submitted for every student on the frame, students were sorted by gender and race/ethnicity before the sample was selected to implicitly stratify the sample.

In the certainty jurisdictions, all students were sampled in all schools. Otherwise, the sample size for grade 4 was 63 students. The sample sizes for grade 8 varied due to the pilot and special study samples. Very small states did not have any pilot test/special study sample and larger states had somewhat more pilot test/special study sample than other states. This is so that the samples for the pilot tests and special studies would be reasonably nationally representative. The sample size for grade 8 ranged from 63 to 114 students with 89 students sampled in the majority of the jurisdictions. Larger schools may have been selected more than once in the sampling process and thus may have a larger sample size. In addition, most fourth-grade schools chose the option of taking all students when enrollment was less than 120. This increased the fourth-grade sample size in many states beyond the designated target.

Some students enrolled in the school after the sample was selected. In such cases, new enrollees were sampled at the same rate as the students on the original list.

In fourth-grade schools, sampled students were randomly assigned to mathematics, reading, and pilot as follows: 30 for mathematics, 30 for reading, and 3 for pilot. In eighth-grade Bureau of Indian Education (BIE) schools that did not participate in the state science assessment, the sampled students were randomly assigned to reading, mathematics, and science as follows: 30 for mathematics, 30 for reading, and 3 for national science. In the other eighth-grade schools that did participate in the state science assessment, students were randomly assigned to mathematics, reading, and science as follows: 30 for mathematics, 30 for reading, 24 for science, plus a varying number for pilot and special studies. This was implemented by spiraling: the booklets assigned to sampled students were provided from booklet packets that had, on average, equal numbers of each of the relevant assessments in a randomized order.

Some of the students who were English language learners (ELL) or students with disabilities (SD) were excluded from the assessment because they could not be assessed with the accommodations NAEP provides.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_studsamp.aspx

Target Population for the 2011 State Assessment

The target population for the 2011 state assessment included all students in public schools in the United States who were enrolled in the fourth or eighth grades. In addition, students attending Bureau of Indian Education (BIE) and Department of Defense Education Activity (DoDEA) schools who were enrolled in the fourth and eighth grades were also included. U.S. territories, although included in some past NAEP assessments, were not included in NAEP 2011.

http://nces.ed.gov/nationsreportcard/dw/sample_design/2011/2011_sampdsn_state_targpop.aspx