

## NAEP Frequently Asked Questions (FAQs)

### NAEP and Its Uses

#### **What is NAEP?**

NAEP, or the National Assessment of Educational Progress, produces the Nation's Report Card, to inform the public about the academic achievement of elementary and secondary students in the United States. Sponsored by the Department of Education, NAEP assessments have been conducted periodically in reading, mathematics, science, writing, U.S. history, civics, geography, and other subjects, beginning in 1969. NAEP collects and reports academic achievement at the national level, and for certain assessments, at the state and district levels. The results are widely reported by the national and local media, and are an integral part of our nation's evaluation of the condition and progress of education.

#### **What are the goals of the NAEP program?**

NAEP has two major goals: to compare student achievement in states and other jurisdictions and to track changes in achievement of fourth-, eighth-, and twelfth-graders over time in mathematics, reading, writing, science, and other content domains. To meet these dual goals, NAEP selects nationally representative samples of students who participate in either the main NAEP assessments or the long-term trend NAEP assessments.

#### **Why is NAEP important?**

Long considered to be the "gold standard" of assessments, NAEP serves as the federal government's official measure of how well students in states and across the nation are performing in core academic subjects over time. Additionally, NAEP has taken on a greater prominence under the No Child Left Behind Act and serves to externally confirm results of state assessments.

#### **What is the difference between state NAEP and national NAEP?**

The NAEP sample in each state is designed to be representative of the students in that state. At the state level, results are currently reported for public school students only and are broken down by several demographic groupings of students. When NAEP is conducted at the state level (i.e., in mathematics, reading, science, and writing), results are also reported for the nation. The national NAEP sample is then composed of all the state samples of public school students, as well as a national sample of nonpublic school students. If there are states that do not participate, a certain number of schools and students are selected to complete the national-level sample.

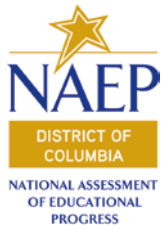
For assessments conducted at the national level only, samples are designed to be representative of the nation as a whole. Data are reported for public and nonpublic school students as well as for several major demographic groups of students.

#### **What are the key differences between NAEP and state assessments?**

State assessments are administered to all students in specific grades while NAEP state level assessments are administered to representative samples of 4<sup>th</sup>, 8<sup>th</sup>, and 12<sup>th</sup> graders. State tests directly measure students' knowledge of state standards. NAEP measures the cumulative knowledge of students and not necessarily what they have been taught in the current school year. NAEP exams are timed assessments that require essay and short-answer responses. As the National Assessment Governing Board stresses *"NAEP and state assessments serve different purposes and are used together to inform educational policy."*

#### **Are the data confidential?**

Federal law dictates complete privacy for all test takers and their families. Under the National Assessment of Educational Progress Authorization Act (Public Law 107-279 III, section 303), the Commissioner of the National



## NAEP Frequently Asked Questions (FAQs)

Center for Education Statistics (NCES) is charged with ensuring that NAEP tests do not question test-takers about personal or family beliefs or make information about their personal identity publicly available.

After publishing NAEP reports, NCES makes data available to researchers but withholds students' names and other identifying information. The names of all participating students are not allowed to leave the schools after NAEP assessments are administered. Because it might be possible to deduce from data the identities of some NAEP schools, researchers must promise, under penalty of fines and jail terms, to keep these identities confidential.

### **Who evaluates NAEP?**

Because NAEP findings have an impact on the public's understanding of student academic achievement, precautions are taken to ensure the reliability of these findings. In its current legislation, as in previous legislative mandates, Congress has called for an ongoing evaluation of the assessment as a whole. In response to these legislative mandates, the National Center for Education Statistics (NCES) has established various panels of technical experts to study NAEP, and panels are formed periodically by NCES or external organizations, such as the National Academy of Sciences, to conduct evaluations. The Buros Center for Testing, in collaboration with the University of Massachusetts/Center for Educational Assessment and the University of Georgia, more recently conducted an external evaluation of NAEP.

### **What subjects does NAEP assess and how are the subjects chosen?**

Since its inception in 1969, NAEP has assessed numerous academic subjects, including mathematics, reading, science, writing, the arts, civics, economics, foreign language, geography, technology and engineering literacy, and U.S. history.

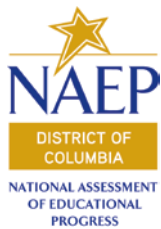
Beginning with the 2003 assessments, NAEP national and state assessments are conducted in mathematics and reading at least once every two years at grades 4 and 8. These assessments are conducted in the same year and initial results are released six months after administration, in the fall of that year. Results from all other assessments are released about one year after administration, usually in the spring of the following year. Many NAEP assessments are conducted at the national level for grade 12, as well as at grades 4 and 8.

Since 1988, the National Assessment Governing Board has selected the subjects assessed by NAEP. Furthermore, the Governing Board oversees creation of the frameworks that underlie the assessments and the specifications that guide the development of the assessment instruments. The framework for each subject area is determined through a collaborative development process that involves teachers, curriculum specialists, subject-matter specialists, school administrators, parents, and members of the general public.

### **What process is used to develop the assessments?**

To meet the nation's growing need for information about what students know and can do, the NAEP assessment instruments must meet the highest standards of measurement reliability and validity. They must measure change over time and must reflect changes in curricula and instruction in diverse subject areas.

Developing the assessment instruments—from writing questions to analyzing pilot test results to constructing the final instruments—is a complex process that consumes most of the time during the interval between assessments. In addition to conducting national pilot tests, developers oversee numerous reviews of the assessment instruments by internal NAEP measurement experts, by the National Assessment Governing Board, and by external groups that include representatives from each of the states and jurisdictions that participate in the NAEP program. To find out



## NAEP Frequently Asked Questions (FAQs)

how a typical assessment is developed, see [How Are NAEP Assessments Developed?](#) For more technical details on development of the assessments, see [NAEP Instruments](#).

### NAEP Results

#### **What results does NAEP provide?**

Subject-matter achievement is reported in two ways—scale scores and achievement levels—so that student performance can be more easily understood. NAEP scale score results provide a numeric summary of what students know and can do in a particular subject and are presented for groups of students. Achievement levels categorize student achievement as Basic, Proficient, and Advanced, using ranges of performance established for each grade. (A fourth category, below Basic, is also reported for this scale.) Achievement levels are used to report results in terms of a set of standards for what students should know and be able to do.

NAEP provides results about subject-matter achievement, instructional experiences, and school environment and reports these results for populations of students (e.g., fourth-graders) and groups within those populations (e.g., male students or Hispanic students). NAEP does not provide individual scores for the students or schools assessed.

Because NAEP scales are developed independently for each subject, scale score and achievement level results cannot be compared across subjects. However, these reporting metrics greatly facilitate performance comparisons within a subject from year to year and from one group of students to another in the same grade. Examples of student responses can be accessed through the [NAEP Questions Tool](#).

#### **How are the results reported to the public?**

NAEP has developed a number of different publications and web-based tools that provide direct access to assessment results at the state and national level. For every major assessment release, web-specific content is developed that is suitable to the web environment.

The Nation's Report Card is a website developed especially to display the results of each assessment in a clear format and comprehensive manner. To locate this useful information, there are links to the most recent results from any subject information page on this website. See, for instance, the link on the mathematics subject page. State Profiles present state-level results and a history of state participation in NAEP; District Profiles report results for the Trial Urban District Assessment (TUDA) participants.

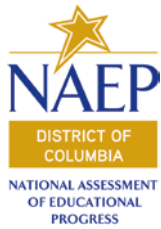
The NAEP Data Explorer and State Comparisons provide comprehensive information on student performance. Explore NAEP Questions links users to the Questions Tool and Item Maps that provide student responses, scoring guides, and other information on the questions that have been released to the public.

#### **How does NAEP reliably score and process millions of student-composed responses?**

While multiple-choice questions allow students to select an answer from a list of options, constructed-response questions require students to provide their own answers. Qualified and trained raters score constructed-response questions.

Scoring a large number of constructed responses with a high level of reliability and within a limited time frame is essential to NAEP's success. (In a typical year, over three million constructed responses are scored.) To ensure reliable, quick scoring, NAEP takes the following steps:

- develops focused, explicit scoring guides that match the criteria delineated in the assessment frameworks;
- recruits qualified and experienced scorers, trains them, and verifies their ability to score particular questions through qualifying tests;



## NAEP Frequently Asked Questions (FAQs)

- employs an image-processing and scoring system that routes images of student responses directly to the scorers so they can focus on scoring rather than paper routing;
- monitors scorer consistency through ongoing reliability checks;
- assesses the quality of scorer decision-making through frequent monitoring by NAEP assessment experts; and
- documents all training, scoring, and quality control procedures in the technical reports.

NAEP assessments generally contain both constructed-response and multiple-choice questions. The constructed responses are scored using the image-processing system, whereas the responses to the multiple-choice questions are scored by scanning the test booklets.

### How does NAEP analyze the assessment results?

Before the data are analyzed, responses from the groups of students assessed are assigned sampling weights to ensure that their representation in NAEP results matches their actual percentage of the school population in the grades assessed.

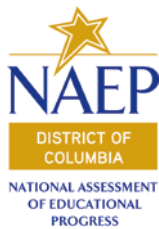
Data for national and state NAEP assessments in most subjects are analyzed by a process involving the following steps:

- **Check Item Data and Performance:** The data and performance of each item are checked in a number of ways, including scoring reliability checks, item analyses, and differential item functioning (DIF), to assure fair and reliable measures of performance in the subject of the assessment.
- **Set the Scale for Assessment Data:** Each subject assessed is divided into subskills, purposes, or content domains specified by the subject framework. Separate scales are developed relating to the content domains in an assessment subject area. A special statistical procedure, Item Response Theory scaling, is used to estimate the measurement characteristics of each assessment question.
- **Estimate Group Performance Results:** Because NAEP must minimize the burden of time on students and schools by keeping assessment administration brief, no individual student takes more than a small portion of the assessment for a given content domain. NAEP uses the results of scaling procedures to estimate the performance of groups of students (e.g., of all fourth-grade students in the nation, of female eighth-grade students in a state).
- **Transform Results to the Reporting Scale:** Results for assessments conducted in different years are linked to reporting scales to allow comparison of year-to-year trend results for common populations on related assessments.
- **Create a Database:** A database is created and used to make comparisons of all results, such as scale scores, percentiles, percentages at or above achievement levels, and comparisons between groups and between years for a group. All comparisons are subjected to testing for statistical significance, and estimates of standard errors are computed for all statistics.

To ensure reliability of NAEP results, extensive quality control and plausibility checks are carefully conducted as part of each analysis step. Quality control tasks are intended to verify that analysis steps have not introduced errors or artifacts into the results. Plausibility checks are intended to encourage thinking about the results, whether they make sense, and what story they tell.

### What contextual data are provided in NAEP?

In addition to assessing subject-area achievement, NAEP collects information that serves to fulfill reporting requirements of federal legislation and to provide a context for reporting student performance. The legislation



## NAEP Frequently Asked Questions (FAQs)

requires that, whenever feasible, NAEP include information on special groups (e.g., information reported by race, ethnicity, socioeconomic status, gender, disability, and limited English proficiency).

As part of most NAEP assessments, several types of questionnaires are used to collect information. The questionnaires appear in separately timed blocks of questions in the assessment booklets, such as the student questionnaires, or, as in the case of questionnaires for the teachers, schools, and students with disabilities or who are classified as English language learners, they are printed separately.

### **When will the results be available?**

Beginning with the 2003 assessment, results in mathematics and reading are to be released six months after the administration of the assessments, except when either of these assessments is based on a new framework. For instance, for the 2009 reading assessment, the Governing Board issued a new framework which resulted in changes to the assessment requiring additional data analyses to examine the measurement of trends over time; consequently, the Reading Report Card for 2009 was released in the spring of 2010. Results from all other assessments will be released one year after administration.

### **Can my school get school-level or individual student-level results?**

No. By design, information will not be available at these levels. Reports traditionally disclose state, regional, and national results. In 2002, NAEP began to report (on a trial basis) results from several large urban districts in the Trial Urban District Assessment (TUDA), but school-level results are not yet reportable. Because NAEP is a large-group assessment, each student takes only a small part of the overall assessment. In most schools, only a small portion of the total grade enrollment is selected to take the assessment and these students may not reliably or validly represent the total school population. Only when the student scores are aggregated at the state or national level are the data considered reliable and valid estimates of what students know and can do in the content area; consequently, school- or student-level results are never reported.

## Administration of NAEP

### **Does the state select the schools that will participate in NAEP?**

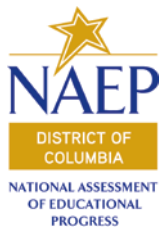
No. As part of a contract with the U.S. Department of Education, Westat, a statistical survey research organization located in Rockville, MD, selects the schools and students that will participate in NAEP.

### **How are schools selected for NAEP?**

NAEP uses a multistage sampling design that relies on stratification (i.e., classification into groups having similar characteristics) to choose samples of schools and students. Samples are randomly selected from groups of schools that have been classified according to variables such as the extent of urbanization, the percentage of minority enrollment, and school-level performance results on state assessments.

### **Who are the students assessed by NAEP?**

The national results are based on a representative sample of students in public schools, private schools, Bureau of Indian Education schools, and Department of Defense schools. Private schools include Catholic, Conservative Christian, Lutheran, and other private schools. The state results are based on public school students only. The main NAEP assessment is usually administered at grades 4 and 8 (at the state level) plus grade 12 at the national level.



## NAEP Frequently Asked Questions (FAQs)

### **Who administers the NAEP to students?**

NAEP representatives, hired by the U.S. Department of Education, administer all assessment sessions. School officials, including classroom teachers, may observe the assessments. The field staff will bring all necessary testing materials and hardware including computers.

### **How many schools and students participate in NAEP and when are the data collected during the school year?**

The number of students selected to be in a NAEP sample depends on whether it is a national-only sample or a combined state and national sample. In the national-only sample, there are approximately 10,000 to 20,000 students. In a combined national and state sample, there are approximately 3,000 students per participating jurisdiction from approximately 100 schools. Typically, 45 to 55 jurisdictions participate in such an assessment.

Other NAEP special studies can occur at different points throughout the school year.

The numbers of schools and students for each recent assessment are available on the website for The Nation's Report Card. Technical information about this may be found in Study Design and the Data Collection Plan.

### **Duties of the School Coordinator**

#### **When is the NAEP testing window?**

The NAEP 2016 assessment window is January 25, 2016 through March 4, 2016.

#### **What are the responsibilities of selected schools?**

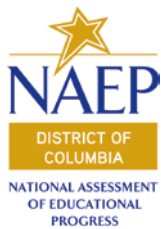
Schools are responsible for identifying students who may need accommodations, planning testing locations, grouping the students taking the assessment, notifying parents, and getting students to their testing locations. School administrators and teachers are not required to be involved in the actual administration of the assessment; however, school staff members should be available for proctoring each group session. Although the state's fall student data collection is used to provide NAEP with the initial needed student information, schools will be asked to furnish an updated student list with demographics the first week in January. Beginning with the 2013-14 school year, pre-assessments activities are conducted online by school level test coordinators designated by their principals.

NAEP requires that parents of students selected for NAEP assessments be notified in writing that their child has been or may be selected for assessment and that each child's participation is voluntary. Prior to the assessment, a dated copy of the information given to parents must be provided to The NAEP representatives. Schools also will need to keep a log of any parent refusals. A sample parent notification letter that may be adapted to satisfy this requirement is available.

#### **How are students with disabilities and English language learners included in the NAEP assessments?**

The NAEP program has always endeavored to assess all students selected as a part of its sampling process. In all NAEP schools, accommodations will be provided as necessary for students with disabilities (SD) and/or English language learners (ELL).

Inclusion in NAEP of an SD or ELL student is encouraged if that student (a) participated in the regular state academic assessment in the subject being tested, and (b) if that student can participate in NAEP with the accommodations NAEP allows. Even if the student did not participate in the regular state assessment, or if he/she needs accommodations NAEP does not allow, school staff are asked whether that student could participate in NAEP with the allowable accommodations. (Examples of accommodations not allowed in NAEP are giving the reading



## NAEP Frequently Asked Questions (FAQs)

assessment in a language other than English, or reading the reading passages aloud to the student. Also, extending testing over several days is not allowed for NAEP because NAEP administrators are in each school only one day.)

### **How can educators use NAEP resources such as frameworks, released questions, and reports in their work?**

NAEP materials such as frameworks, released questions, and reports have many uses in the educational community. For instance, frameworks can serve as models for designing an assessment or revising curricula. Also, released constructed-response questions and their corresponding scoring guides can serve as models of innovative assessment practices.

### **If a student is selected for NAEP, must he/she participate?**

Student participation in NAEP is voluntary. Under federal law, parental notification by schools prior to testing is required to inform families that students who are sampled for the assessment may opt not to participate. To ensure that our state is part of NAEP, and that educators, parents, policymakers, and citizens can learn how our state performs compared to other states and the nation as a whole, it is very important that all students selected for NAEP actually participate.

### **How much time is required of students participating in NAEP?**

Students selected to participate in the NAEP pencil and pencil version will take a 50-minute test in one subject area. Students also complete a short background questionnaire. In all, no more than 90 minutes is required for the entire experience. Students selected to take the test on a NAEP supplied tablet may take a little longer to complete the assessment.

## NAEP Assessment Sample Design

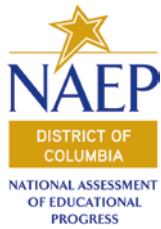
Each assessment cycle, a sample of students in designated grades within both public and private schools throughout the United States (and sometimes specified territories and possessions) is selected for assessment. In addition, in state assessment years, of which 2017 is an example, the samples of public schools and their students in each state are large enough to support state-level estimates. In all cases, the selection process utilizes a probability sample design in which every school and student has a chance to be selected, and standard errors can be calculated for the derived estimates.

### *Public School Selection in State Assessment Years*

The selection of a sample of public school students for state assessment involves a complex multistage sampling design with the following stages:

- Select public schools within the designated areas,
- Select students in the relevant grades within the designated schools, and
- Allocate selected students to assessment subjects.

The Common Core of Data (CCD) file, a comprehensive list of operating public schools in each jurisdiction that is compiled each school year by the National Center for Education Statistics (NCES), is used as the sampling frame for the selection of sample schools. The CCD also contains information about grades served, enrollment, and location of each school. In addition to the CCD list, a set of specially sampled jurisdictions is contacted to determine if there are



## NAEP Frequently Asked Questions (FAQs)

any newly formed public schools that were not included in the lists used as sampling frames. Considerable effort is expended to increase the survey coverage by locating public schools not included in the most recent CCD file.

As part of the selection process, public schools are combined into groups known as strata on the basis of various school characteristics related to achievement. These characteristics include the physical location of the school, extent of minority enrollment, state-based achievement scores, and median income of the area in which the school is located. Stratification of public schools occurs within each state. Grouping schools within strata by such selected characteristics provides a more ordered selection process with improved reliability of the assessment results.

On average, a sample of approximately 100 grade-eligible public schools is selected within each jurisdiction; within each school, about 60 students are selected for assessment. Both of these numbers may vary somewhat, depending on the number and enrollment size of the schools in a jurisdiction, and the scope of the assessment in the particular year. Students are sampled from a roster of individual names, not by whole classrooms. The total number of schools selected is a function of the number of grades to be assessed, the number of subjects to be assessed, and the number of states participating.

### *National-Only Assessment Years*

In years when the NAEP samples are intended only to provide representation at the national level and not for each individual state, the public and private school selection process is somewhat different. Rather than selecting schools directly from lists of schools, the first stage of sampling involves selecting a sample of some 50 to 100 geographic primary sampling units (PSUs). Each PSU is composed of one or more counties. They vary in size considerably, and generally about 1,000 PSUs are created in total, from which a sample is selected. Within the set of selected PSUs, public and private school samples are selected using similar procedures to those described above for the direct sampling of schools from lists. The samples are clustered geographically, which results in a more efficient data collection process. The selection of PSUs is not necessary when the sample sizes are large in each state, as in state assessment years.

National Center for Educational Statistics, (2011). NAEP Technical Documentation. *NAEP Assessment Sample Design*. Retrieved from [http://nces.ed.gov/nationsreportcard/tdw/sample\\_design/](http://nces.ed.gov/nationsreportcard/tdw/sample_design/)