

# Frequently Asked Questions About NAEP Sampling

*Developed in support of reporting NAEP 2011 Science results*

## Who is assessed by NAEP?

- NAEP assesses representative samples of students in certain grades or at certain ages in public and nonpublic schools in the United States.
- For the national assessments, NAEP samples students from grades 4, 8, and 12 in public and nonpublic schools.
- For long-term trend assessments, NAEP samples students at ages 9, 13, and 17.
- For the state assessments, NAEP samples students from grades 4 and 8 and assesses students in reading, mathematics, writing, and science. As part of the *No Child Left Behind* legislation, NAEP assesses fourth- and eighth-grade students every two years in reading and mathematics (beginning in 2003). Other subjects (science, writing) vary from year to year.
- NAEP does not provide scores for individual students or schools; instead it offers aggregate results regarding subject-matter achievement, instructional experiences, and school environment for populations of students (e.g., grade 4) and subgroups of those populations (e.g., female students, Hispanic students). NAEP results are based on samples of these student populations of interest.

## How many students are selected for the state and national assessments?

- For the state assessments, NAEP samples approximately 2,500 to 3,500 students in the state in each grade for each subject area. To do this, NAEP samples between 100 and 300 schools in the state at each grade (4 and 8), and students are sampled within those schools.
- From each school in the state assessment, NAEP sampled about 30 students per grade for each subject. In 2011, approximately 70 students were sampled for grade 4 from each school for reading and mathematics. For grade 8 approximately 90 students were assessed from each school in states participating in all three state subject assessments—reading, mathematics and science. At grade 4, schools were allowed an option to assess all students (up to 120). It is possible large schools may have a larger sample sizes. This is related to the student population size and the number of schools in the state.

- In a handful of states (Washington, Oregon, Arizona, North Carolina, Minnesota, and Utah) the public school sample at grades 4 and 8 was increased somewhat, so as to give publishable results for AI/AN students. This affected school sampling only. There was no special student sampling procedures for this purpose. This was achieved by increasing, by an appropriate factor, the measures of size of school.
- The national samples for the 4<sup>th</sup> and 8<sup>th</sup> grade 2011 reading and mathematics assessments contain the combined sample of students assessed in reading and mathematics in each state. Beginning with the 2002 assessments, NAEP has selected a combined sample of public schools for state and national NAEP rather than selecting separate state and national samples. This approach reduces the burden on states and schools by decreasing the total number of schools participating in state and national NAEP. The full data set is analyzed together, allowing all data to contribute to the final results, and setting a single scale for the assessment and improving the reliability of the national estimates.
- A separate, national sample of nonpublic schools is also selected for grades 4, 8 and 12. This sample is designed to produce national estimates and estimates for two types of nonpublic schools (Catholic and non-Catholic Private).

### **How does NAEP select the schools and the students for the assessments?**

- As the Nation's Report Card, NAEP must report accurate results for populations of students and subgroups of these students (e.g., minority students, students from low-income families). To ensure accurate results, the relatively small samples of students must be truly representative of the entire student population in the nation (for the national assessments) or the state (for the state assessments) or the district (for the TUDA assessments).
- NAEP uses a multistage sampling design that relies on stratification of schools (i.e., classification into groups having similar characteristics). To ensure an accurate representation of public schools, schools are stratified by variables such as the extent of urbanization, race/ethnicity composition, and school level results on state achievement tests or median household income levels in the zip code where the school is located. After the stratification process, systematic sampling is utilized to choose the samples of schools. For national assessments not involving state-by-state samples, region of the country and median household income of the area where the school is located are also used. A similar approach is used for nonpublic schools, but using somewhat different characteristics.
- NAEP selects a representative sample of students by first selecting schools using a probability sampling design and then selecting the students within those schools who will participate in a given NAEP assessment. Every eligible school has a known nonzero chance of being selected for the sample. Within a selected school, all students within a participating grade have an equal chance of being selected. The probability of students and schools being selected into the sample varies based on factors such as grade, subject, public and nonpublic school status, school size and

so on. Those probabilities are important in producing NAEP results, and NAEP takes them into account in the calculation of results through the process of applying sampling weights. The overall goal of the sampling process is that every eligible student within a state has the same probability of selection.

### **Why are some schools always selected?**

- NAEP usually selects around 100 to 300 public schools for each subject at each grade for the state's sample. If a school is chosen repeatedly, typically it is because it has more than about 1 percent of the state's student enrollment in the grade. Other schools, with about 0.5–1 percent of the enrollment, are selected frequently though not *always*—however, it probably seems like *always* to those schools. Smaller schools may be selected repeatedly with certainty in states with a small student population in the target grades. An example in 2009 is North Dakota 8th grade, with an estimated total of about 8,000 students statewide. In order to meet the target student sample size of approximately 9,000 students for three subjects, NAEP took all schools and all students for 8th grade in North Dakota for the 2009 assessments.

### **Why does NAEP use sampling? What are the benefits of sampling for NAEP?**

- Sampling produces accurate estimates of student achievement while reducing the amount of time and cost to administer and score the assessment. Administering NAEP to all students in a state or the nation would be very expensive—there are many constructed-response questions, which are very expensive to score.
- NAEP does not report data for individual students, schools, or districts—except for a few districts participating in the TUDA assessment—and, therefore, it is not necessary to assess and report results for every student in every school.
- Sampling minimizes the assessment time required per student while allowing complete coverage of the subject being assessed. In order to have a valid and reliable assessment of the NAEP content, several hundred assessment questions are needed. Testing a student on the entire collection of assessment questions that make up each NAEP assessment is too time consuming and impractical. Hence, no single student takes the entire assessment.

## **What is matrix sampling and what are its advantages for NAEP?**

- In matrix sampling, different portions from the entire pool of assessment questions are printed in separate booklets and administered to different but equivalent samples of students. Matrix sampling allows NAEP to assess the entire subject area within a reasonable amount of testing time.
- Students are only assessed in one subject area taking two blocks of items.
- In each school, some students are assessed in each subject. In most states, about an equal number of students in each school will take reading, math, science