



# NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP) MATHEMATICS ASSESSMENT

## Summary of Frameworks and Assessment Design

### Chronology of NAEP Mathematics assessment administration

- National and State samples:
  - Grade 12 at the national level in 1996, 2000, 2005, 2009, 2011, and 2013. Grade 12 at the state level was administered for the first time in 11 states in 2009. In 2013 it will be administered as a state assessment in 13 states and as a national assessment.
  - Grade 4 in 1992, 1996, 2003, 2005, 2007, 2009, 2011, and 2013. (Mathematics was assessed in 2000 but Florida did not participate).
  - Grade 8 in 1990, 1992, 1996, 2003, 2005, 2007, 2009, 2011, and 2013. (Mathematics was assessed in 2000 but Florida did not participate).
- TUDA samples:
  - TUDA Grades 4 and 8 in 2002, 2003, 2005, 2007, 2009, 2011, and 2013.

### Students chosen to be assessed in NAEP

- NAEP assessments are administered to a stratified random sample of students from grades 4, 8, and 12.
- Both public school and nonpublic school students in grades 4 and 8 are assessed at the national level.
- At the state level, only the results of public school students are reported.
- Fifty-two jurisdictions participate in the NAEP Mathematics assessment—the 50 states, the District of Columbia Public Schools, and the Department of Defense Education Activity schools.
- Accommodations are offered to English language learners (ELLs), students with a 504 plan, and students with disabilities (SD) who have Individual Education Plans (IEPs). The most typical accommodations include:
  - extra testing time
  - individual or small-group administrations
  - large-print booklets

### Mathematics content assessed

- All NAEP frameworks are developed under the guidance of the National Assessment Governing Board (NAGB).
- The NAEP 2011 Mathematics Framework was used for the first time for grades 4 and 8 in the early 1990's.

- The grade 12 mathematics framework for 2009 was an update of the grade 12 framework adopted by NAGB in 2005. The 2005 framework was developed to reflect changes in high school standards and coursework.
- A framework specifies what is to be assessed and how it is to be assessed. The mathematics framework can be accessed at <http://www.nagb.org/publications/frameworks/math-framework09.pdf>.
- A NAEP framework does not specify how a subject should be taught nor does it prescribe a particular curricular approach to teaching.
- The NAEP mathematics framework specifies:
  - five broad areas of content
    1. Number Properties and Operations
    2. Measurement
    3. Geometry
    4. Data Analysis, Statistics, and Probability
    5. Algebra
  - three levels of mathematical complexity (attempts to focus on the cognitive demands of the items made on students)
    1. Low complexity – questions typically specify what a student is to do, which is often to carry out a routine mathematical procedure
    2. Moderate complexity – questions involve more flexibility of thinking and often require a response with multiple steps
    3. High complexity – questions make heavier demands and often require abstract reasoning or analysis in a unique situation

### **Mathematics assessment administration**

- Assessment takes approximately 90 minutes to administer. This includes two separately timed blocks of items, each 25 minutes in length, and background questions.
- Questions are presented in two formats:
  - multiple-choice
  - constructed-response
    - short answer constructed-response questions - require a one- or two-sentence answer
    - extended answer constructed-response questions - require a paragraph or full-page response
- Distribution of items on the NAEP Mathematics assessment:
  - differs by grade level to reflect the knowledge and skills appropriate for each
  - distribution also reflects the relative importance and value given to each content area
  - percentages refer to the numbers of items, not the amount of student testing time

## Percentage Distribution by Content Area and Grade

Content Area	Grade 4	Grade 8	Grade 12
Number properties and operations	40%	20%	10%
Measurement	20%	15%	30%
Geometry	15%	20%	30%
Data analysis, Statistics, and Probability	10%	15%	25%
Algebra	15%	30%	35%

- Some of the assessment booklets include manipulative materials, where possible, to measure a student's ability to represent his or her understanding and ability to use tools to solve problems (e.g., rulers, protractors).
- Calculators:
  - Students with calculator booklets are allowed to use calculators on approximately one-third of the assessment.
  - The assessment contains blocks of questions for which calculators are not allowed and blocks of questions that require calculators.
    - Grade 4 students use a four-function calculator supplied by NAEP
    - Grade 8 students use their own scientific calculator or one supplied by NAEP
- Testing time on NAEP Mathematics is divided evenly among multiple-choice items and both types of constructed-response items.
- Half of the total testing time is spent on items of moderate complexity, with the remainder of the total time spent equally on items of low and high complexity.
- Because each block is spiraled with other blocks and is administered to a representative sample of students, the results can be combined to produce average group and subgroup results based on the entire assessment.
- Background questions gather information that is used for analyzing a number of student demographic and instructional factors related to student achievement. Questions such as, "How many books are there in your home?" are asked and four different number ranges are provided for the student to choose from.

## Scoring criteria for constructed-response items used on all NAEP assessments

- Unique scoring guides are developed for each question.
- Scoring guides describe the specific criteria for assigning a score level to student responses.
- Scoring process:
  - Expert scorers are extensively trained how to determine what score level to assign to student responses.
  - Scoring is monitored to ensure the scoring standards are being adhered to reliably
  - Monitoring measures the consistency of scoring the same items administered in different assessments—therefore ensuring consistency of application standards across assessment years
  - The percentage of exact agreement among raters of the same student responses are tracked

## Reporting NAEP scores on all NAEP assessments

- Results are used to compile national, state, and selected urban district data. No results are generated for schools or individual students.
- National results reflect the performance of all grades 4, 8, and 12 students in public schools, private schools, Bureau of Indian Education (BIE) schools, and Department of Defense Activity schools.
- National Public and State results reflect the performance of students in public schools in grades 4 and 8.
- Overall results are reported for various groups of students: by race/ethnicity, eligibility for the National School Lunch Program, students with disabilities (SD), English language learners (ELL), and by gender.
- NAEP reports scores in two different ways: average scale scores and achievement levels. Both scores are based on the performance of samples of students, not the entire population.
- Mathematics average scale scores indicate how much a student *knows and can do* based on a 0–500 scale. The scores are reported as:
  - Average scale scores (range from 0–500)
  - Percentiles (10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, and 90<sup>th</sup>), which show trends in performance for lower-, middle-, and higher-performing students
- Achievement levels
  - offer a means of identifying percentages of students who have demonstrated certain proficiencies
  - are performance standards based on scale scores and define the degree to which student performance meets expectations of what students *should know and be able to do*
  - are set by NAGB
    - *Advanced* represents superior performance
    - *Proficient* represents solid academic performance. Students reaching this level have demonstrated competency over challenging subject matter
    - *Basic* denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at a given grade
  - Below *Basic* is reported, but is not considered to be an achievement level.
  - Achievement level descriptors for grade 4 mathematics can be found at <http://nces.ed.gov/nationsreportcard/mathematics/achieveall.asp#grade4>.
  - Achievement level descriptors for grade 8 mathematics can be found at <http://nces.ed.gov/nationsreportcard/mathematics/achieveall.asp#grade8>.
  - Achievement level descriptors for grade 12 mathematics can be found at <http://nces.ed.gov/nationsreportcard/mathematics/achieveall.asp#grade12>.

## Interpreting NAEP scores for all NAEP assessments

- Differences between average scale scores or between achievement-level percentages are discussed in NAEP reports only when they are statistically significant. Statistically significant means it is unlikely the differences in scores occurred by chance. The differences are referred to as “significant differences” or as being “significantly different.”
- NAEP assesses a representative sample of students in each state and in selected urban districts. The number of students tested determines the standard error for a particular jurisdiction. Because of sample design, performance standard error must be considered in reporting NAEP results. Statistical tests that factor in the standard errors are used to determine whether the differences are significant at the 0.05 level.
- Estimates based on smaller groups are likely to have relatively large standard errors. In these cases, some seemingly large differences may not be statistically significant.
- Results data for all NAEP assessments can be found at the NAEP Data Explorer Web site at <http://nces.ed.gov/nationsreportcard/nde>. Reports of Florida’s NAEP results can be found at <http://www.fldoe.org/asp/naep/>.