Assessment and Item Specifications for the NAEP 2009 Mathematics Assessment

National Assessment of Educational Progress

MATHEMATICS

NAEP

2009

June 2007

National Assessment Governing Board
U.S. Department of Education

CHAPTER ONE

INTRODUCTION

Since 1973 the National Assessment of Educational Progress (NAEP) has gathered information about student achievement in mathematics. Results of these periodic assessments, produced in print and web-based formats, provide valuable information to a wide variety of audiences. They inform citizens about the nature of students' comprehension of the subject, curriculum specialists about the level and nature of student achievement, and policymakers about factors related to schooling and its relationship to student proficiency in mathematics.

The NAEP assessment in mathematics is accomplished in two different ways. One assessment measures long-term trends in achievement among 9, 13, and 17-year old students using the same basic design each time. This unique measure has allowed for comparisons of students' knowledge of mathematics since the assessment was first administered in 1973. The main NAEP assessment is administered at the national, state, and selected urban district levels. Results are reported on student achievement in grades 4, 8, and 12 at the national level and for grades 4 and 8 at the state and trial urban district level. The main NAEP assessment is based on a framework that can be updated periodically. The *Assessment and Item Specifications for the NAEP 2009 Mathematics Assessment* only reflects changes in grade 12 from 2005. Mathematics content objectives for grades 4 and 8 have not changed from 2005. Therefore, main NAEP trendlines from the early 1990's can continue at 4th and 8th grades for the 2009 assessment. Taken together, the NAEP assessments provide a rich, broad, and deep picture of student mathematics achievement in the United States.

WHAT IS AN ASSESSMENT SPECIFICATIONS DOCUMENT?

This document is a companion to the NAEP *Mathematics Framework for 2009.* The framework lays out the basic design of the assessment by describing the mathematics content that should be tested and the types of assessment questions that should be included. It also describes how the various design factors should be balanced across the assessment. The assessment and item specifications give more detail about development of the items and conditions for the 2009 NAEP mathematics assessment. It contains much of the same information that is in the framework about the mathematics content and other dimensions of the assessment, but adds further detail. The intended audience for the specifications is test developers and item writers.

THE NEED FOR A NEW FRAMEWORK AND SPECIFICATIONS AT GRADE 12

For several years, the National Assessment Governing Board has been focusing special attention on ways to improve the assessment of 12th graders by the National Assessment of Educational Progress. The goal for this 12th grade initiative is to enable reporting on how well prepared 12th grade students are for post-secondary education and training. To accomplish this goal the content of the assessments, as described in the 2005 mathematics framework, was analyzed and revisions considered. The challenge was to find the essential mathematics that can form the foundation for these post-secondary paths. This must include use of quantitative tools, broad competence in mathematical reasoning, mathematics required for postsecondary courses, and the ability to integrate and apply mathematics to diverse problem-solving contexts. Analysis of the 2005 framework revealed that some revisions would be necessary to meet this challenge.

THE FRAMEWORK AND SPECIFICATIONS DEVELOPMENT PROCESS

To implement this change at the 12th grade, the Governing Board contracted with Achieve, Inc. to examine NAEP's mathematics assessment framework in relation to benchmarks set by the American Diploma Project. An Achieve panel of mathematicians, mathematics educators, and policymakers proposed increasing the scope and rigor of 12th grade NAEP. New assessment objectives were developed by Achieve, and then a panel of mathematicians and mathematics educators (including classroom teachers) reviewed and revised the objectives and matched them against the current set of objectives for grades 4 and 8. The panel conducted focus groups with the Association of State Supervisors of Mathematics and survey reviews with various NAEP constituents, using repeated rounds of reviews. The final set of grade 12 objectives was approved by the Governing Board in August 2006.

CHANGES FROM THE 2005 FRAMEWORK AND SPECIFICATIONS

The chart below compares the 2009 to the 2005 mathematics framework and specifications:

Mathematics content	• Objectives for grades 4 and 8 remain the same • New subtopic of "mathematical reasoning" at grades 4, 8, and 12 • Distribution of items for each content area at all grades remains the same • New objectives for grade 12
Mathematical complexity	New clarifications and new examples to describe the levels of mathematical complexity
Calculator policy	Remains the same
Item formats	Remains the same
Tools and manipulatives	Remains the same

CONCLUSION AND PREVIEW OF DOCUMENT

The assessment and item specifications for the NAEP 2009 mathematics assessment follow. The bullets below summarize each chapter:

- **Mathematics content**
  Chapter Two contains descriptions of the five major content areas of mathematics (Number Properties and Operations, Measurement, Geometry, Data Analysis, Statistics, and Probability, and Algebra), as well as the specific objectives for grades 4, 8 and 12 that will be assessed. Further specifications are added to some objectives to clarify the intent for item writers.
- **Mathematical complexity**

Each NAEP mathematics test item is designed to measure a specific level of thinking, called the mathematical complexity of the item. Chapter Three describes the three levels and offers examples of each.

- **Item development**
  Chapter Four describes considerations for good item writing, with multiple examples of how each characteristic of an item might be met. The chapter also contains a description of the item tryout and review process.
- **Design of the assessment**
  Each form of the NAEP mathematics assessment must be balanced according to a number of different factors, including content, level of complexity, and format. In Chapter Five the guidelines for balancing each factor are described. The chapter also addresses other issues of design, such as sampling, use of calculators, tools and manipulatives, and accessibility for all students.

A valuable resource for learning more about NAEP can be found on the Internet at http://nces.ed.gov/nationsreportcard/. This site contains reports describing results of recent assessments, as well as a searchable tool for viewing released items. The items can be searched by many different features, such as grade level and content area. Information about the items includes student performance and any applicable scoring rubrics. NAEP released items that are used as examples in this document are marked with the designation that matches the item name in the NAEP Sample Question Tool, found on the website.

CHAPTER TWO

ITEM SPECIFICATIONS BY GRADE LEVEL AND CONTENT AREA

This chapter presents an overview of the content areas, a description of the matrix format, and a detailed description of each content area followed by the specific objectives of the mathematics framework. In addition there are guidelines provided for item writers intended to help clarify content area specifications as well as individual objectives.

CONTENT AREAS

Although the names of the content areas in previous NAEP frameworks, as well as some of the topics in those areas, may have changed somewhat from one assessment to the next, there has remained a consistent focus toward collecting information on student performance in five key areas. The framework for the 2009 Mathematics Assessment is anchored in these same five broad areas of mathematical content:

- **Number Properties and Operations** (including computation and the understanding of number concepts)
- **Measurement** (including use of instruments, application of processes, and concepts of area and volume)
- **Geometry** (including spatial reasoning and applying geometric properties)
- **Data Analysis, Statistics, and Probability** (including graphical displays, and statistics)
- **Algebra** (including representations and relationships)

These divisions are not intended to separate mathematics into discrete elements. Rather, they are intended to provide a helpful classification scheme that describes the full spectrum of mathematical content assessed by NAEP. Classifying items into one primary content area is not always clear cut, but doing so brings us closer to the goal of ensuring that important mathematical concepts and skills are assessed in a balanced way.

At grade 12, the five content areas are collapsed into four, with geometry and measurement combined into one for assessment development purposes. This reflects the fact that the majority of measurement topics suitable for twelfth-grade students are geometrical in nature. Separating these two areas of mathematics at grade 12 becomes forced and unnecessary.

It is important to note that there are certain aspects of mathematics that occur in all the content areas. The best example of this is computation. Computation is the skill of performing operations on numbers. It should not be confused with the content area of NAEP called Number Properties and Operations, which encompasses a wide range of concepts about our numeration system. Certainly the area of Number Properties and Operations includes a variety of computational skills, ranging from operations with whole numbers to work with decimals and fractions and finally real numbers. But computation is also critical in Measurement and Geometry, such as in calculating the perimeter of a rectangle, estimating the height of a building, or finding the hypotenuse of a right triangle. Data analysis often involves computation, such as calculating a mean or the range of a set of data. Probability often entails work with rational numbers. Solving algebraic equations usually involves numerical computation as well. Computation, therefore, is a foundational skill in every content area. While the main NAEP assessment is not designed to report a separate score for computation, results from the long-term NAEP assessment can provide insight into students' computational abilities.

As described in Chapter One, one of the changes made from the 2005 framework is the addition of a subtopic for mathematical reasoning that appears in Number, Geometry, Data Analysis, Statistics and Probability, and Algebra. At grades 4 and 8, no new objectives were written, but some of the objectives from the 2005 framework were moved into this new subtopic area. This reflects a new emphasis on the importance of mathematical reasoning across each of the content areas.

2009 NAEP Mathematics Objectives Organization

The specifications matrix is organized by the five NAEP content areas: Number Properties and Operations; Measurement, Geometry; Data Analysis, Statistics and Probability; and Algebra. Though such an organization brings with it the danger of fragmentation, the intent is the test items will, in many cases, cross some boundaries of these content areas, although an item will emphasize a primary content area.

The specifications matrix depicts the particular objectives appropriate for assessment under each subtopic. Within Number, for example, and the subtopic of Number Sense, specific objectives are listed for assessment at grade 4, grade 8, and grade 12. The same topic at different grade levels depicts a developmental sequence for that concept or skill. An empty cell in the matrix is used to convey the fact that a particular objective is not appropriate for assessment at that grade level. Guidelines for item writers (in italics) are included when needed to clarify the scope or measurement intent of the objectives.

In order to fully understand the objectives and their intent, please note the following:

- The objectives included in the matrix describe what is to be assessed on the 2009 NAEP. They should not be interpreted as a complete description of mathematics that should be taught at these grade levels.
- Some of the 12th grade objectives are marked with a "*". This denotes objectives that describe mathematics content that is beyond that typically taught in a standard 3-year course of study (the equivalent of one year of geometry and two years of algebra). Therefore, these objectives will be selected less often than the others for inclusion on the assessments.
- While all test items will be assigned a primary classification, some test items could potentially fall into more than one objective or more than one content area.
- When the guidelines use "include," this means that items can include these features, not that most items should include these features.

- When the word “or” is used in an objective, it should be understood as inclusive; that is, an item may assess one or more of the concepts included.

A valuable resource for learning more about NAEP can be found on the Internet at http://nces.ed.gov/nationsreportcard/. This site has reports describing results of recent assessments, as well as a searchable tool for viewing released items. The items can be searched by different features, such as grade level and content area. Information about the items includes student performance and any applicable scoring rubrics.

Mathematical Content Areas

NUMBER PROPERTIES AND OPERATIONS

Numbers are our main tools for describing the world quantitatively. As such they deserve a privileged place in the NAEP Mathematics framework. With whole numbers, we can count collections of discrete objects of any type. We can also use numbers to describe fractional parts, and even to describe continuous quantities such as length, area, volume, weight, and time, and more complicated derived quantities such as rates—speed, density, inflation, interest, and so forth. Thanks to Cartesian coordinates, we can use pairs of numbers to describe points in a plane or triples of numbers to label points in space. Numbers let us talk in a precise way about anything that can be counted, measured, or located in space.

Numbers are not simply labels for quantities. They form systems with their own internal structure. The arithmetic operations (addition and subtraction, multiplication and division) help us model basic real world operations. For example, joining two collections, or laying two lengths end to end, can be described by addition, while the concept of rate depends on division. Multiplication and division of whole numbers lead to the beginnings of number theory, including concepts of factorization, remainder, and prime number. Besides the arithmetic operations, the other basic structure of the real number system is its ordered nature. This allows comparison of numbers, as to which is greater or lesser. Ordering and comparing reflect our intuitions about the relative size of quantities and provide a basis for making sensible estimates.

The accessibility and usefulness of arithmetic is greatly enhanced by our efficient means for representing numbers — the Hindu-Arabic decimal place value system. In its full development, this remarkable system includes decimal fractions, which let us approximate any real number as closely as we wish. Decimal notation allows us to do arithmetic by means of simple, routine algorithms, and it also makes size comparisons and estimation easy. The decimal system achieves its efficiency through sophistication, as all the basic algebraic operations are implicitly used in writing decimal numbers. To represent ratios of two whole numbers exactly, we supplement decimal notation with fractions.

Comfort in dealing with numbers effectively is called *number sense*. It includes firm intuitions about what numbers tell us; an understanding of the ways to represent them symbolically (including facility with converting between different representations); ability to calculate, either exactly or approximately, and by several means—mentally, with paper and pencil, or with a calculator, as appropriate; and skill in estimation. Ability to deal with proportion, including percents, is another important part of number sense.

Number sense is a major expectation of the NAEP Mathematics Assessment. At 4th grade, students are expected to have a solid grasp of the whole numbers, as represented by the decimal system, and to have the beginnings of understanding fractions. By 8th grade, they should be comfortable with rational numbers, represented either as decimal fractions (including percents) or as common fractions. They should be able to use them to solve problems involving proportionality and rates. In middle school also, number should begin to coalesce with geometry, via the idea of the number line. This should be connected with ideas of approximation and the use of scientific notation. Eighth graders should also have some acquaintance with naturally occurring irrational numbers, such as square roots and pi. By twelfth grade, students should be comfortable dealing with all types of real numbers.

Number Properties and Operations

The word “expressions” refers to *numerical* expressions in this content area.
Italicized print in the matrix indicates item development guidelines.

1) Number sense		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Identify the place value and actual value of digits in whole numbers.	a) Use place value to model and describe integers and decimals.	
b) Represent numbers using models such as base 10 representations, number lines, and two-dimensional models.	b) Model or describe rational numbers or numerical relationships using number lines and diagrams.	
c) Compose or decompose whole quantities by place value (e.g., write whole numbers in expanded notation using place value: 342 = 300 + 40 + 2). *Items may use numbers through 999,999.*		

1) Number sense		
GRADE 4	**GRADE 8**	**GRADE 12**
d) Write or rename whole numbers (e.g., 10: 5 + 5, 12 – 2, 2 x 5).	d) Write or rename rational numbers.	d) Represent, interpret or compare expressions for real numbers, including expressions utilizing exponents and logarithms. *Negative and fractional exponents may be used.* *Expressions may include, for example, π, square root of 2, and numerical relationships using number lines, models or diagrams.*
e) Connect model, number word, or number using various models and representations for whole numbers, fractions, and decimals.	e) Recognize, translate between, or apply multiple representations of rational numbers (fractions, decimals, and percents) in meaningful contexts.	
	f) Express or interpret numbers using scientific notation from real-life contexts.	f) Represent or interpret expressions involving very large or very small numbers in scientific notation. *Negative exponents may be used.* *Items may include interpreting calculator or computer displays given in scientific notation.*
	g) Find or model absolute value or apply to problem situations.	g) Represent, interpret or compare expressions or problem situations involving absolute values.
	h) Order or compare rational numbers (fractions, decimals, percents, or integers) using various models and representations (e.g., number line).	
i) Order or compare whole numbers, decimals, or fractions.	i) Order or compare rational numbers including very large and small integers, and decimals and fractions close to zero.	i) Order or compare real numbers, including very large and very small real numbers.

2) Estimation		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Use benchmarks (well-known numbers used as meaningful points for comparison) for whole numbers, decimals, or fractions in contexts (e.g., 1/2 and .5 may be used as benchmarks for fractions and decimals between 0 and 1.00).	a) Establish or apply benchmarks for rational numbers and common irrational numbers (e.g.,) in contexts.	
b) Make estimates appropriate to a given situation with whole numbers, fractions, or decimals by: • Knowing when to estimate, • Selecting the appropriate type of estimate, including over- estimate, underestimate, and range of estimate, or • Selecting the appropriate method of estimation (e.g., rounding).	b) Make estimates appropriate to a given situation by: • Identifying when estimation is appropriate, • Determining the level of accuracy needed, • Selecting the appropriate method of estimation, or • Analyzing the effect of an estimation method on the accuracy of results.	b) Identify situations where estimation is appropriate, determine the needed degree of accuracy, and analyze the effect of the estimation method on the accuracy of results.
c) Verify solutions or determine the reasonableness of results in meaningful contexts.	c) Verify solutions or determine the reasonableness of results in a variety of situations including calculator and computer results.	c) Verify solutions or determine the reasonableness of results in a variety of situations. *Items may include using estimation and order of magnitude to determine the reasonableness of technology-aided computations and interpreting results in terms of the context.*
	d) Estimate square or cube roots of numbers less than 1,000 between two whole numbers. *Eighth grade items should be limited to numbers that are between more familiar perfect squares (1 through 144) or more familiar perfect cubes (1 through 125).*	d) Estimate square or cube roots of numbers less than 1,000 between two whole numbers.

3) Number operations		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Add and subtract: • Whole numbers, or • Fractions with like denominators, or • Decimals through hundredths. *Include items that are not placed in a context and require computation with common and decimal fractions, as well as items that use a context.*	a) Perform computations with rational numbers. *Include items that are not placed in a context and require computation with common and decimal fractions, as well as items that use a context.*	a) Find integral or simple fractional powers of real numbers. *Items also should include numbers expressed with negative exponents.* *For example, evaluate* $27^{1/3}$.
b) Multiply whole numbers: • No larger than two-digit by two-digit with paper and pencil computation, or • Larger numbers with use of calculator. *Money is an exception: multiplication problems involving money, with decimal places, can be included on calculator blocks.*		b) Perform arithmetic operations with real numbers, including common irrational numbers. *Include items that are not placed in a context and require computation with common and decimal fractions (decimals that can be written as a standard fraction) as well as items that use a context.* *Order of operations may be a component of items addressing this objective (i.e., the computation may require students knowing the appropriate order of operations).* *Items should not include absolute values (absolute value is addressed in A3c).*
c) Divide whole numbers: • Up to three-digits by one-digit with paper and pencil computation, or • Up to five-digits by two-digits with use of calculator. *Items written for calculator blocks should not have remainders.*		c) Perform arithmetic operations with expressions involving absolute value.

3) Number operations		
GRADE 4	**GRADE 8**	**GRADE 12**
d) Describe the effect of operations on size (whole numbers).	d) Describe the effect of multiplying and dividing by numbers including the effect of multiplying or dividing a rational number by: • Zero, or • A number less than zero, or • A number between zero and one, • One, or • A number greater than one.	d) Describe the effect of multiplying and dividing by numbers including the effect of multiplying or dividing a real number by: • Zero, or • A number less than zero, or • A number between zero and one, or • One, or • A number greater than one. *An item at eighth grade might ask, for example, about the effect of multiplying a fraction by a fraction less than one, or a fraction by a fraction greater than one. Twelfth grade items could include, for example, what is the effect of multiplying* $2\sqrt{3}$ *by ½?*
e) Interpret whole number operations and the relationships between them. *Interpret subtracting a number as the inverse operation to adding a number. Interpret dividing by a number as the inverse operation to multiplying a number.* *Interpret multiplication as repeated addition.*	e) Interpret rational number operations and the relationships between them. *Use the four operations, roots, and powers; additive inverses, multiplicative inverses.*	
f) Solve application problems involving numbers and operations. *Use the same limitations on computation as in 3a – 3d.*	f) Solve application problems involving rational numbers and operations using exact answers or estimates as appropriate.	f) Solve application problems involving numbers, including rational and common irrationals.

4) Ratios and proportional reasoning		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Use simple ratios to describe problem situations.	a) Use ratios to describe problem situations.	
	b) Use fractions to represent and express ratios and proportions.	

4) Ratios and proportional reasoning		
GRADE 4	**GRADE 8**	**GRADE 12**
	c) Use proportional reasoning to model and solve problems (including rates and scaling).	c) Use proportions to solve problems (including rates of change). *Items should not include scale drawings (scale drawings are included in B2f).*
	d) Solve problems involving percentages (including percent increase and decrease, interest rates, tax, discount, tips, or part/whole relationships).	d) Solve multi-step problems involving percentages, including compound percentages.

5) Properties of number and operations		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Identify odd and even numbers.	a) Describe odd and even integers and how they behave under different operations..	
b) Identify factors of whole numbers. *Limit numbers to 2 through 12.*	b) Recognize, find, or use factors, multiples, or prime factorization. *Lowest common multiple, greatest common factor, common multiple for reasonably small numbers.* *Without calculator, numbers to be less than 400. With calculator, numbers to be less than 1,000*	
	c) Recognize or use prime and composite numbers to solve problems.	c) Solve problems using factors, multiples, or prime factorization. *Items should include problems involving prime numbers.*
	d) Use divisibility or remainders in problem settings.	d) Use divisibility or remainders in problem settings.
e) Apply basic properties of operations. *Properties include order and grouping.*	e) Apply basic properties of operations. *Properties include commutative, associative, and distributive properties of addition and multiplication.*	e) Apply basic properties of operations, including conventions about the order of operations. *Properties include commutative, associative, and distributive properties of addition and multiplication.* *The emphasis should be on properties rather than computation.*

5) Properties of number and operations		
GRADE 4	**GRADE 8**	**GRADE 12**
		f) Recognize properties of the number system—whole numbers, integers, rational numbers, real numbers, and complex numbers—recognize how they are related to each other, and identify examples of each type of number. *Items can include questions about identifying irrational numbers. For example, which if the following is irrational: 0.333, 0.333…, 3.14,* $\sqrt{3}$ *?*

6) Mathematical reasoning using number		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Explain or justify a mathematical concept or relationship (e.g., explain why 15 is an odd number or why 7–3 is not the same as 3–7).	a) Explain or justify a mathematical concept or relationship (e.g., explain why 17 is prime).	a) Give a mathematical argument to establish the validity of a simple numerical property or relationship.
	b) Provide a mathematical argument to explain operations with two or more fractions.	b) * Analyze or interpret a proof by mathematical induction of a simple numerical relationship. *For example, for a proof that the sum of the first n odd numbers is n squared, students should be expected to recognize or complete, but not compose, the proof.*

MEASUREMENT

Measuring is the process by which numbers are assigned in order to describe the world quantitatively. This process involves selecting the attribute of the object or event to be measured, comparing this attribute to a unit, and reporting the number of units. For example, in measuring a child, we may select the attribute of height and the inch as the unit for the comparison. In comparing the height to the inch, we may find that the child is about 42 inches. If considering only the domain of whole numbers, we would report that the child is 42 inches tall. However, since height is a continuous attribute, we may consider the domain of rational numbers and report that the child is 41 3/16 inches tall (to the nearest sixteenth of the inch). Measurement also allows us to model positive and negative numbers as well as the irrational numbers.

This connection between measuring and number makes measuring a vital part of the school curriculum. Measurement models are often used when students are learning about number and operations. For example, area and volume models can help students understand multiplication and the properties of multiplication. Length models, especially the number line, can help students understand ordering and rounding numbers. Measurement also has a strong connection to other areas of school mathematics and to the other subjects in the school curriculum. Problems in algebra are often drawn from measurement situations. One can also consider measurement to be a function or a mapping of the attribute to a set of numbers. Much of school geometry focuses on the measurement aspect of geometric figures. Statistics also provides ways to measure and to compare sets of data. These are some of the ways that measurement is intertwined with the other four content areas.

In this NAEP mathematics framework, attributes such as capacity, weight/mass, time and temperature are included as well as the geometric attributes of length, area, and volume. Although many of these attributes are included in the grade 4 framework, the emphasis is on length, including perimeter, distance, and height. More emphasis is placed on area and angle in grade 8. By grade 12, volumes and rates constructed from other attributes, such as speed, are emphasized.

Units involved in items on the NAEP assessment include non-standard, customary, and metric units. At grade 4, common customary units such as inch, quart, pound, and hour and the common metric units such as centimeter, liter, and gram are emphasized. Grades 8 and 12 include the use of both square and cubic units for measuring area, surface area, and volume, degrees for measuring angles, and constructed units such as miles per hour. Converting from one unit in a system to another such as from minutes to hours is an important aspect of measurement included in problem situations. Understanding and using the many conversions available is an important skill. There are a limited number of common, everyday equivalencies that students are expected to know.

Items classified in this content area depend on some knowledge of measurement. For example, an item that asks the difference between a 3-inch and a 1 3/4 inch line segment is a number item, while an item comparing a 2-foot segment with an 8-inch line segment is a measurement item. In many secondary schools, measurement becomes an integral part of geometry; this is reflected in the proportion of items recommended for these two areas.

General Guidelines for Measurement

Any attribute, unit, instrument, conversion factor, or formula included in the list at a lower grade is also appropriate for the higher grade(s).

<u>Attributes</u>

The following attributes may be included in items:

- Grade 4 – Length, time, temperature, capacity, weight, and area with emphasis on length (length includes perimeter, height, and distance).
- Grade 8 – Angle and volume. Emphasis is on area. Attributes such as speed, measured in terms of the attributes of time and distance, are also appropriate.
- Grade 12 – The emphasis is on area (including surface area) and volume, but any attribute used in grade 8 is appropriate. Rates constructed from other attributes such as speed or flow rate are appropriate.

<u>Units</u>

- Grade 4 – Non-standard units, common customary units (inch, foot, mile, cup, quart, gallon, pound, hour, minute, day, year) and metric units (centimeter, millimeter, meter, liter, gram) for the allowed attributes at this grade level.
- Grade 8 – Square units and cubic units, degrees of angles, and constructed units such as miles per hour; metric units most commonly used for each of the attributes are appropriate also.
- Grade 12 – Same as grade 8.

<u>Instruments</u>

The following instruments are commonly found in curricula; variations based on the same principles could be used (e.g., graduated cup measures):

- Grade 4 – Ruler, clock, thermometer, graduated cylinder, balance scales, scales.
- Grade 8 – Protractor.
- Grade 12 – Same as grade 8.

<u>Conversions</u>

Items should be based on students' knowing specific equivalences, as follows:

- Grade 4 – Feet/inches, hours/minutes, and meters/centimeters; other simple conversions should be given, such as 2 pints = 1 quart.
- Grade 8 – Square and cubic unit conversions; students should also know all common time equivalences, all common metric equivalence. Otherwise, conversions should be based on provided equivalences.
- Grade 12 – Conversions from constructed units such as miles per hour to feet per minute.

<u>Formulas</u>

Grade 4 students are not expected to know any measurement formulas; however, they are expected to know how to find the perimeter and area of a rectangle. Both grade 8 and grade 12 students should know formulas for the area of a rectangle, triangle, and circle, the circumference of a circle, and the volume of a cylinder and rectangular solid. If other formulas are used, they should be given. (See General Guidelines for Geometry for more information about formulas for area, circumference, and volume.)

Measurement

Italicized print in the matrix indicates item development guidelines.

1) Measuring physical attributes		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Identify the attribute that is appropriate to measure in a given situation.		
b) Compare objects with respect to a given attribute, such as length, area, volume, time, or temperature.	b) Compare objects with respect to length, area, volume, angle measurement, weight, or mass.	b) Determine the effect of proportions and scaling on length, areas and volume.
c) Estimate the size of an object with respect to a given measurement attribute (e.g., length, perimeter, or area using a grid).	c) Estimate the size of an object with respect to a given measurement attribute (e.g., area).	c) Estimate or compare perimeters or areas of two-dimensional geometric figures.
		d) Solve problems of angle measure, including those involving triangles or other polygons or parallel lines cut by a transversal. *Students are expected to know that the sum of the interior angles of a triangle is 180^o and to know about angles formed by parallel lines cut by a transversal.*
e) Select or use appropriate measurement instruments such as ruler, meter stick, clock, thermometer, or other scaled instruments.	e) Select or use appropriate measurement instrument to determine or create a given length, area, volume, angle, weight, or mass.	
f) Solve problems involving perimeter of plane figures.	f) Solve mathematical or real-world problems involving perimeter or area of plane figures such as triangles, rectangles, circles, or composite figures.	f) Solve problems involving perimeter or area of plane figures such as polygons, circles, or composite figures.
g) Solve problems involving area of squares and rectangles.		
	h) Solve problems involving volume or surface area of rectangular solids, cylinders, prisms, or composite shapes.	h) Solve problems by determining, estimating, or comparing volumes or surface areas of three-dimensional figures.
	i) Solve problems involving rates such as speed or population density.	i) Solve problems involving rates such as speed, density, population density, or flow rates.

2) Systems of measurement		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Select or use appropriate type of unit for the attribute being measured such as length, time, or temperature.	a) Select or use appropriate type of unit for the attribute being measured such as length, area, angle, time, or volume.	a) Recognize that geometric measurements (length, area, perimeter, and volume) depend on the choice of a unit, and apply such units in expressions, equations, and problem solutions.
b) Solve problems involving conversions within the same measurement system such as conversions involving inches and feet or hours and minutes. *Items may include conversions such as pints to quarts, given the conversion information (e.g., 2 pints = 1 quart).*	b) Solve problems involving conversions within the same measurement system such as conversions involving square inches and square feet.	b) Solve problems involving conversions within or between measurement systems, given the relationship between the units. *Items may include cubic units and rates such as miles per hour to feet per second.* *Refer to the list of conversions student should know on page 15 of the specifications document.*
	c) Estimate the measure of an object in one system given the measure of that object in another system and the approximate conversion factor. For example: • Distance conversion: 1 kilometer is approximately 5/8 of a mile. • Money conversion: US dollars to Canadian dollars. • Temperature conversion: Fahrenheit to Celsius	
d) Determine appropriate size of unit of measurement in problem situation involving such attributes as length, time, capacity, or weight.	d) Determine appropriate size of unit of measurement in problem situation involving such attributes as length, area, or volume.	d) Understand that numerical values associated with measurements of physical quantities are approximate, are subject to variation, and must be assigned units of measurement.
e) Determine situations in which a highly accurate measurement is important.	e) Determine appropriate accuracy of measurement in problem situations (e.g., the accuracy of each of several lengths needed to obtain a specified accuracy of a total length) and find the measure to that degree of accuracy.	e) Determine appropriate accuracy of measurement in problem situations (e.g., the accuracy of measurement of the dimensions to obtain a specified accuracy of area) and find the measure to that degree of accuracy. *For example, if you measured the size of a rectangle to the nearest inch and found it to be 3" by 5", what is the range that the area of the rectangle could actually be?*

2) Systems of measurement		
GRADE 4	**GRADE 8**	**GRADE 12**
	f) Construct or solve problems (e.g., floor area of a room) involving scale drawings.	f) Construct or solve problems involving scale drawings. *The scale drawing does not have to be given.* *For example, determine the number of rolls of insulation needed for insulating a house.*

3) Measurement in Triangles		
GRADE 4	**GRADE 8**	**GRADE 12**
	a) Solve problems involving indirect measurement such as finding the height of a building by comparing its shadow with the height and shadow of a known object.	a) Solve problems involving indirect measurement. *For example, find the height of a building by finding the distance to the base of the building and the angle of elevation to the top.*
		b) Solve problems using the fact that trigonometric ratios (sine, cosine, and tangent) stay constant in similar triangles. *For example, explain why the tangents of corresponding angles of two similar triangles are equal.*
		c) Use the definitions of sine, cosine, and tangent as ratios of sides in a right triangle to solve problems about length of sides and measure of angles. *Students should know the definitions of sine, cosine, and tangent.* *Students should know the side relationships for triangles with angle measurements of 45-45-90 and 30-60-90.*
		d) Interpret and use the identity $\sin^2\theta + \cos^2\theta = 1$ for angles θ between 0° and 90°; recognize this identity as a special representation of the Pythagorean theorem. *Students should know that* $\sin^2\theta + \cos^2\theta = 1$.

3) Measurement in Triangles		
GRADE 4	**GRADE 8**	**GRADE 12**
		e) * Determine the radian measure of an angle and explain how radian measurement is related to a circle of radius 1. *Angles should be restricted to π/6, π/4, π/3, π/2 and angles in other quadrants with these same referent angles.*
		f) * Use trigonometric formulas such as addition and double angle formulas. *Students should be provided with trigonometric formulas (law of cosines, double-angle formula, etc.)* *For example, explain why the following is true or false: sin 20º = 2sin10º..*
		g) * Use the law of cosines and the law of sines to find unknown sides and angles of a triangle. *Students should be provided with trigonometric formulas (law of cosines, double-angle formula, etc.).*

GEOMETRY

Geometry began as a practical collection of rules for calculating lengths, areas, and volumes of common shapes. In classical times, the Greeks turned it into a subject for reasoning and proof, and Euclid organized their discoveries into a coherent collection of results, all deduced using logic from a small number of special assumptions, called postulates. Euclid's *Elements* stood as a pinnacle of human intellectual achievement for over 2000 years.

The 19th century saw a new flowering of geometric thought, going beyond Euclid, and leading to the idea that geometry is the study of the possible structures of space. This had its most striking application in Einstein's theories of relativity, which described the behavior of light, and also of gravity, in terms of a four-dimensional geometry, which combines the usual three dimensions of space with time as an additional dimension.

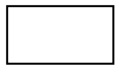
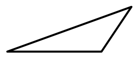
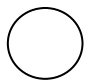
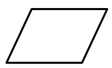
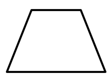
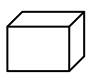
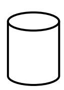
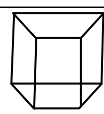
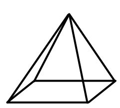
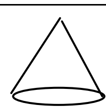
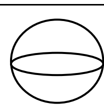
A major insight of the 19th century is that geometry is intimately related to ideas of symmetry and transformation. The symmetry of familiar shapes under simple transformations—that our bodies look more or less the same if reflected across the middle, or that a square looks the same if rotated by 90^{o}—is a matter of everyday experience. Many of the standard terms for triangles (scalene, isosceles, equilateral) and quadrilaterals (parallelogram, rectangle, rhombus, square) refer to symmetry properties. Also, the behavior of figures under changes of scale is an aspect of symmetry with myriad practical consequences. At a deeper level, the fundamental ideas of geometry itself (for example, congruence) depend on transformation and invariance. In the 20th century, symmetry ideas were seen to underlie much of physics also, not only Einstein's relativity theories, but atomic physics and solid state physics (the field that produced computer chips).

School geometry roughly mirrors the historical development through Greek times, with some modern additions, most notably symmetry and transformations. By grade 4, students are expected to be familiar with a library of simple figures and their attributes, both in the plane (lines, circles, triangles, rectangles, and squares), and in space (cubes, spheres, and cylinders). In middle school, understanding of these shapes deepens, with study of cross-sections of solids, and the beginnings of an analytical understanding of properties of plane figures, especially parallelism, perpendicularity, and angle relations in polygons. Right angles and the Pythagorean Theorem are introduced, and geometry becomes more and more mixed with measurement. The basis for analytic geometry is laid by study of the number line. In high school, attention is given to Euclid's legacy and the power of rigorous thinking. Students are expected to make, test, and validate conjectures. Via analytic geometry, the key areas of geometry and algebra are merged into a powerful tool that provides a basis for calculus and the applications of mathematics that helped create the modern technological world in which we live.

Symmetry is an increasingly important component of geometry. Elementary students are expected to be familiar with the basic types of symmetry transformations of plane figures, including flips (reflection across lines), turns (rotations around points), and slides (translations). In middle school, this knowledge becomes more systematic and analytical, with each type of transformation being distinguished from other types by their qualitative effects. For example, a rigid motion of the plane that leaves at least two points fixed (but not all points) must be a reflection in a line. In high school, students are expected to be able to represent transformations algebraically. Some may also gain insight into their systematic structure, such as the classification of rigid motions of the plane as reflections, rotations, or translations, and what happens when two or more isometries are performed in succession (composition).

General Guidelines for Geometry

Students are expected to know the basic formulas for shapes as indicated in the chart below.

Shape: Formulas for Area and Circumference		Grade 4	Grade 8	Grade 12
Rectangle		(Find area and perimeter, but not use the formula.)		
Triangle				
Circle				
Parallelogram				
Trapezoid				
Figure: Formulas for Volume and Surface Area		**Grade 4**	**Grade 8**	**Grade 12**
Rectangular Prism				
Right Circular Cylinder				
General Prisms				
Square Pyramid				
Right Circular Cone				
Sphere				

Key:
- **Not tested**
- **Students are expected to know the formula**
- **Formula should be provided**

Geometry

Italicized print in the matrix indicates item development guidelines.

1) Dimension and shape		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Explore properties of paths between points.	a) Draw or describe a path of shortest length between points to solve problems in context. *For example, find the shortest path between two locations when there are buildings in between the locations.*	
b) Identify or describe (informally) real-world objects using simple plane figures (e.g., triangles, rectangles, squares, and circles) and simple solid figures (e.g., cubes, spheres, and cylinders). *For example, identify rectangles in a picture of a room.*	b) Identify a geometric object given a written description of its properties.	
c) Identify or draw angles and other geometric figures in the plane.	c) Identify, define, or describe geometric shapes in the plane and in three-dimensional space given a visual representation. *Three-dimensional shapes should be simple ones such as a sphere, tetrahedron, prism, pyramid.*	c) Give precise mathematical descriptions or definitions of geometric shapes in the plane and in three-dimensional space. *Include the full set of Platonic solids (e.g., cube, regular tetrahedron).*
	d) Draw or sketch from a written description polygons, circles, or semicircles.	d) Draw or sketch from a written description plane figures and planar images of three-dimensional figures. *Figures can include isosceles triangles, regular polygons, polyhedra, spheres, and hemispheres.*
	e) Represent or describe a three-dimensional situation in a two-dimensional drawing from different views. *Figures should be simple, standard ones such as a cube, regular tetrahedron, rectangular solid.*	e) Use two-dimensional representations of three-dimensional objects to visualize and solve problems.
f) Describe attributes of two- and three-dimensional shapes.	f) Demonstrate an understanding about the two- and three-dimensional shapes in our world through identifying, drawing, modeling, building, or taking apart.	f) Analyze properties of three-dimensional figures including spheres and hemispheres.

2) Transformation of shapes and preservation of properties		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Identify whether a figure is symmetrical, or draw lines of symmetry. *Items should address line symmetry only.* *Items can involve a single or more than one line of symmetry.*	a) Identify lines of symmetry in plane figures or recognize and classify types of symmetries of plane figures. *Items may include point, line, and rotational symmetry.*	a) Recognize or identify types of symmetries (e.g., point, line, rotational, self-congruence) of two- and three-dimensional figures.
		b) Give or recognize the precise mathematical relationship (e.g., congruence, similarity, orientation) between a figure and its image under a transformation. *Transformations can include reflections, rotations, translations, and dilations.*
c) Identify the images resulting from flips (reflections), slides (translations), or turns (rotations).	c) Recognize or informally describe the effect of a transformation on two-dimensional geometric shapes (reflections across lines of symmetry, rotations, translations, magnifications, and contractions).	c) Perform or describe the effect of a single transformation on two- and three-dimensional geometric shapes (reflections across lines of symmetry, rotations, translations, and dilations).
d) Recognize which attributes (such as shape and area) change or don't change when plane figures are cut up or rearranged.	d) Predict results of combining, subdividing, and changing shapes of plane figures and solids (e.g., paper folding, tiling, and cutting up and rearranging pieces).	d) Identify transformations, combinations or subdivisions of shapes that preserve the area of two-dimensional figures or the volume of three-dimensional figures. *Items can include the comparison of the areas of two different shapes.*
e) Match or draw congruent figures in a given collection.	e) Justify relationships of congruence and similarity, and apply these relationships using scaling and proportional reasoning. *Two-dimensional figures only.*	e) Justify relationships of congruence and similarity, and apply these relationships using scaling and proportional reasoning. *Justifications should be less formal than the proofs called for in Geometry 5e, such as giving reasons why figures are congruent or similar.* *The scaling and proportional reasoning may be applied to both two- and three-dimensional figures.*

2) Transformation of shapes and preservation of properties		
GRADE 4	**GRADE 8**	**GRADE 12**
	f) For similar figures, identify and use the relationships of conservation of angle and of proportionality of side length and perimeter. *Include triangles, with an emphasis on right triangles and quadrilaterals.*	
		g) Perform or describe the effects of successive transformations. *For example, describe the result of a series of three reflections over three parallel lines.*

3) Relationships between geometric figures		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Analyze or describe patterns of geometric figures by increasing number of sides, changing size or orientation (e.g., polygons with more and more sides).		
b) Assemble simple plane shapes to construct a given shape.	b) Apply geometric properties and relationships in solving simple problems in two and three dimensions. *Properties include geometric similarity, congruence, angle sum.* *Include angle relationships and transversal properties of quadrilateral angles.* *Eligible figures include parallel and perpendicular lines, triangles, circles, cylinders, and cones.*	b) Apply geometric properties and relationships to solve problems in two and three dimensions. *Problems can involve multiple steps.* *Figures can include parallel and perpendicular lines, triangles (including 45-45-90 and 30-60-90 triangles), cylinders, cones, prisms, and pyramids.* *The emphasis should be on solving problems.*
c) Recognize two-dimensional faces of three-dimensional shapes.	c) Represent problem situations with simple geometric models to solve mathematical or real world problems.	c) Represent problem situations with geometric models to solve mathematical or real-world problems. *Grade 12 items will be more complex that grade 8 items. For example, grade 12 items might involve more figures, or more properties.* *The emphasis should be on representations or models.*

3) Relationships between geometric figures		
GRADE 4	**GRADE 8**	**GRADE 12**
	d) Use the Pythagorean theorem to solve problems. *Students are expected to recall the Pythagorean theorem.*	d) Use the Pythagorean theorem to solve problems in two- or three-dimensional situations. *Students will not be provided the Pythagorean Theorem, but will be expected to know and apply it.*
		e) Recall and interpret definitions and basic properties of congruent and similar triangles, circles, quadrilaterals, polygons, parallel, perpendicular and intersecting lines, and associated angle relationships. *The emphasis should be on definitions or defining properties*
f) Describe and compare properties of simple and compound figures composed of triangles, squares, and rectangles.	f) Describe or analyze simple properties of, or relationships between, triangles, quadrilaterals, and other polygonal plane figures. *For example, given a pair of parallel lines cut by a transversal, identify the angles that have the same measure.*	f) Analyze properties or relationships of triangles, quadrilaterals, and other polygonal plane figures. *Items can include rhombi, parallelograms, trapezoids, being sure to avoid situations in which the definition of a trapezoid must be assumed.* *The emphasis should be on analyzing properties.*
	g) Describe or analyze properties and relationships of parallel or intersecting lines.	g) Analyze properties and relationships of parallel, perpendicular, or intersecting lines, including the angle relationships that arise in these cases. *The emphasis should be on analyzing properties.*
		h) Analyze properties of circles and the intersection of circles and lines (inscribed angles, central angles, tangents, secants, and chords). *Items can ask, for example, about angles inscribed in a semicircle, or the relationship between tangents, secants, chords and radii.*

4) Position, direction, and coordinate geometry		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Describe relative positions of points and lines using the geometric ideas of parallelism or perpendicularity.	a) Describe relative positions of points and lines using the geometric ideas of midpoint, points on common line through a common point, parallelism, or perpendicularity.	a) Solve problems involving the coordinate plane such as the distance between two points, the midpoint of a segment, or slopes of perpendicular or parallel lines. *Items can include finding the slope of a line given two points.*
	b) Describe the intersection of two or more geometric figures in the plane (e.g., intersection of a circle and a line).	b) Describe the intersections of lines in the plane and in space, intersections of a line and a plane, or of two planes in space.
	c) Visualize or describe the cross section of a solid. *Cross-sections should be of standard, familiar solids such as a sphere, cylinder, and rectangular solid.*	c) Describe or identify conic sections and other cross sections of solids. *Cross-sections should be of standard, familiar solids such as a cone, sphere or cylinder, and some Platonic solids (e.g., cube, regular tetrahedron).*
d) Construct geometric figures with vertices at points on a coordinate grid. *Emphasis is on geometric properties.*	d) Represent geometric figures using rectangular coordinates on a plane.	d) Represent two-dimensional figures algebraically using coordinates and/or equations.
		e) * Use vectors to represent velocity and direction; multiply a vector by a scalar and add vectors both algebraically and graphically.
		f) Find an equation of a circle given its center and radius and, given an equation of a circle, find its center and radius. *Students are expected to know the equation of a circle.* *Items may require the student to derive the center or radius.*

4) Position, direction, and coordinate geometry		
GRADE 4	**GRADE 8**	**GRADE 12**
		g) *Graph ellipses and hyperbolas whose axes are parallel to the coordinate axes and demonstrate understanding of the relationship between their standard algebraic form and their graphical characteristics. *The formulas for ellipses and hyperbolas will be provided in standard form.* *Items may require knowledge of general characteristics of these functions (e.g., drawing a graph), but should not require knowledge of technical characteristics (e.g., equations of asymptotes or foci).*
		h) * Represent situations and solve problems involving polar coordinates.

5) Mathematical reasoning in Geometry		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Distinguish which objects in a collection satisfy a given geometric definition and explain choices.	a) Make and test a geometric conjecture about regular polygons.	a) Make, test, and validate geometric conjectures using a variety of methods including deductive reasoning and counterexamples.
		b) Determine the role of hypotheses, logical implications, and conclusion, in proofs of geometric theorems.
		c) Analyze or explain a geometric argument by contradiction *For example, explain why, in a scalene triangle the bisector of an angle cannot be perpendicular to the opposite side.*
		d) Analyze or explain a geometric proof of the Pythagorean theorem. *For example, complete missing steps in the proof based on similar triangles.*

5) Mathematical reasoning in Geometry		
GRADE 4	**GRADE 8**	**GRADE 12**
		e) Prove basic theorems about congruent and similar triangles and circles. *Items should allow for a variety of representations of the proof (e.g., flow diagrams, paragraph, two-column proofs).* *Examples include standard SAS, SSS, or ASA congruence proofs with corresponding parts.*

DATA ANALYSIS, STATISTICS, AND PROBABILITY

Data analysis covers the entire process of collecting, organizing, summarizing, and interpreting data. This is the heart of the discipline called statistics and is in evidence whenever quantitative information is used in determining a course of action. To emphasize the spirit of statistical thinking, data analysis should begin with a question to be answered—not with the data. Data should be collected only with a specific question (or questions) in mind and only after a plan (usually called a design) for collecting data relevant to the question is thought out. Beginning at an early age, students should grasp the fundamental principle that looking for questions in an existing data set is far different from the scientific method of collecting data to verify or refute a well-posed question. A pattern can be found in almost any data set if one looks hard enough, but a pattern discovered in this way is often meaningless, especially from the point of view of statistical inference.

In the context of data analysis, or statistics, probability can be thought of as the study of potential patterns in outcomes that have not yet been observed. We say that the probability of a balanced coin coming up heads when flipped is one half because we believe that about half of the flips would turn out to be heads if we flipped the coin many times. Under random sampling, patterns for outcomes of designed studies can be anticipated and used as the basis for making decisions. If the coin actually turned up heads 80% of the time, we would suspect that it was not balanced. The whole probability distribution of all possible outcomes is important in most statistics problems because the key to decision-making is to decide whether or not a particular observed outcome is unusual (located in a tail of the probability distribution) or not. For example, four as a grade point average is unusually high among most groups of students, four as the pound weight of a baby is unusually low, and four as the number of runs scored in a baseball game is not unusual in either direction.

By grade 4, students should be expected to apply their understanding of number and quantity to pose questions that can be answered by collecting appropriate data. They should be expected to organize data in a table or a plot, and summarize the essential features of center, spread, and shape both verbally and with simple summary statistics. Simple comparisons can be made between two related data sets, but more formal inference based on randomness should come later. The basic concept of chance and statistical reasoning can be built into meaningful contexts, though, such as “If I draw two names from among those of the students in the room, am I likely to get two girls?” Such problems can be addressed through simulation.

Building on the same definition of data analysis and the same principles of describing distributions of data through center, spread, and shape, grade 8 students will be expected to use a wider variety of organizing and summarizing techniques. They can also begin to analyze statistical claims through designed surveys and experiments that involve randomization, with simulation being the main tool for making simple statistical inferences. They will begin to use more formal terminology related to probability and data analysis.

Students in grade 12 will be expected to use a wide variety of statistical techniques for all phases of the data analysis process, including a more formal understanding of statistical inference (but still with simulation as the main inferential analysis tool). In addition to comparing univariate data sets, students at this level should be able to recognize and describe possible associations between two variables by looking at two-way tables for categorical variables or scatterplots for measurement variables. Association between variables is related to the concepts of independence and dependence, and an understanding of these ideas requires knowledge of conditional probability. These students should be able to use statistical models (linear and non-linear equations) to describe possible associations between measurement variables and be familiar with techniques for fitting models to data.

General Guidelines for Data Analysis, Statistics, and Probability

<u>Data Representation</u>

- Items should include interpretation of uncommon representations of data such as those found in newspapers and magazines.
- Bar and line graphs should increase in complexity (e.g., through using more complex scales) from grade to grade.
- Descriptions of data sets at grade 4 may be informal.

The following representations of data are indicated for each grade level. Objectives in which only a subset of these representations is applicable are indicated in the parenthesis associated with the objective.

Grade 4
Pictographs, bar graphs, circle graphs, line graphs, line plots, tables, and tallies

Grade 8
Histograms, line graphs, scatterplots, box plots, bar graphs, circle graphs, stem and leaf plots, frequency distributions, and tables

Grade 12
Histograms, line graphs, scatterplots, box plots, bar graphs, circle graphs, stem and leaf plots, frequency distributions, and tables, including two-way tables

Data Analysis, Statistics, and Probability

Italicized print in the matrix indicates item development guidelines.

1) Data representation		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Read or interpret a single set of data.	a) Read or interpret data, including interpolating or extrapolating from data.	a) Read or interpret graphical or tabular representations of data.
b) For a given set of data, complete a graph (limits of time make it difficult to construct graphs completely).	b) For a given set of data, complete a graph and then solve a problem using the data in the graph (histograms, line graphs, scatterplots, circle graphs, and bar graphs).	b) For a given set of data, complete a graph and solve a problem using the data in the graph (histograms, scatterplots, line graphs)

1) Data representation		
GRADE 4	**GRADE 8**	**GRADE 12**
c) Solve problems by estimating and computing within a single set of data.	c) Solve problems by estimating and computing with data from a single set or across sets of data.	c) Solve problems involving univariate or bivariate data. *Items can require using multiple sets of data. For example, construct and compare three box plots based on given data sets.*
	d) Given a graph or a set of data, determine whether information is represented effectively and appropriately (histograms, line graphs, scatterplots, circle graphs, and bar graphs).	d) Given a graphical or tabular representation of a set of data, determine whether information is represented effectively and appropriately.
	e) Compare and contrast the effectiveness of different representations of the same data. *For example, compare the effects of scale change on various graphs*	e) Compare and contrast different graphical representations of univariate and bivariate data. *For example, compare the effects of scale change on various graphs.*
		f) Organize and display data in a spreadsheet in order to recognize patterns and solve problems. *Until graphing calculators are required or until the assessment is administered via computer, students will not be asked to manipulate spreadsheets. However, students can be asked to recognize patterns displayed in a spreadsheet and use the data to solve problems.*

2) Characteristics of data sets		
GRADE 4	**GRADE 8**	**GRADE 12**
	a) Calculate, use, or interpret mean, median, mode, or range.	a) Calculate, interpret, or use summary statistics for distributions of data including measures of typical value (mean, median), position (quartiles, percentiles), and spread (range, interquartile range, variance, standard deviation).

2) Characteristics of data sets		
GRADE 4	**GRADE 8**	**GRADE 12**
b) Given a set of data or a graph, describe the distribution of the data using median, range, or mode.	b) Describe how mean, median, mode, range, or interquartile ranges relate to the shape of the distribution.	b) Recognize how linear transformations of one-variable data affect mean, median, mode, range, interquartile range, and standard deviation. *For example, what is the effect on the mean of adding a constant to each data point?*
	c) Identify outliers and determine their effect on mean, median, mode, or range.	c) Determine the effect of outliers on mean, median, mode, range, interquartile range, or standard deviation.
d) Compare two sets of related data.	d) Using appropriate statistical measures, compare two or more data sets describing the same characteristic for two different populations or subsets of the same population. *Items can use mean, median, mode, and range.*	d) Compare data sets using summary statistics (mean, median, mode, range, interquartile range, or standard deviation) describing the same characteristic for two different populations or subsets of the same population.
	e) Visually choose the line that best fits given a scatterplot and informally explain the meaning of the line. Use the line to make predictions.	e) Approximate a trend line if a linear pattern is apparent in a scatterplot or use a graphing calculator to determine a least-squares regression line, and use the line or equation to make a prediction. *Until graphing calculators are required or until the assessment is administered via computer, students will not be asked to use a graphing calculator to construct a least-squares regression line.*
		f) Recognize that the correlation coefficient is a number from –1 to +1 that measures the strength of the linear relationship between two variables; visually estimate the correlation coefficient (e.g., positive or negative, closer to 0, .5, or 1.0) of a scatterplot. *Items may ask students to construct scatterplots for correlations of 0, ±0.5, or ±1.0.*

2) Characteristics of data sets		
GRADE 4	**GRADE 8**	**GRADE 12**
		g) Know and interpret the key characteristics of a normal distribution such as shape, center (mean), and spread (standard deviation).

3) Experiments and samples		
GRADE 4	**GRADE 8**	**GRADE 12**
	a) Given a sample, identify possible sources of bias in sampling.	a) Identify possible sources of bias in sample surveys, and describe how such bias can be controlled and reduced.
	b) Distinguish between a random and nonrandom sample.	b) Recognize and describe a method to select a simple random sample.
		c) * Draw inferences from samples, such as estimates of proportions in a population, estimates of population means, or decisions about differences in means for two "treatments".
	d) Evaluate the design of an experiment.	d) Identify or evaluate the characteristics of a good survey or of a well-designed experiment.
		e) * Recognize the differences in design and in conclusions between randomized experiments and observational studies. *Items can ask, for example, about different sources of bias between the two types of studies, how randomness is considered in each type, or how changes in variables are treated.*

4) Probability		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Use informal probabilistic thinking to describe chance events (i.e., likely and unlikely, certain and impossible).	a) Analyze a situation that involves probability of an independent event.	a) Recognize whether two events are independent or dependent.

4) Probability		
GRADE 4	**GRADE 8**	**GRADE 12**
b) Determine a simple probability from a context that includes a picture.	b) Determine the theoretical probability of simple and compound events in familiar contexts. *Items should use familiar contexts such as number cubes, flipping coins, spinners.*	b) Determine the theoretical probability of simple and compound events in familiar or unfamiliar contexts. *Compound events included in items should be independent.*
	c) Estimate the probability of simple and compound events through experimentation or simulation. *Items should use familiar contexts such as number cubes, flipping coins, spinners.*	c) Given the results of an experiment or simulation, estimate the probability of simple or compound events in familiar or unfamiliar contexts. *For example, explain how the relative frequency of occurrences of a specified outcome of an event is not the same as its probability but can be used to estimate the probability of the outcome (for example: if Anita flipped a coin 10 times and got 7 heads, the probability of a head is not 0.7).*
	d) Use theoretical probability to evaluate or predict experimental outcomes.	d) Use theoretical probability to evaluate or predict experimental outcomes. *Items at the 12th grade should be more complex than those at the 8th grade. For example, they would involve more events.*
e) List all possible outcomes of a given situation or event.	e) Determine the sample space for a given situation.	e) Determine the number of ways an event can occur using tree diagrams, formulas for combinations and permutations, or other counting techniques. *Students should demonstrate understanding of how to generate sample spaces.*
	f) Use a sample space to determine the probability of the possible outcomes of an event.	
g) Represent the probability of a given outcome using a picture or other graphic.	g) Represent probability of a given outcome using fractions, decimals, and percents.	

4) Probability		
GRADE 4	**GRADE 8**	**GRADE 12**
	h) Determine the probability of independent and dependent events. (Dependent events should be limited to a small sample size.)	h) Determine the probability of independent and dependent events. *Compound events included in items should be dependent.*
		i) Determine conditional probability using two-way tables.
	j) Interpret probabilities within a given context.	j) Interpret and apply probability concepts to practical situations.
		k) *Use the binomial theorem to solve problems. *The binomial theorem will be given to students.* *For example, given a binomial problem situation with the probability of an event being 0.1, determine the probability of that event occurring 3 out of 11 times.*

5) Mathematical Reasoning With Data		
GRADE 4	**GRADE 8**	**GRADE 12**
		a) Identify misleading uses of data in real-world settings and critique different ways of presenting and using information.
		b) Distinguish relevant from irrelevant information, identify missing information, and either find what is needed or make appropriate approximations.
		c)* Recognize, use, and distinguish between the processes of mathematical (deterministic) and statistical modeling. *For example, distinguish between calculating a line of best fit for a scatterplot and finding the equation of a line through 2 points.*
		d) Recognize when arguments based on data confuse correlation with causation.

5) Mathematical Reasoning With Data		
GRADE 4	**GRADE 8**	**GRADE 12**
		e) * Recognize and explain the potential errors caused by extrapolating from data. *For example, explain the danger of using a line of best fit to make predictions for values well beyond the range of the given data.*

ALGEBRA

Algebra was pioneered in the Middle Ages by mathematicians in the Middle East and Asia as a method of solving equations easily and efficiently by manipulation of symbols, rather than by the earlier geometric methods of the Greeks. The two approaches were eventually united in the analytic geometry of René Descartes. Modern symbolic notation, developed in the Renaissance, greatly enhanced the power of the algebraic method, and from the 17th century forward, algebra in turn promoted advances in all branches of mathematics and science.

The widening use of algebra led to study of its formal structure. Out of this were gradually distilled the "rules of algebra," a compact summary of the principles behind algebraic manipulation. A parallel line of thought produced a simple but flexible concept of function and also led to the development of set theory as a comprehensive background for mathematics. When it is taken liberally to include these ideas, algebra reaches from the foundations of mathematics to the frontiers of current research.

These two aspects of algebra, a powerful representational tool and a vehicle for comprehensive concepts such as function, form the basis for the expectations throughout the grades. By grade 4, students are expected to be able to recognize and extend simple numeric patterns as one foundation for a later understanding of function. They can begin to understand the meaning of equality and some of its properties, as well as the idea of an unknown quantity, as a precursor to the concept of variable.

As students move into middle school, the ideas of function and variable become more important. Representation of functions as patterns, via tables, verbal descriptions, symbolic descriptions, and graphs can combine to promote a flexible grasp of the idea of function. Linear functions receive special attention. They connect to the ideas of proportionality and rate, forming a bridge that will eventually link arithmetic to calculus. Symbolic manipulation in the relatively simple context of linear equations is reinforced by other means of finding solutions, including graphing by hand or with calculators.

In high school, students should become comfortable in manipulating and interpreting more complex expressions. The rules of algebra should come to be appreciated as a basis for reasoning. Non-linear functions, especially quadratic functions, and also power and exponential functions, are introduced to solve real-world problems. Students should become accomplished at translating verbal descriptions of problem situations into symbolic form. Expressions involving several variables, systems of linear equations, and the solutions to inequalities are encountered by grade 12.

General Guidelines for Algebra

Overall, items at grade 4 emphasize informal algebra. For example, there is an emphasis on "completing number sentences" instead of "solving equations." At grade 8, items cover some formal algebra, but the expectation is that less formal algebra content will be included. For example, solution of higher degree polynomial equations or systems of linear or non-linear equations is not expected at the 8th grade level, but these topics are included at the 12th grade level.

At grade 12, the types of functions eligible for use in all items are linear, quadratic, rational, exponential, and trigonometric. Rational functions are limited to the following set: those with a constant or linear numerator and a linear or quadratic denominator. Rational expressions are limited in the same way. Trigonometric functions are limited to sine, cosine and tangent.

Logarithmic functions can be used in * items only.

Algebra

Italicized print in the matrix indicates item development guidelines.

1) Patterns, relations, and functions		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Recognize, describe, or extend numerical patterns.	a) Recognize, describe, or extend numerical and geometric patterns using tables, graphs, words, or symbols. *Pattern types can include rational numbers, powers, simple recursive patterns, regular polygons, three-dimensional shapes. The complexity of patterns should be higher than grade 4.*	a) Recognize, describe, or extend numerical patterns, including arithmetic and geometric progressions. *Items can use same pattern types as grade 8 but can be more complex. Pattern types can include rational numbers, powers, simple recursive patterns, linear patterns, quadratic patterns, exponential patterns, regular polygons, 3-dimentional shapes.* *Items must be carefully worded so as not to give the mistaken idea that there is only one correct solution.* *The nature of the pattern should be clearly defined in the problem.* *Items can use patterns with multiple solutions if the student is asked to explain his/her answer.* *Responses can include verbal descriptions or equations.*

1) Patterns, relations, and functions		
GRADE 4	**GRADE 8**	**GRADE 12**
b) Given a pattern or sequence, construct or explain a rule that can generate the terms of the pattern or sequence.	b) Generalize a pattern appearing in a numerical sequence or table or graph using words or symbols.	b) Express linear and exponential functions in recursive and explicit form given a table, verbal description, or some terms of a sequence. *Items can require the student to provide the explicit form of a function, given a recursive form.* *Students may be asked to write an equation of a line given a table of points.*
c) Given a description, extend or find a missing term in a pattern or sequence.	c) Analyze or create patterns, sequences, or linear functions given a rule.	
d) Create a different representation of a pattern or sequence given a verbal description.		
e) Recognize or describe a relationship in which quantities change proportionally. *Items should include relating input to output.*	e) Identify functions as linear or nonlinear or contrast distinguishing properties of functions from tables, graphs, or equations. *Items can include properties \ such as whether a given function is represented by a line or curve, slopes and intercepts.*	e) Identify or analyze distinguishing properties of linear, quadratic, rational, exponential, or *trigonometric functions from tables, graphs, or equations. *Items can include properties such as rate of change, intercepts, periodicity or symmetry.*
	f) Interpret the meaning of slope or intercepts in linear functions.	
		g) Determine whether a relation, given in verbal, symbolic, tabular, or graphical form, is a function.
		h) Recognize and analyze the general forms of linear, quadratic, rational, exponential, or *trigonometric functions *Items can include examining parameters and their effect on the graph of linear and quadratic functions. For example, in $y = ax + b$, recognize the roles of a and b.*

1) Patterns, relations, and functions		
GRADE 4	**GRADE 8**	**GRADE 12**
		i) Determine the domain and range of functions given in various forms and contexts. *Eligible functions are linear, quadratic, inverse proportionality (y=k/x), exponential, and trigonometric functions.* *Items can include characteristics of domain and range in problem contexts, or in functions such as f(x)=|x-3|.*
		j) * Given a function, determine its inverse if it exists, and explain the contextual meaning of the inverse for a given situation. *For example, if f(t) = the population in year t, what does* $f^{-1}(3000) = 1965$ *mean?*

2) Algebraic representations		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Translate between the different forms of representations (symbolic, numerical, verbal, or pictorial) of whole number relationships (such as from a written description to an equation or from a function table to a written description).	a) Translate between different representations of linear expressions using symbols, graphs, tables, diagrams, or written descriptions.	a) Create and translate between different representations of algebraic expressions, equations, and inequalities (e.g., linear, quadratic, exponential, or *trigonometric) using symbols, graphs, tables, diagrams, or written descriptions. *Items can include those that require students to construct graphs.* *The stimulus can include symbols, graphs, tables, diagrams, or written descriptions.* *Items should require either translating between two different forms of representation or, given one form of representation, creating a different form of representation.*

2) Algebraic representations		
GRADE 4	**GRADE 8**	**GRADE 12**
	b) Analyze or interpret linear relationships expressed in symbols, graphs, tables, diagrams, or written descriptions. *Items can include identification of strengths and weaknesses of different representations.*	b) Analyze or interpret relationships expressed in symbols, graphs, tables, diagrams (including Venn diagrams), or written descriptions and evaluate the relative advantages or disadvantages of different representations to answer specific questions.
c) Graph or interpret points with whole number or letter coordinates on grids or in the first quadrant of the coordinate plane. *The emphasis is on use of coordinates.* *Items can include maps.*	c) Graph or interpret points that are represented by ordered pairs of numbers on a rectangular coordinate system. *Items should include rational number coordinates only.*	
	d) Solve problems involving coordinate pairs on the rectangular coordinate system. *Items can include finding areas of simple geometric figures.*	d) Perform or interpret transformations on the graphs of linear, quadratic, exponential, and *trigonometric functions. *The graph of the function should be given in the stem.* *For example, give the vertex of the new parabola if $y = x^2$ is translated up 3 units and right 5 units, and then reflected over the line $y = x$.*
		e) Make inferences or predictions using an algebraic model of a situation.
	f) Identify or represent functional relationships in meaningful contexts including proportional, linear, and common nonlinear (e.g., compound interest, bacterial growth) in tables, graphs, words, or symbols. *Non-linear functions should have whole number powers***.**	f) Given a real-world situation, determine if a linear, quadratic, rational, exponential, logarithmic, or *trigonometric function fits the situation. *Items can include, for example, Celsius/Fahrenheit conversions, projectile motion, half-life, bacterial growth, Richter scale for earthquakes, or logarithmic scales in graphs.*
		g) Solve problems involving exponential growth and decay.

2) Algebraic representations		
GRADE 4	**GRADE 8**	**GRADE 12**
		h) *Analyze properties of exponential, logarithmic, and rational functions. *Items can include, for example, points of discontinuity or asymptotes (vertical and horizontal).* *Items should not require determining domains and ranges (determining domains and ranges is in Algebra 1i.).*

3) Variables, expressions, and operations		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Use letters and symbols to represent an unknown quantity in a simple mathematical expression.		
b) Express simple mathematical relationships using number sentences.	b) Write algebraic expressions, equations, or inequalities to represent a situation. *Use linear and simple quadratic functions in a contextual situation.*	b) Write algebraic expressions, equations, or inequalities to represent a situation. *Items can include finding or writing the equation of a line given the slope and a point or given two points.* *Twelfth grade items can include terms of higher degree, while 8th grade items should be restricted to expressions, equations, or inequalities with first degree terms.*
	c) Perform basic operations, using appropriate tools, on linear algebraic expressions (including grouping and order of multiple operations involving basic operations, exponents, roots, simplifying, and expanding).	c) Perform basic operations, using appropriate tools, on algebraic expressions including polynomial and rational expressions.
		d) Write equivalent forms of algebraic expressions, equations, or inequalities to represent and explain mathematical relationships. *Items should address equivalent forms within one type of representation, not translating between different representations.*

3) Variables, expressions, and operations		
GRADE 4	**GRADE 8**	**GRADE 12**
		e) Evaluate algebraic expressions, including polynomials and rational expressions.
		f) Use function notation to evaluate a function at a specified point in its domain and combine functions by addition, subtraction, multiplication, division, and composition.
		g) * Determine the sum of finite and infinite arithmetic and geometric series. *Students will be provided with formulas for the sum of a finite or infinite series.* *For example, find the total distance traveled by a ball dropped from 20 feet that bounces to 75% of its height.*
		h) Use basic properties of exponents and *logarithms to solve problems.

4) Equations and inequalities		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Find the value of the unknown in a whole number sentence. *Use equalities and simple inequalities (e.g., $x + a < b$).*	a) Solve linear equations or inequalities (e.g., $ax + b = c$ or $ax + b = cx + d$ or $ax + b > c$). *Use rational coefficients if the item is in a non-calculator block.*	a) Solve linear, rational or quadratic equations or inequalities, including those involving absolute value. *Items should not use complex roots.* *Items can include real number coefficients.* *Students are expected to know the quadratic formula.*
	b) Interpret "=" as an equivalence between two expressions and use this interpretation to solve problems.	

4) Equations and inequalities		
GRADE 4	**GRADE 8**	**GRADE 12**
	c) Analyze situations or solve problems using linear equations and inequalities with rational coefficients symbolically or graphically (e.g., ax + b = c or ax + b = cx + d).	c) Analyze situations, develop mathematical models, or solve problems using linear, quadratic, exponential, or logarithmic equations or inequalities symbolically or graphically. *Items should not use complex roots.* *Items can include real number coefficients.*
	d) Interpret relationships between symbolic linear expressions and graphs of lines by identifying and computing slope and intercepts (e.g., know in y = ax + b, that a is the rate of change and b is the vertical intercept of the graph).	d) Solve (symbolically or graphically) a system of equations or inequalities and recognize the relationship between the analytical solution and graphical solution. *Systems of equations should be limited to two linear equations or one linear and one quadratic equation.* *Items can assess compound inequalities.*
	e) Use and evaluate common formulas [e.g., relationship between a circle's circumference and diameter (C = pi d), distance and time under constant speed]. *Formulas must be from contextual situation.*	e) Solve problems involving special formulas such as: $A = P(I + r)^t$, $A = Pe^{rt}$]. *Special formulas are given to the student. All variables in special formulas are defined for the student.*
		f) Solve an equation or formula involving several variables for one variable in terms of the others.
		g) Solve quadratic equations with complex roots. *Students are expected to know the quadratic formula.*

5) Mathematical Reasoning in Algebra		
GRADE 4	**GRADE 8**	**GRADE 12**
a) Verify a conclusion using algebraic properties. *For example, if Sam is 3 years older than Ned, 20 years from now Sam will still be 3 years older than Ned.*	a) Make, validate, and justify conclusions and generalizations about linear relationships. *Identify the limits of generalizations in concrete situations, modeled algebraically using equations or graphs and algebraic properties.* *Items should require inductive and deductive reasoning.*	a) Use algebraic properties to develop a valid mathematical argument. *Properties include properties of equality and properties of operations. For example, explain why division by zero is undefined.*
		b) Determine the role of hypotheses, logical implications, and conclusions in algebraic argument. *For example, understand that either of the following statements cannot be reversed: $y = x - 1$ implies $y^2 = (x-1)^2$ or $f(x) = 0$ implies $g(x)*f(x) = 0$.*
		c) Explain the use of relational conjunctions (and, or) in algebraic arguments. *For example, for what values of x and y is $(x-1)(y+1) > 0$? Explain why.*

CHAPTER THREE

MATHEMATICAL COMPLEXITY OF ITEMS

Each NAEP item assesses an objective that can be associated with a content area of mathematics, such as number or geometry. The item also makes certain demands on students' thinking. These demands determine the mathematical complexity of the item, which is the second dimension of the mathematics framework. The three levels of mathematical complexity in NAEP assessment are low, moderate, and high.

The demands on thinking that an item expects—what it asks the student to recall, understand, reason about, and do—assume that students are familiar with the mathematics of the task. For example, a task with low complexity might ask students simply to state the formula to find the distance between two points. Those students who had never learned anything about distance formula would not be successful on the task even though the demands were low. Items are developed for administration at a given grade level on the basis of the framework, and complexity of those items is independent of the particular curriculum a student has experienced.

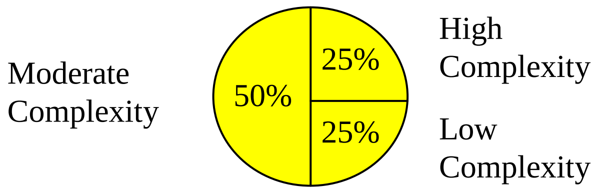
Mathematical complexity deals with *what* the students are asked to do in a task. It does not take into account *how* they might undertake it. In the distance formula task, for instance, some students who had studied the formula might simply reproduce it from memory. Others, however, who could not recall the exact formula, might end up deriving it from the Pythagorean theorem, engaging in a different kind of thinking than the task presupposed.

The categories—low complexity, moderate complexity, and high complexity—form an ordered description of the demands an item may make on a student. Items at the low level of complexity, for example, may ask a student to recall a property. At the moderate level, an item may ask the student to make a connection between two properties; at the high level, an item may ask a student to analyze the assumptions made in a mathematical model. This is an example of the distinctions made in item complexity to provide balance in the item pool. The ordering is not intended to imply that mathematics is learned or should be taught in such an ordered way. Using the levels of complexity to describe that dimension of each item allows for a balance of mathematical thinking in the design of the assessment.

The mathematical complexity of an item is not directly related to its format (multiple-choice, short constructed response, or extended constructed response). Items requiring that the student generate a response tend to make somewhat heavier demands on students than items requiring a choice among alternatives, but that is not always the case. Any type of item can deal with mathematics of greater or less depth and sophistication. There are multiple-choice items that assess complex mathematics, and constructed response items can be crafted to assess routine mathematical ideas.

The remainder of this chapter gives brief descriptions of each level of complexity as well as examples from previous NAEP assessments to illustrate each level. A brief rationale is included to explain why an item is so classified. All example items found in this chapter can also be found in the companion document, the NAEP *Mathematics Framework for 2009*.

Items in the NAEP assessment should be balanced according to levels of complexity, as described in more detail in Chapter Five. The ideal balance should be as follows:



Percent of Testing Time at Each Level of Complexity

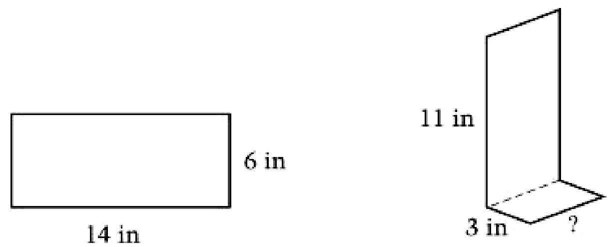
LOW COMPLEXITY

Low-complexity items expect students to recall or recognize concepts or procedures specified in the framework. Items typically specify what the student is to do, which is often to carry out some procedure that can be performed mechanically. It is not left to the student to come up with an original method or to demonstrate a line of reasoning. The following examples are items that have been classified at the low complexity level.

EXAMPLE 1: LOW COMPLEXITY Grade 4 Number Properties and Operations: Number sense	Source: 1996 NAEP 4M9 #1 Percent correct: 50% No calculator
How many fourths make a whole? Answer: __________	Correct Answer: 4
Rationale: This item is of low complexity since it explicitly asks students to recognize an example of a concept (four-fourths make a whole).	

EXAMPLE 2: LOW COMPLEXITY
Grade 4
Geometry: Transformations of shapes

Source: 2005 NAEP 4M12 #12
Percent correct: 54%
No calculator



A piece of metal in the shape of a rectangle was folded as shown above. In the figure o the right, the "?" symbol represents what length?

A. 3 inches
B. 6 inches
C. 8 inches
D. 11 inches

Correct Answer: B

Rationale: Although this is a visualization task, it is of low complexity since it requires only a straightforward recognition of the change in the figure. Students in the 4th grade are expected to be familiar with sums such as 11 + 3, so this does not increase the complexity level for these students.

EXAMPLE 3: LOW COMPLEXITY
Grade 8
Algebra: Algebraic representations

Source: 2005 NAEP 8M12 #17
Percent correct: 54%
No calculator

x	0	1	2	3	10
y	-1	2	5	8	29

Which of the following equations represents the relationship between *x* and *y* shown in the table above?

A. $y = x^2 + 1$
B. $y = x + 1$
C. $y = 3x – 1$
D. $y = x^2 - 3$
E. $y = 3 x^2 – 1$

Correct Answer: C

Rationale: This item would be at the moderate level if it were written as follows, “Write the equation that represents the relationship between *x* and *y*.” In generating the equation students would first have to decide if the relationship was linear.

EXAMPLE 4: LOW COMPLEXITY Grade 8 Data Analysis, Statistics, and Probability: Characteristics of data sets	Source: 2005 NAEP 8M12 #6 Percent correct: 51% No calculator

The prices of gasoline in a certain region are \$1.41, \$1.36, \$1.57, and \$1.45 per gallon. What is the median price per gallon for gasoline in this region?

A. \$1.41
B. \$1.43
C. \$1.44
D. \$1.45
E. \$1.47

Correct Answer: B

Rationale: Students do not have to decide what to do, but to recall the concept of a median and the procedure for handling a set of data with an even number of entries.

EXAMPLE 5: LOW COMPLEXITY Grade 12 Algebra: Equations and inequalities	Source: 2005 NAEP B3M3#12 Percent correct: 31% No calculator

$$x + 2y = 17$$
$$x - 2y = 3$$

The graphs of the two equations shown above intersect at the point (x, y).
What is the value of x at the point of intersection?

A. 3 ½
B. 5
C. 7
D. 10
E. 20

Correct Answer: D

Rationale: This item is of low complexity since it involves a procedure that should be carried out mechanically by grade 12.

EXAMPLE 6: LOW COMPLEXITY
Grade 12
Algebra: Variables, expressions, and operations

Source: 2005 NAEP B3M3 #16
Percent correct: 26%
No calculator

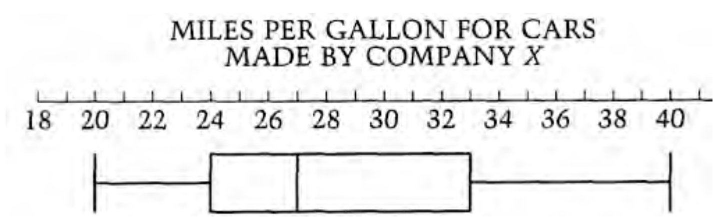
If $f(x) = x^2 + x$ and $g(x) = 2x + 7$, what is the expression for $f(g(x))$?

Correct Answer: $4x^2 + 30x + 56$

Rationale: Although the content of the task could be considered advanced, it involves recognizing the notation for composition of two functions and carrying out a procedure.

EXAMPLE 7: LOW COMPLEXITY
Grade 12
Data Analysis, Statistics, and Probability: Data representation

Source: 2005 NAEP B3M3 #11
Percent correct: 39%
No calculator



According to the box-and-whisker plot above, three-fourths of the cars made by Company X got fewer than how many miles per gallon.

A. 20
B. 24
C. 27
D. 33
E. 40

Correct Answer: D

Rationale: This item is of low complexity since it requires reading a graph and recalling that the four sections of the box-and-whisker plot are quartiles each represent one-fourth of the data.

MODERATE COMPLEXITY

Items in the moderate-complexity category involve more flexibility of thinking and choice among alternatives than do those in the low-complexity category. The student is expected to decide what to do, and how to do it, bringing together concepts and processes from various domains. For example, the student may be asked to represent a situation in more than one way, to draw a geometric figure that satisfies multiple conditions, or to solve a problem involving multiple unspecified operations. Students might be asked to show or explain their work, but would not be expected to justify it mathematically. The following examples are items that have been classified at the moderate complexity level.

EXAMPLE 8: MODERATE COMPLEXITY Grade 4 Equations and inequalities	Source: 2005 NAEP 4M4 #12 Percent correct: 34% (Full credit) Algebra: 22% (Partial credit) No calculator, tiles provided

Questions 11-14 [these questions included this item] refer to the number tiles or the paper strip. Please remove the 10 number tiles and the paper strip from your packet and put them on your desk.

Jan entered four numbers less than 10 on his calculator. He forgot what his second and fourth numbers were. This is what he remembered doing.

$$8 + \boxed{} - 7 + \boxed{} = 10$$

List a pair of numbers that could have been the second and fourth numbers.
(You may use the number tiles to help you.)

________ , ________

List a different pair that could have been the second and fourth numbers.

________ , ________

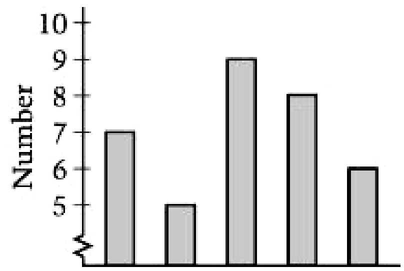
Correct Answer: Any two of these combinations.

(0,9)	(9,0)
(1,8)	(8,1)
(2,7)	(7,2)
(3,6)	(6,3)
(4,5)	(5,4)

Rationale: This item is of moderate complexity because students have to decide what to do and how to do it. It requires some flexibility in thinking since students at this grade level are not expected to have a routine method to determine two missing numbers and they also have to find two different solutions.

EXAMPLE 9: MODERATE COMPLEXITY
Grade 4
Data Analysis, Statistics, and Probability: Data representation

Source: 2005 NAEP 4M12 #11
Percent correct: 52%
No calculator



Jim made the graph above. Which of these could be the title for the graph?

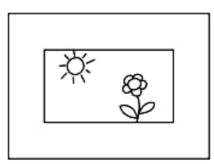
A. Number of students who walked to school on Monday through Friday
B. Number of dogs in five states
C. Number of bottles collected by three students
D. Number of students in ten clubs

Correct Answer: A

Rationale: Students must analyze the graph and the choices for a title and eliminate choices because of knowledge of dogs and clubs and the structure of the graph (5 sets of data) in order to choose an appropriate title for the graph.

EXAMPLE 10: MODERATE COMPLEXITY
Grade 8
Measurement: Measuring physical attributes

Source: 2005 NAEP 8M3 #3
Percent correct: 44% (Full credit)
13%(Partial credit)
No calculator, ruler provided



The figure above shows a picture and its frame.

In the space below, draw a rectangular picture 2 inches by 3 inches and draw a 1-inch-wide frame around it.

Rationale: Students must plan their drawing, whether to begin with the inside or outside rectangle and how the other rectangle is related to the one chosen. Often creating a drawing that satisfies several conditions is more complex than describing a given figure.

EXAMPLE 11: MODERATE COMPLEXITY	Source: 2005 NAEP 8M3 #10
Grade 8	Percent correct: 34%
Algebra: Patterns, relations, and functions	No calculator

In the equation $y = 4x$, if the value of x is increased by 2, what is the effect on the value of y?

A. It is 8 more than the original amount.
B. It is 6 more than the original amount.
C. It is 2 more than the original amount.
D. It is 16 times the original amount.
E. It is 8 times the original amount

Correct Answer: A

Rationale: This item is of moderate complexity because it involves more flexibility and a choice of alternative ways to approach the problem than a low complexity level in which more clearly states what to be done. At grade 8, students have not learned a procedure for answering this type of question.

EXAMPLE 12: MODERATE COMPLEXITY	Source: 2005 NAEP 8M3 #14
Grade 8	Percent correct: 28%
Geometry: Relationships in geometric figures	No calculator

A certain 4-sided figure has the following properties.

- Only one pair of opposite sides are parallel.
- Only one pair of opposite sides are equal in length.
- The parallel sides are not equal in length.

Which of the following must be true about the sides that are equal in length?

A. They are perpendicular to each other.
B. They are each perpendicular to an adjacent side.
C. They are equal in length to one of the other two sides.
D. They are not equal in length to either of the other two sides.
E. They are not parallel.

Correct Answer: E

Rationale: This item is of moderate complexity since it requires some visualization and reasoning, but no mathematical justification for the answer chosen.

EXAMPLE 13: MODERATE COMPLEXITY Grade 12 Number Properties and Operations: Number operations	Source: 2005 NAEP B3M3 Percent Correct: 22% No calculator

The remainder when a number *n* is divided by 7 is 2. Which of the following is the remainder when 2*n* + 1 is divided by 7?

A. 1
B. 2
C. 3
D. 4
E. 5

Correct Answer: E

Rationale: Although the problem could be approached algebraically ($n = 7m + 2$, for some whole number m, and $2n + 1= 2(7m + 2) + 1$ or $14 m + 5$, so the remainder is 5), students can solve the problem by using a value for n that satisfies the condition that it has a remainder of 2 when divided by 7. If the students were asked to justify their solution algebraically, then this would be an item of high complexity.

EXAMPLE 14: MODERATE COMPLEXITY Grade 12 Measurement: Measuring physical attributes	Source: 2005 NAEP B3M12 #15 Percent Correct: 41% Calculator available

A cat lies crouched on level ground 50 feet away from the base of a tree. The cat can see a bird's nest directly above the base of the tree. The angle of elevation from the cat to the bird's nest is 40°. To the nearest foot, how far above the base of the tree is the bird's nest?

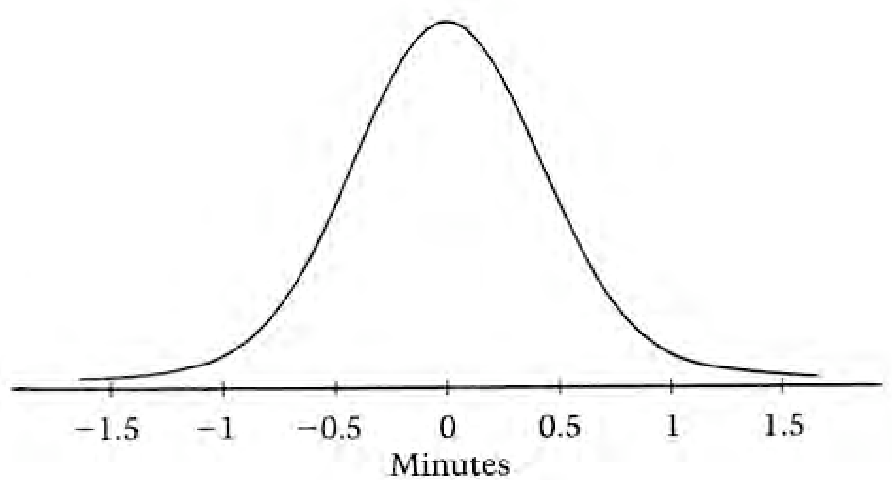
A. 32
B. 38
C. 42
D. 60
E. 65

Correct Answer: C

Rationale: Students must draw or visualize the situation, recall the appropriate trigonometric function, and use a calculator to determine the value of that function.

EXAMPLE 15: MODERATE COMPLEXITY Grade 12 Data Analysis, Statistics, and Probability: Characteristics of data sets	Source: 2005 NAEP B3M12 #16 Percent Correct: 12% Calculator available

A clock manufacturer has found that the amount of time their clocks gain or lose per week is normally distributed with a mean of 0 minutes and a standard deviation of 0.5 minute, as shown below.



In a random sample of 1,500 of their clocks, which of the following is closest to the expected number of clocks that would gain or lose more than 1 minute per week?

A. 15
B. 30
C. 50
D. 70
E. 90

Correct Answer: D

Rationale: Students must recall information about the normal curve (that the region between the mean ± 2 standard deviations contains 95% of the data), and apply that information to solve the problem.

HIGH COMPLEXITY

High-complexity items make heavy demands on students, who are expected to use reasoning, planning, analysis, judgment, and creative thought. Students may be expected to justify mathematical statements or construct a mathematical argument. Items might require students to generalize from specific examples. Items at this level take more time than those at other levels due to the demands of the task, but not due to the number of parts or steps. In the example items at the moderate level, several suggestions were made in the rationale that would make the items at a high level of complexity. The following examples are items that have been classified at the high complexity level.

EXAMPLE 16: HIGH COMPLEXITY
Grade 4
Algebra: Patterns, relations and functions

Source: 2003 NAEP 4M7 #20
Percent Correct: 3% (Extended)
6% (Satisfactory) 13% (Partial)
27 (Minimal)
Calculator available

The table below shows how the chirping of a cricket is related to the temperature outside. For example, a cricket chirps 144 times each minute when the temperature is 76°.

Number of Chirps Per Minute	Temperature
144	76°
152	78°
160	80°
168	82°
176	84°

The table below shows how the chirping of a cricket is related to the temperature outside. For example, a cricket chirps 144 times each minute when the temperature is 76°.

What would be the number of chirps per minute when the temperature outside is 90° if this pattern stays the same?

Answer: ____________________________

Explain how you figured out your answer.

Correct Answer: 200

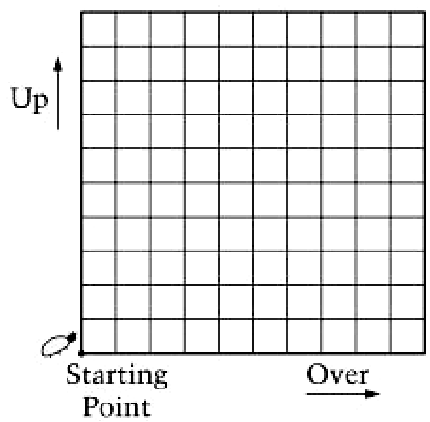
Rationale: To receive full credit for this item, students must give the correct number of chirps and explain that for every 2 degrees rise in the temperature, the number of chirps increases by 8. The item requires creative thought for students at this grade as well as planning a solution strategy. Additionally, it requires a written justification of their answer, more than just showing work.

EXAMPLE 17 : HIGH COMPLEXITY

Algebra: Patterns, relations, and functions

Source: 2005 NAEP 8M4 #11 Grade 8
Percent correct: 12% (Full credit)
24% (Partial credit)
No calculator

If the grid in Question 10 [the previous question] were large enough and the beetle continued to move in the same pattern [over 2 and up 1], would the point that is 75 blocks up and 100 blocks over from the starting point be on the beetle's path?



Ⓐ Yes Ⓑ No

Give a reason for your answer.

Rationale: Students must justify their yes or no answer by using the concept of slope showing that moving over 2 and up 1 repeatedly would result in the beetle being at a point 100 blocks over and 50 blocks up. This requires analysis of the situation as well as a mathematical explanation of the thinking. Since it is not realistic to extend the grid, students are expected to generalize about the ratio.

EXAMPLE 18: HIGH COMPLEXITY
Grade 12
Number Properties and Operations: Number Sense

Modified Item
No calculator

Which of the following is false for all values of *x* if *x* is any real number?

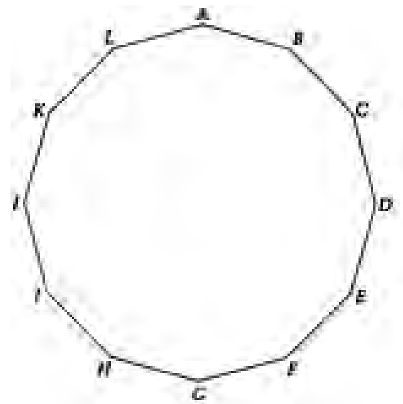
A. $x < x^2 < x^3$
B. $x^3 < x < x^2$
C. $x^2 < x < x^3$
D. $x < x^3 < x^{2'}$
E. $x^3 < x^2 < x$

Correct Answer: C

Rationale: This multiple-choice item requires planning, deciding what strategy to use, and reasoning about which statement is always false.

EXAMPLE 19: HIGH COMPLEXITY
Grade 12
Geometry: Mathematical reasoning

MODIFIED NAEP item
No protractor



Each of the 12 sides of the regular figure above has the same length.

1. Which of the following angles has a measure of 90°?

   A. Angle *ABI*
   B. Angle *ACG*
   C. Angle *ADF*
   D. Angle *ADI*
   E. Angle *AEH*

Correct Answer: B

2. Prove that no angle formed by joining three vertices of the figure could have a measure of 50 degrees.

Modification: This item (2005 NAEP B3M3 #1) has been modified to illustrate a high complexity item. The original item allowed the use of protractor and did not ask for a proof.

Rationale: There are several ways to approach part 1 of this problem so students must decide what method to bring to it. Part 2 raises the complexity to high since it requires students to present a mathematical argument requiring creative thought and the bringing together of information about circle arcs and inscribed angles. They could argue that no angle can be 50° because all angles must be multiples of 15°.

Additional examples of items and their classifications can be found in *2009 NAEP Mathematics Assessment and Item Specifications* as well as on the National Center for Education Statistics' website: http://nces.ed.gov/nationsreportcard/itmrls/. All the released items from recent mathematics assessments can be accessed from this site. The complexity classification is available only for items beginning with the 2005 assessment since this was the first year that the framework specified this dimension.

CHAPTER FOUR

ITEM DEVELOPMENT

This chapter discusses specifications that apply to all grade levels assessed by the NAEP in mathematics. Chapter Two: Item Specifications by Content Area contains specifications by content area for each grade level. Chapter Three: Mathematical Complexity of Items describes the three levels of complexity of NAEP mathematics items. The guidelines in these three chapters are focused on translating the intent of the content framework into development of items used on the assessment.

These guidelines highlight only some of the critical considerations in item development and concentrates on topics specific to the NAEP mathematics assessment. Item writers should refer to directions for developing items provided by the assessment development contractor in addition to the information in Chapters Two, Three, and Four.

ITEM CHARACTERISTICS

Each item written for the NAEP mathematics assessment reflects two major dimensions: mathematical content area (see Chapter Two) and mathematical complexity (see Chapter Three).

Items will also vary by

- format (see Chapter Five, page xx, and this chapter, pages xx-xx),
- calculator status (see Chapter Five, page xx),
- whether manipulatives are required (see Chapter Five, page xx), and
- whether they are part of a family of items (see Chapter Five, page xx).

Each of these characteristics is discussed in the specifications, and examples of items illustrating these characteristics are given.

GENERAL PRINCIPLES OF ITEM WRITING

The following principles of good item writing should be used as appropriate, depending upon the measurement intent of the item. For example, a principle under graphics states, “represent each important part of the item in the visual images.” This principle is appropriate for items with graphics that are used to support students’ interpretation of the text but may not be appropriate for graphics that are used in other ways. The guidelines should be followed unless the targeted construct precludes doing so.

1. <u>Clear Measurement Intent</u>
   A critical step in good item writing is making sure that the measurement intent of the item is clear and that students understand what is being measured and what type of response is expected.

a) Clear intent in development

- Item writers should provide a clear description of what objective(s) and, if applicable, the portion(s) of the objective(s) each item is intended to measure, as well as specify the mathematical complexity of the item along with a rationale for the level of complexity. This will help classify items according to assessment specifications, help develop clear scoring rubrics and scoring materials, reduce confusion in reviews, and provide evidence of the degree of alignment of the assessment to the framework.
- In order to clearly measure the targeted objective(s), each item should be independent. The response to one item should not depend upon the response to another. For example, students should not be asked to determine the cost of an article in one item and then use that cost to determine sales tax in another item. Items can be related to one another, for example through being based on the same graphic display or by having the same theme, but their response requirements must be independent.

b) Clear intent for test takers

- It should be clear to the student what is being asked in each item. Writers should be careful not to make assumptions about how students will interpret an item's implicit requirements. Directions should be as straightforward as possible.
- Constructed-response items should contain clear directions to students about how they can respond. For example, can the response incorporate graphics or does it require the student to produce a verbal description? Is more than one type of response appropriate? Is more credit given for providing multiple solutions to a problem?
- If a constructed-response item requires students to show their work, the directions must clearly specify that students should include their work, provide direction for how students can show their work, and note that student work will be used in scoring the response.

Examples 20, 21, and 22 illustrate clear directions for students.

Example 21 has directions to students in how they should respond and includes required characteristics of the response.

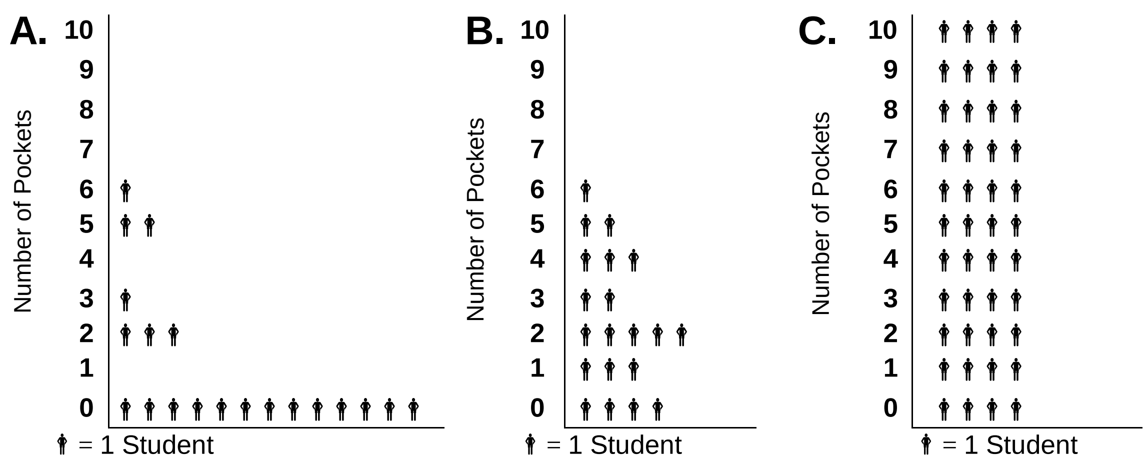
EXAMPLE 20 Grade 12 Geometry	Source: 1996 NAEP 12M13 #2 Low Complexity
In the space below, use your ruler to draw a parallelogram that has perpendicular diagonals. Show the diagonals in your sketch.	

Example 21 asks students to explain their answers, and it provides structure for them to do so. The item also includes direction to students about the format of their responses.

EXAMPLE 21 Grade 12 Data Analysis, Statistics and Probability	Source: 1992 NAEP 4M15 #10 Moderate Complexity

Think carefully about the following question. Write a complete answer. You may use drawings, words, and numbers to explain your answer. Be sure to show all of your work.

There are 20 students in Mr. Pang's class. On Tuesday, <u>most</u> of the students in the class said they had pockets in the clothes they were wearing.



Which of the graphs most likely shows the number of pockets that each child had? ____

Explain why you chose that graph.

Explain why you did not choose the other graphs.

Example 22 specifies that students should give mathematical evidence for their answer.

EXAMPLE 22 Grade 12 Number Properties and Operations	Source: 1996 NAEP 12M12 #8 Moderate Complexity

Luis mixed 6 ounces of cherry syrup with 53 ounces of water to make a cherry-flavored drink. Martin mixed 5 ounces of the same cherry syrup with 42 ounces of water. Who made the drink with the stronger cherry flavor?

Give mathematical evidence to justify your answer.

2. Plain Language
The purpose of using plain language is to clearly convey meaning without altering what items are intended to measure. All items should use plain language. Even when the intent of the item is for students to define, recognize, or use mathematics vocabulary correctly, the surrounding text should be in plain language. Plain language guidelines often increase access and minimize confusion for students[1].

- Write questions using brief, "simple" sentences or stems.
- Use the same structure for paragraphs throughout the assessment as much as possible (e.g., topic sentence, supporting sentences, and concluding sentence).
- Use present tense and active voice.
- Minimize paraphrasing.
- Avoid using pronouns.
- Use high-frequency words as much as possible.
- Avoid colloquialisms.
- When using words with multiple meanings, make sure the intended meaning is clear.
- Avoid using unnecessary descriptive information.
- Use format to clarify text (e.g., use bullets, allow space between pieces of text, and use boxes and lines judiciously).

Example 23 illustrates the use of present tense and active voice. No unnecessary descriptive information is included.

EXAMPLE 23 Grade 4 Number Properties and Operations	Source: 1996 NAEP 4M #22 (rev) Moderate Complexity

A fourth-grade class needs 5 leaves each day to feed its 2 caterpillars.
How many leaves do they need each day for 12 caterpillars?

Answer: ______________________

Use drawings, words, or numbers to show how you got your answer.

[1] In its review of the literature on effects of accommodations, the National Research Council suggests that "…language simplification may have removed factors that were construct-irrelevant for all test-takers." (p. 38)

Example 24 uses a simple, straightforward format.

EXAMPLE 24 — Source: 1990 NAEP 8M9 #16

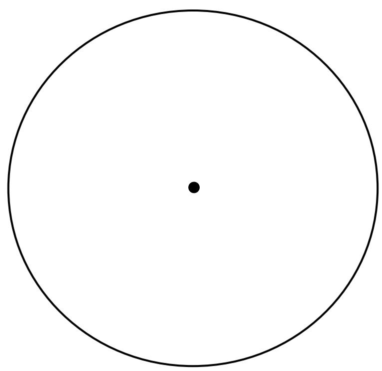
Grade 8 — Moderate Complexity

Data Analysis, Statistics and Probability

HAIR COLOR SURVEY
RESULTS

Color of Hair	Percentage
Blond	17
Brown	50
Black	33
Total	**100**

The table above shows the results of a survey of hair color. On the circle below, make a circle graph to illustrate the data in the table. Label each part of the circle graph with the correct hair color.



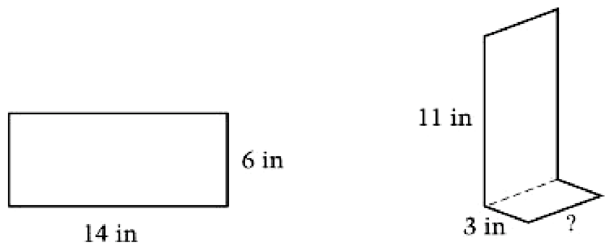
3. <u>Graphics</u>

   Graphics such as pictures, charts, and diagrams can be used to help students understand item text by depicting information in an item or a set of items. Graphics can be very effective in supporting text, illustrating mathematical concepts in the text, and increasing item access with minimal cost. If used improperly, however, graphics can add substantial confusion and distract test takers from what items are asking students to do. When using graphics,

   - use visuals that mirror and parallel the wording and expectations of the text.
   - illustrate only necessary information in the graphics.
   - represent each important part of the item in the visual images.

Example 25 is an item that illustrates only necessary information and parallels the expectations of the text.

EXAMPLE 25 Grade 4 Algebra	Source: 2005 NAEP 4M12 #13 Moderate Complexity



A piece of metal in the shape of a rectangle was folded as shown above. In the figure on the right, the "?" symbol represents what length?

A. 3 inches
B. 6 inches
C. 8 inches
D. 11 inches

Correct Answer: B

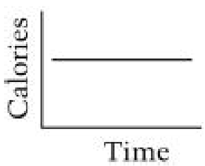
Example 26 shows the effective use of graphics as response options.

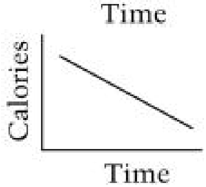
EXAMPLE 26
Grade 12
Algebra

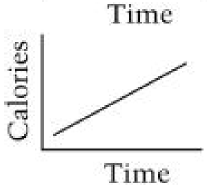
Source: 2005 NAEP 12M3 #5
Moderate Complexity

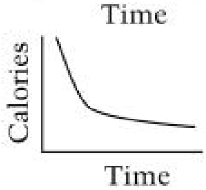
Exercise Time (minutes)	Total Calories Burned
10	37
12	42
14	49
16	58
18	64
20	70

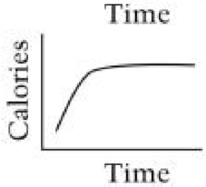
1. Which of the following graphs best illustrates the relationship between exercise time and total calories burned, as shown in the table above?

A) 

B) 

C) 

D) 

E) 

Correct Answer: C

Example 27 uses pictures to illustrate important information in the text.

EXAMPLE 27 Grade 4 Number Properties and Operations	Source: 1996 NAEP 4M #22 (rev) Moderate Complexity

A fourth-grade class needs 5 leaves each day to feed its 2 caterpillars.
How many leaves do they need each day for 12 caterpillars?

Answer: ______________________

Use drawings, words, or numbers to show how you got your answer.

4. Using Contextual Information Appropriately
   Often, items will be designed so that they measure mathematics in context. Contextual information includes problem scenarios, explanations, more thorough directions, and background text. Using contextual information judiciously can place mathematical concepts in fuller, often more realistic, conditions, measure students' ability to apply mathematical concepts, and provide necessary background information. However, the contextual information should not interfere with the mathematics being assessed or become a barrier to a student's ability to demonstrate his or her mathematical knowledge. The context should be the minimum necessary to set up the problem. Guidelines for using contextual information appropriately include the following:
   - Use contexts that are meaningful to the mathematics being assessed.
   - Use contexts that are appropriate for the grade level assessed.
   - Use realistic contexts.
   - Use familiar contexts; avoid contexts that may confuse or be unfamiliar to some students taking the assessment.

Examples 27 (above) and 28 show the use of contexts that are appropriate for the grade level. The contexts are likely to be familiar to students in 4th and 12th grade, respectively.

EXAMPLE 28 — Source: 1996 NAEP 12M12 #10

Grade 12 — Moderate Complexity

Data Analysis, Statistics and Probability

The table below shows the daily attendance at two movie theaters for 5 days and the mean (average) and the median attendance.

	Theater A	Theater B
Day 1	100	72
Day 2	87	97
Day 3	90	70
Day 4	10	71
Day 5	91	100
Mean (average)	75.6	82
Median	90	72

(a) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater A? Justify your answer.

(b) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater B? Justify your answer.

.

5. Writing Items with Multiple Access Points

Students vary in their abilities to access information and respond to tasks through visual, spatial, auditory, kinesthetic, and tactile pathways. When possible, items should be designed to allow students to approach and respond to the item in different ways. Incorporating multiple pathways appropriately in both constructed-response and multiple-choice items can increase the ability of the assessment to elicit responses from students across the range of achievement without affecting the mathematical content being assessed.

a) Multiple access in how items are written or presented to the students
- Use visuals to accompany and explain the text when appropriate.
- Write both multiple-choice and constructed-response items that allow for various approaches to determining the solution.
- Use manipulatives.
- Use a brief activity when appropriate to address students with auditory, kinesthetic, or tactile strengths.

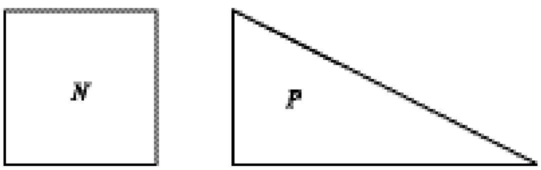
b) Multiple response opportunities
- Write constructed-response items that allow for multiple response formats. For example, students might be allowed to show their answers through illustrations, diagrams, or formulas.
- When possible, write constructed-response items and scoring rubrics that measure the mathematical knowledge and skills of students with a range of achievement.

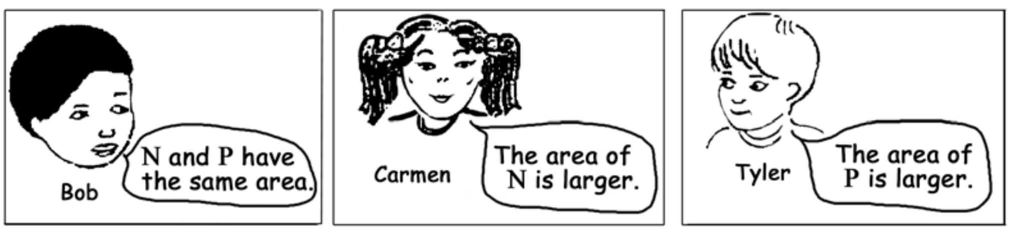
Example 27 (shown earlier) allows for a variety of response formats. Students can give verbal explanations, draw, or construct charts, for example.

Example 29 provides access through the use of dialogue, personalizing the mathematics. The item uses dialogue in a grade-appropriate way.

EXAMPLE 29 Grade 4 Measurement	Source: 1996 NAEP 4M10 #6 Moderate Complexity

The following shapes were provided to students. (Shapes were larger than shown.)





Who was correct? ______________________________

Use pictures and words to explain why.

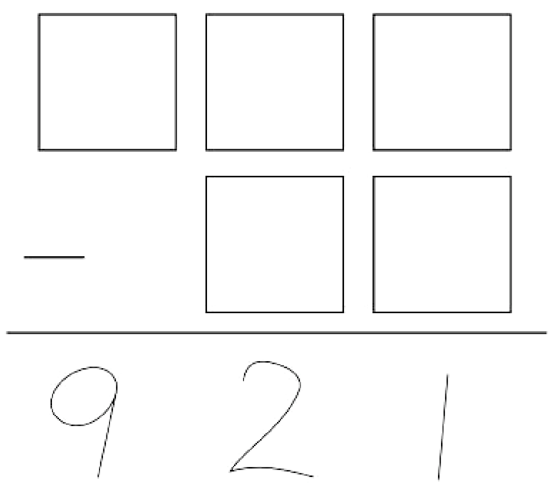
Example 30 provides access through the use of manipulatives, which allow students to use both tactile and kinesthetic pathways.

EXAMPLE 30	Source: 2005 NAEP 4M4 #11
Grade 4	Moderate Complexity
Number Properties and Operations	

This question refers to the number tiles. Please remove the 10 number tiles and the paper strip from your packet and put them on your desk.

Audrey used only the number tiles with the digits 2, 3, 4, 6, and 9. She placed one tile in each box below so the difference was 921.

Write the numbers in the boxes below to show where Audrey placed the tiles.



INCLUSION OF ENGLISH LANGUAGE LEARNERS AND STUDENTS WITH DISABILITIES

The mathematics assessments should be developed to allow for the participation of the widest possible range of students, so that interpretation of scores lead to valid inferences about levels of performance of the nation's students as well as valid comparisons across states. All students should have the opportunity to demonstrate their knowledge of the concepts and ideas that the NAEP Mathematics Assessment is intended to measure.

According to the National Research Council:

> Fairness, like validity, cannot be properly addressed as an afterthought once the test has been developed, administered, and used. It must be confronted throughout

> the interconnected phases of the testing process, from test design and development to administration, scoring, interpretation, and use (1999, pp. 80-81).

Current NAEP inclusion criteria for students with disabilities and English language learners are as follows (from NCES, 2007, Current Policy section):[2]

> Inclusion in NAEP of an SD [student with disabilities] 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 testing accommodations not allowed in NAEP are giving the reading 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.)

Most students with disabilities are eligible to be assessed in the NAEP program. Similarly, most students who are learning English as their second language are also eligible to participate in the NAEP assessment. There are two ways that NAEP addresses the issue of accessibility. One is to follow careful item and assessment development procedures to build accessibility into the standard assessment. The other is to provide accommodations for students with disabilities and for English language learners.

ACCOMMODATIONS

For many students with disabilities and students whose native language is other than English, the standard administration of the NAEP assessment will be most appropriate. For some students with disabilities and some English language learners, the use of one or more accommodations will be more suitable. How to select and provide appropriate accommodations is an active area of research, and new insights are emerging on how to best apply accommodation guidelines to meet the needs of individual students[3]. The NAEP mathematics accommodations policy allows for a variety of accommodations, depending upon the needs of each student. Most accommodations

[2] Identification of special needs students is problematic in and of itself. While NAEP defines inclusion criteria, student identification and classification is under the purview of states and localities. Consequently, there is wide variation from state to state on the percentage of students identified and classified as either SWD or ELL status. For example, the percentage of students identified as having a qualifying disability for the 2005 NAEP mathematics assessment varied across states from 10% to 24% at grade 4 and from 9% to 18% at grade 8 (Perie, Grigg, & Dion, 2005, p. 36). Similarly, the percentage of students identified as ELL for the 2005 NAEP mathematics assessment varied across states from 1% to 33% at grade 4 and from 1% to 21% at grade 8 (Perie, et al., 2005, p. 36). Of course, much of the ELL variation depends on the overall demographics of particular states, but some of the variation is influenced by identification criteria and procedures. Indeed, Koretz and Barton (2003-2004) and Abedi (2004) have raised questions about the reliability and validity of student classification into SWD and ELL, respectively.

[3] For example, the South Carolina Department of Education is working with a consortium of states to refine a research-based decision-making process for assigning accommodations to ELLs based on individual student characteristics (Kopriva and Carr, 2006) and to develop a decision-making framework for IEP teams to use in assigning accommodations to SWDs (South Carolina Department of Education, 2006).

that schools routinely provide in their own testing programs are allowed in the mathematics assessment, as long as they do not affect the construct tested. For example, it would NOT be appropriate to allow a student to use a calculator on a non-calculator block.

The most frequently used accommodations included (from the *NAEP 2006 Assessment Administrator Manual*, p. 2.28):

- Bilingual booklet (mathematics operational and science only)
- Bilingual dictionary (mathematics operational and science only)
- Large-print booklet
- Extended time in regular session
- Read aloud (mathematics and science only)
- Small group
- One-on-one
- Scribe or use of a computer to record answers
- Other—includes format or equipment accommodations such as a sign language translator or amplification devices (if provided by the school)
- Breaks during test
- Magnification device
- School staff administers

Accommodations are offered in combination as needed; for example, students who receive one-on-one testing generally also use extended time.

In a very small number of cases, students will not be able to participate in the assessment, even with the accommodations offered by NAEP:

- Students with disabilities whose Individualized Education Plan (IEP) teams or equivalent groups have determined that they cannot participate, **or** whose cognitive functioning is so severely impaired that they cannot participate, **or** whose IEP requires an accommodation that NAEP does not allow.
- Limited English proficient students who have received mathematics instruction primarily in English for less than three school years **and** who cannot participate in the assessment when it is administered in English, with or without an accommodation, or when the bilingual English/Spanish form is used.

ITEM WRITING CONSIDERATIONS TO MAXIMIZE ACCESSIBILITY

Appropriate specialists (e.g., specialists in educating students with disabilities, school psychologists, linguists, specialists in second language acquisition) should be involved in the entire process of test development, along with psychometricians, teachers, and content specialists. While such specialists often participate in item review, their expertise can be useful in the item development process. For example, in the case of addressing accessibility for English language learners, many issues regarding vocabulary (meaning of individual words), semantics (meaning of words in a sentence), syntax (grammatical structure of sentences), and pragmatics (interpretation of words and sentences in context) cannot be properly addressed without applying the methods and reasoning from linguistics.

This section addresses techniques for maximizing access to items for English language learners and students with disabilities. Chapter Five includes a section on maximizing accessibility at the assessment level. In addition, many of guidelines in the section, General Principles of Item Writing, address accessibility issues.

Item Writing Considerations for English Language Learners

Some students who are learning English as their second language will take the standard, English-only version of the test. These students are diverse both across and within their language groups. This is particularly the case with Spanish language speakers who come from various countries in Latin America and the Caribbean. Among the Spanish speaking population there are linguistic differences (mainly in vocabulary), cultural differences, and differences in educational and socio-economic backgrounds. English language learners may have trouble understanding what items are asking for on assessment forms administered in English.[4]

Although this section is specific to making items accessible to English language learners, the guidelines below can be applied to the development of mathematics items for all students. Many are extensions of the plain language principles found in the section, General Principles of Item Writing.

Vocabulary

The vocabulary of both mathematical English and ordinary English must be considered when developing mathematics items (Shorrocks-Taylor and Hargreave, 1999). In general, familiar words and natural language should be used in problems (Prins and Ulijn, 1998); mathematics-specific vocabulary should be used when appropriate, ordinary vocabulary when feasible, that is, when vocabulary knowledge is not part of targeted measurement construct and when ordinary English will not convey the precise meaning that a mathematical term will. For English language learners, the problem context or conventions that allow “shorthand” communication may not be enough to provide the meaning of specific words (Wong Fillmore and Snow, 2000). Avoid when possible the use of words with different meanings in ordinary English and mathematical English (e.g., odd), and clarify meanings of multi-meaning words when it is necessary to use them (e.g., largest can refer to size or quantity) (Kopriva, 2000; Brown, 1999).

- Avoid the use of idioms.
- Use concrete rather than abstract words. For example, instead of using “object” in a stem, refer to an actual object.
- Use shorter words, as long as they convey precise meaning.
- Use active verbs whenever possible.
- Limit the use of pronouns.
- Avoid using ambiguous referents as shorthand (for example, It is a good idea to…).
- Use cognates when appropriate (cognates are words that are similar in form and meaning in two languages).

[4] For more information about designing assessments that are accessible to English language learners, see Kopriva (in press) and Abedi (2006).

- Avoid the use of false cognates (words that are similar in form in two languages but have different meanings). For example, “once” means 11 in Spanish.

Text Structure

English and other languages often have different rules of syntax or word order. While students may know the basic differences between their primary language and English, subtle differences that can lead to confusion based on syntax or word order should be avoided. There are several features of test items that can impact comprehension.

- Use active voice (Abedi, 2006). For example, use “Laura drew a circle with a radius of 7 inches.” rather than “A circle was drawn with a radius of 7 inches.”
- Avoid long noun phrases (nouns with several modifiers) (Abedi, 2006).
- Use simple sentences as much as possible. However, do not compress sentences to the extent that important connecting information is omitted (Wong Fillmore and Snow, 2000; Shorrocks-Taylor and Hargreave, 1999).
  - Avoid long stems that contain adverbials (e.g., Although …., Gerald ….) and conditional (e.g., if/then) clauses (Abedi, 2006).
  - Avoid relative clauses and prepositional phrases (Abedi, 2006).
    For example, use “Marcie pays $2.00 for pens. Each pen costs $0.50. How many pens does Marcie buy?” rather than “If Marcie pays $2.00 for pens and each pen costs $0.50, how many pens does Marcie buy?”
- Write questions in the positive; avoid using negatives in the stem and item options.
- Avoid mixing mathematical symbols and text (Prins & Ulijn, 1998).
- Use formatting to keep items clear, separating key ideas (e.g., use bullets or frames).

Item Writing Considerations for Students with Disabilities

Most students with disabilities will take the standard assessment without accommodations, and those who take the assessment with accommodations will use the standard version of the test also. Item writers and the assessment developer should minimize item characteristics that could hinder accurately measuring the mathematics achievement of students with disabilities. Using the item writing considerations for English language learners will minimize some linguistic characteristics that can affect the responses of some students with disabilities. In addition, item writers should attend to the following recommendations.

- Avoid layout and design features that could interfere with the ability of the student to understand the requirements and expectations of the item.
- Use illustrations carefully. Thompson, Johnstone, and Thurlow (2002) provide guidance for the use of illustrations:
  - Minimize the use of purely decorative illustrations, since they can be distracting.
  - Use illustrations that can be enlarged or viewed with a magnifier or other assistive technology.
  - Use simple black and white line drawings when possible.
- Avoid item contexts that assume background experiences that may not be common to some students with sensory or physical disabilities.
- Develop items so that they can be used with allowed accommodations.

- Address alternatives for students who are not able to use the manipulatives necessary for responding to an item.

SCORING RESPONSES FROM ENGLISH LANGUAGE LEARNERS

Students' literacy status and varied background experiences have an impact on how well scorers can properly read, understand, and evaluate the responses of English language learners to constructed-response items.[5]

Literacy Issues

Responses sometimes can be difficult to read because language confusion arises between the students' native language and English. While this is developmentally appropriate in terms of language acquisition, many scorers are not trained to interpret these types of systematic errors. The following procedures should be used to score responses from English language learners appropriately:

- Scoring leaders should have additional training in recognizing and properly interpreting responses from English language learners.
- Experts in reading responses of English language learners should be available to scorers throughout the scoring process.
- Scorer training materials and benchmark student responses should illustrate features typically observed in the writing of ELLs in English. These features include:
    1. Code-switching—intermittent use of the student's native language and English.
    2. Use of native language or beginning-stage English phonetic spelling in attempting to respond in English or (e.g., *de ticher sed*—"the teacher said").
    3. Use of writing conventions from the native language when students are responding in English (e.g., *today is monday*—the names of the weekdays are not capitalized in standard Spanish).
    4. Word mergers (the condensing of words into one mega-word), transposition of words, and omission of tense markers, articles, plurals, prepositions, or other words.
    5. Use of technical notation conventions from the student's culture (e.g., $25,00 to express twenty-five dollars).
    6. Substitution of common words for more precise terminology. For instance, a student may use the word "fattest" to mean "greatest." If the item is not intended to evaluate students' knowledge and use of correct terminology, then this substitution may be acceptable. On the other hand, if the intent **is** to measure the students' knowledge and use of such terminology in an applied setting, then the substitution would be incorrect.
    7. Inappropriate use of unfamiliar words.
    8. Unusual sentence and paragraph structures that reflect discourse structures in the native language.

[5] For more information about scoring responses from English language learners, see Kopriva and Saez (1997) and Solano-Flores, Lara., Sexton, and Navarrete (2001).

9. Other features such as the transposition of words (e.g., *the cat black*) and omission of tense markers (e.g., *yesterday he learn at lot*), articles (*I didn't see it in notebook)*, plurals (*the horse are gone*), prepositions (e.g., *explain me what you said*), or other words.
10. Over-reliance on non-verbal forms of communication, such as charts or pictures.

Varied Background Experiences

Novel interpretations and responses are common for English language learners and often reflect background experiences quite different from that of most native English speakers. Scorers should evaluate responses based on the measurement intent of the item and recognize when an unusual response is actually addressing that intent.

At times, scoring rubrics implicitly or explicitly favor writing styles that mirror what is taught in language arts curricula in the U.S. schools. However, circular, indirect, deductive, and abbreviated reasoning writing styles are encouraged by some cultures, and scorers should be trained to appropriately score these types of responses. In addition, some cultures discourage children from giving long responses to questions, especially when authority figures ask the questions. Such a pattern of communication can be reflected in the written responses of ELLs to constructed-response questions. Despite being short, these responses to constructed-response items may be correct.

When a specific writing style is not the measurement intent of the item, scorers need to understand the nature, conventions, and approaches of these kinds of styles and how to separate the structure and sophistication of the written response from the substantive content being evaluated.

ITEM FORMATS

There are three types of items on the NAEP mathematics assessment: multiple-choice, short constructed-response, and extended constructed-response.

- **Multiple-choice** items require students to select one correct or best answer to a given problem. These items are scored as either correct or incorrect.
- **Short constructed-response** items require students to give a short answer such as a numerical result or the correct name or classification for a group of mathematical objects, draw an example of a given concept, or perhaps write a brief explanation for a given result. Short constructed-response items are scored according to scoring rubrics with two or three categories describing increasing degrees of knowledge and skill.
- **Extended constructed-response** items require students to consider a situation that demands more than a numerical response or a short verbal or graphic communication. If it is a problem to solve, for example, the student may be asked to carefully consider a situation choose a strategy to "solve" the situation, carry out the strategy, and interpret

the solution derived in terms of the original situation. Extended constructed-response items are typically scored according to scoring rubrics with five categories. [6]

Item writers should carefully consider the content and skills they intend to assess when deciding whether to write a multiple-choice or constructed-response item. Each content area includes knowledge and skills that can be measured using each of the three item formats and each level of mathematical complexity can be measured by any of the item formats. Carefully constructed multiple-choice items, for example, can measure any of the levels of mathematical complexity. Although a level of mathematical complexity may lend itself more readily to one item format, each type of item—multiple-choice, short constructed-response, and extended constructed-response—can deal with mathematics of greater or less depth and sophistication.

DEVELOPING MULTIPLE-CHOICE ITEMS

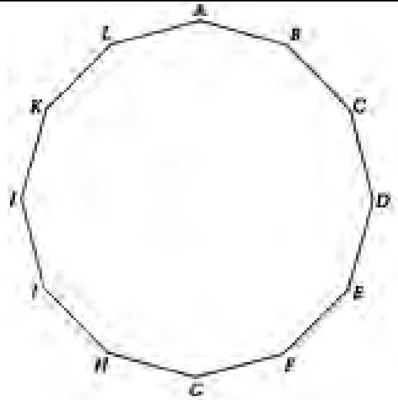
Multiple-choice items are an efficient way to assess knowledge and skills, and they can be developed to measure any of the levels of mathematical complexity. In a well-designed multiple-choice item, the stem clearly presents the problem to the student. The stem may be in the form of a question, a phrase, or a mathematical expression, as long as it conveys what is expected of the student. The stem is followed by either four or five answer choices, or options, only one of which is correct. In developing multiple-choice items, item writers should ensure that

- the stem includes only the information needed to make the student's task clear or needed to set the problem in an appropriate context.
- options are as short as possible.
- options are parallel in structure, syntax, and complexity.
- options do not contain inadvertent cues to the correct answer such as repeating a word from the stem in the correct answer or using specific determiners (e.g., all, never) in the distractors (incorrect options).
- distractors are plausible, but not so plausible as to be possible correct answers.
- distractors are designed to reflect the measurement intent of the item, not to trick students into choices that are not central to the mathematical idea being assessed.

Example 31 illustrates a straightforward stem with a direct question. The distractors are plausible, but only one is clearly correct.

[6] In some cases, it may be appropriate to have four scoring categories for an extended constructed-response item, depending upon the construct assessed and the nature of expected student responses to the item.

EXAMPLE 31 | Source: 2005 NAEP 4M3 #1
Grade 12 | Low Complexity
Geometry



Each of the 12 sides of the regular figure above has the same length.

1. Which of the following angles has a measure of 90°?

A. Angle *ABI*
B. Angle *ACG*
C. Angle *ADF*
D. Angle *ADI*
E. Angle *AEH*

Correct Answer: B

DEVELOPING CONSTRUCTED-RESPONSE ITEMS AND SCORING RUBRICS

The type of constructed-response item, short or extended, that is written should depend on the mathematical construct that is being assessed—the objectives and the level of complexity. Item writers should draft the scoring rubric as they are developing the item so that both the item and rubric reflect the construct being measured.

In developing the scoring rubric for an item, writers should think about what kind of student responses would show increasing degrees of knowledge and understanding. Writers should sketch condensed sample responses for each score category. Item writers also should include a mathematical justification or explanation for each rubric category description. Doing so will assist the writer in drafting a clear scoring rubric as well as provide guidance for scoring the item.

Short Constructed-Response Items

Some short constructed-response items are written to be scored dichotomously (that is, as correct or incorrect). Short constructed-response items with two scoring categories should measure knowledge and skills in a way that multiple-choice items cannot or provide greater evidence of students' understanding. Such short constructed-response items might be appropriate for measuring some computation skills, for example, to avoid guessing or estimation, which could be a factor if a multiple-choice item were used. They are also useful when there is more than one possible correct answer, when there are different ways to display an answer, or when a brief explanation is required. Item writers should take care that short constructed-response items would not be better or more efficiently structured as multiple-choice items—they should not be simply multiple-choice items without the options. Constructed-response items should be developed so that the knowledge and skills they measure are "worth" the additional time and effort to respond on the part of the student and the time and effort it takes to score the response.

Some short constructed-response items are written to be scored on a three-category scale. Short constructed-response items with three scoring categories should measure knowledge and skills that require students to go beyond giving an acceptable answer. These items allow for degrees of accuracy in a response so that a student can receive some credit for demonstrating partial understanding of the concept or skill measured by the item.

Item writers must draft a scoring rubric for each short constructed-response item. For dichotomous items, the rubrics should define the following two categories:

Correct
Incorrect

Examples 32 and 33 show dichotomous constructed-response items and their rubrics.

Example 32 requires students to perform a calculation, and the rubric simply defines a correct result.

EXAMPLE 32 Source: 2003 NAEP 8M7 #13
Grade 8 Low Complexity
Data Analysis, Statistics, and Probability

Score	Number of Students
90	1
80	3
70	4
60	0
50	3

The table above shows the scores of a group of 11 students on a history test. What is the average (mean) score of the group to the nearest whole number?

Answer: ________________________

SCORING GUIDE

1 – Correct response: 69

0 – Incorrect

Example 33 asks students to explain a concept, and the rubric lists some examples and provides guidance for determining whether other student responses are acceptable.

EXAMPLE 33 — Source: 1992 NAEP 8M14 #2

Grade 8 — Moderate Complexity

Number Properties and Operations

Tracy said, "I can multiply 6 by another number and get an answer that is smaller than 6."

Pat said, "No, you can't. Multiplying 6 by another number always makes the answer 6 or larger."

Who is correct? Give a reason for your answer.

SCORING GUIDE

1 – Correct

Tracy, with correct reason given.

OR

No name stated but reason given is correct.

Examples of correct reasons:

- If you multiply by a number *smaller* than 1 the result is less than 6.
- 6 x 0 = 0
- 6 x 1/2 = 3
- 6 x –1 = –6

0 – Incorrect

Tracy with no reason or incorrect reason

OR

Any response that states Pat is correct

OR

No name stated and reason given is incorrect

OR

Any other incorrect response

For items with three score categories, the rubrics should define the following categories:

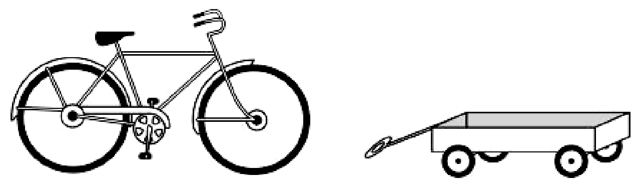
Correct
Partial
Incorrect

Examples 34 and 35 show constructed-response items with three score categories.

Example 34 requires students to demonstrate understanding of a concept, and the scoring rubric lists acceptable responses.

EXAMPLE 34 Grade 4 Algebra	Source: 2003 NAEP 4M7 #6 Moderate Complexity

A schoolyard contains only bicycles and wagons like those in the figure below.



On Tuesday the total number of wheels in the schoolyard was 24. There are several ways this could happen.

a. How many bicycles and how many wagons could there be for this to happen?
   Number of bicycles ________
   Number of wagons ________

b. Find another way that this could happen.
   Number of bicycles ________
   Number of wagons ________

SCORING GUIDE

Solution:
Any <u>two</u> of the following correct responses:

0 bicycles, 6 wagons
2 bicycles, 5 wagons
4 bicycles, 4 wagons
6 bicycles, 3 wagons
8 bicycles, 2 wagons
10 bicycles, 1 wagon
12 bicycles, 0 wagons

2 - Correct
Two correct responses

1 - Partial
One correct response, for either part a or part b
OR
Same correct response in both parts

0 - Incorrect
Any incorrect or incomplete response

Example 35 requires a completely correct answer for the top score category and gives a “partial” score for an answer that demonstrates an understanding of the appropriate ratio.

EXAMPLE 35 Grade 12 Number Properties and Operations	Source: 1996 NAEP 12M12 #8 Moderate Complexity
Luis mixed 6 ounces of cherry syrup with 53 ounces of water to make a cherry-flavored drink. Martin mixed 5 ounces of the same cherry syrup with 42 ounces of water. Who made the drink with the stronger cherry flavor? Give mathematical evidence to justify your answer.	
SCORING GUIDE	
2 – Correct Martin’s drink has the stronger cherry flavor. $6 \div 59 < 5 \div 47$ $0.1017 < 0.1064$ OR $6 \div 53 < 5 \div 42$ $0.1132 < 0.1190$ OR Luis; 1 part CS to 8.8 parts water < Martin: 1 part CS to 8.4 parts water.	
1– Partial Compares a pair of correct ratios for both Luis and Martin but does not give the correct answer.	
0 – Incorrect Incorrect response	

Extended Constructed-Response Items

In general, extended constructed-response items ask students to solve a problem by applying and integrating mathematical concepts or require that students analyze a mathematical situation and explain a concept, or both. Extended constructed-response items typically have five scoring categories:

Extended
Satisfactory
Partial
Minimal
Incorrect

Examples 36, 37, and 38 are extended constructed-response items.

Example 36 asks students to solve a problem and analyze a situation.

EXAMPLE 36 Source: 2003 NAEP 4M7 #20
Grade High Complexity
Algebra

The table below shows how the chirping of a cricket is related to the temperature outside. For example, a cricket chirps 144 times each minute when the temperature is 76°.

Number of Chirps Per Minute	Temperature
144	76°
152	78°
160	80°
168	82°
176	84°

The table below shows how the chirping of a cricket is related to the temperature outside. For example, a cricket chirps 144 times each minute when the temperature is 76°.

What would be the number of chirps per minute when the temperature outside is 90° if this pattern stays the same?

Answer: ________________________

Explain how you figured out your answer.

SCORING GUIDE

Extended
Answers 200 with explanation that indicates number of chirps increases by 8 for every temperature increase of 2°.

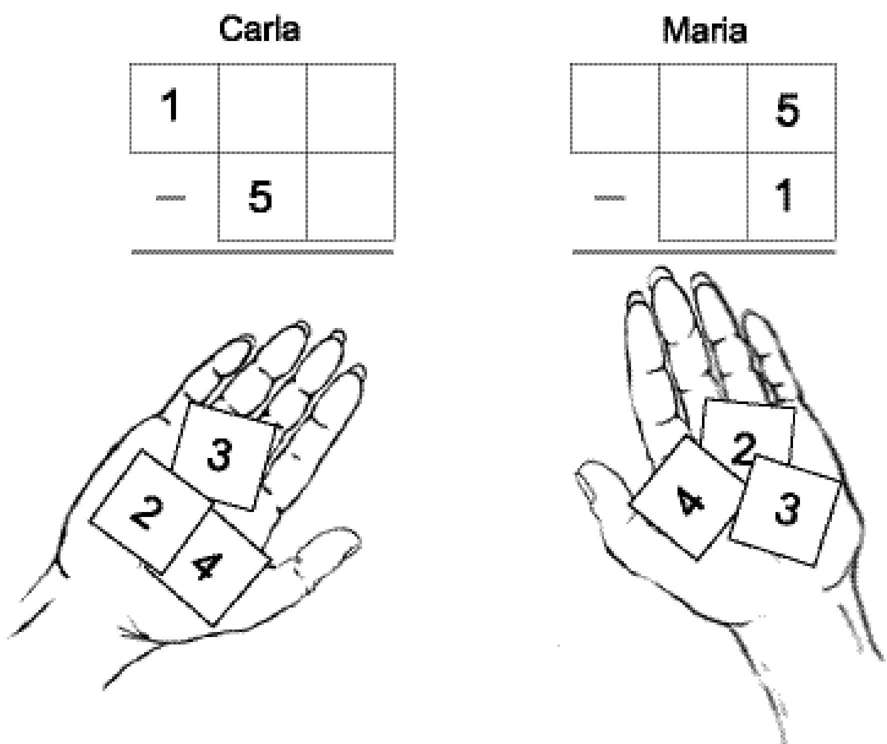
Satisfactory Gives explanation that describes ratio, but does not carry process far enough (e.g., gives correct answer for 86° (184) or 88° (192) or carries process too far (answers 208)). OR Answers 200 and shows 184 86°, 192 88°, 200 90° but gives no explanation. OR Answers 200 with explanation that is not stated well but conveys the correct ratio. OR Gives clear description of ratio and clearly has minor computational error (e.g., adds incorrectly).
Partial Answers between 176 and 208, inclusive, with explanation that says chirps increase as temperature increases. OR Answers between 176 and 208, inclusive, with explanation that they counted by 8 (or by 2). OR Uses a correct pattern or process (includes adding a number 3 times or showing 184 and 86 in chart) or demonstrates correct ratio. OR Has half the chart with 200 on the answer line. OR "I added 24" (with 200 on answer line).
Minimal Answers between 176 and 208, inclusive, with no explanation or irrelevant or incomplete explanation. OR Has explanation that number of chirps increases as temperature increases but number is not in range. OR Has number out of range but indicates part of the process (e.g., I counted by 8's) OR Explanation—as temperature increases the chirps increase but number is out of range.
Incorrect Incorrect response.

Example 37 requires students to explain and justify a solution.

EXAMPLE 37 Grade 8 Number Properties and Operations	Source: 1996 NAEP 8M3 #13 High Complexity

In a game, Carla and Maria are making subtraction problems using tiles numbered 1 to 5. The player whose subtraction problem gives the largest answer wins the game.

Look at where each girl placed two of her tiles.



Who will win the game?________________________

Explain how you know this person will win.

SCORING GUIDE

Explanations:

The following reasons may be given as part of an Extended, Satisfactory, or Partial correct answer:

a. The largest possible difference for Carla is less than 100 and the smallest possible difference for Maria is 194.

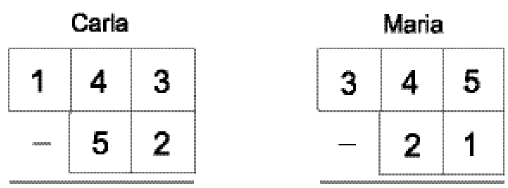
b. Carla can only get a difference of 91 or less, but Maria can get several larger differences.

c. Carla can have only up to 143 as her top number, but Maria can have 435 as her largest number.

d. Carla has only 1 hundred but Maria can have 2, 3, or 4 hundreds.

e. Maria can never take away as much as Carla.

f.



Any combination of problems to show that Maria's difference is greater.

4 – Extended

Student answers Maria and gives explanation such as (a) or (b), or an appropriate combination of the other explanations.

3– Satisfactory

Student answers Maria and gives explanation such as (c), (d), or (e).

2– Partial

Student answers Maria with partially correct, or incomplete but relevant, explanation.

1– Minimal

Student answers Maria and gives sample such as in (f) but no explanation OR answers Maria with an incorrect explanation.

0 – Incorrect

Incorrect response

Example 38 requires students to demonstrate understanding of a concept by explaining its use.

EXAMPLE 38 — Source: 1996 NAEP 12M12 #10
Grade 12 — Moderate Complexity
Data Analysis, Statistics and Probability

The table below shows the daily attendance at two movie theaters for 5 days and the mean (average) and the median attendance.

	Theater A	Theater B
Day 1	100	72
Day 2	87	97
Day 3	90	70
Day 4	10	71
Day 5	91	100
Mean (average)	75.6	82
Median	90	72

(a) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater A? Justify your answer.

(b) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater B? Justify your answer.

SCORING GUIDE

Example Explanations

An appropriate explanation for Theater A should include that the attendance on day 4 is much different than the attendance numbers for any other days for Theater A.

An appropriate explanation for Theater B should include the following ideas:

- There are two clusters of data.
- The median is representative of one of the clusters while the mean is representative of both.

OR

a justification that conveys the idea that 82 is a better indicator of where the "center" of the 5 data points is located

4 – Extended

Indicates mean for Theater B and median for Theater A and gives a complete explanation for each measure.

3– Satisfactory

Indicates mean for Theater B and median for Theater A and gives a complete explanation for one measure.

2– Partial Indicates mean for Theater B and median for Theater A with either no explanation or an incomplete explanation OR selects the better measure for one theater and gives an appropriate explanation.
1– Minimal Indicates the mean for Theater B with no explanation or an incomplete explanation OR indicates the median for Theater A with no explanation or an incomplete explanation.
0 – Incorrect Incorrect response

Item writers must develop a draft scoring rubric specific to each extended constructed-response item. The rubric should clearly reflect the measurement intent of the item. The next section describes some requirements for writing scoring rubrics.

Aligning Items and Rubrics

Item writers should refer back to the measurement intent of the item when they are developing its scoring rubric. The number of categories used in the rubric should be based upon the mathematical demand of the item.

1. <u>Defining the Score Categories</u>
   Each score category must be distinct from the others; descriptions of the score categories should clearly reflect increasing understanding and skill in the targeted mathematics constructs. Distinctions among the categories should suggest the differences in student responses that would fall into each category; the definitions must be clear enough to use in training scorers. Each score level should be supported by the mathematical intent of the item. Factors unrelated to the measurement intent of the item should not be evaluated in the rubric. For example, if an item is not meant to measure writing skills, then the scoring rubric should be clear that the demonstration of mathematics in the response does not need to be tied to how well the response is written. However, if an explanation is part of the item requirement, the rubric should reflect that such explanations should be clear and understandable.

   The three extended response items above (Examples 36, 37, and 38) all have distinct, well-defined score categories. Example 36 is a good illustration of a rubric that clearly differentiates among score categories. The categories describe increasing mathematical understanding as the scores increase, and the descriptions are in terms of student responses.

2. <u>Measuring More than One Concept</u>
   If an item is measuring more than one skill or concept, the description of the score categories in the rubric should clearly reflect increasing understanding and achievement in each area. For instance, if the item is measuring both students' understanding of a topic in algebra and their skill in developing an appropriate problem solving approach, then the description of each category in the rubric should explain how students' understanding and skill are

evaluated. If an item requires both an acceptable solution and an explanation, the rubric should show how these two requirements are addressed in each score category.

Example 38 requires two correct answers accompanied by justifications, and the scoring rubric clearly addresses both.

3. Specifying Response Formats
Unless the item is measuring whether or not a student can use a specified approach to a problem, each score category should allow for various approaches to the item. It should be clear in the rubric that different approaches to the item are allowed. Varied approaches may include

- different mathematical procedures; for example, adding or multiplying may lead to the same solution.
- different representations; for example, using a diagram or algorithms to solve or explain a solution may be appropriate.

Scorer training materials should include examples of student work that illustrate various appropriate approaches to the item.

ITEM TRYOUTS AND REVIEWS

Appendix B contains the *NAEP Item Development and Review Policy Statement* (National Assessment Governing Board, 2002), which explicates the following six guiding principles:

Principle 1
NAEP test questions selected for a given content area shall be representative of the content domain to which inferences will be made and shall match the NAEP assessment framework and specifications for a particular assessment.

Principle 2
The achievement level descriptions for basic, proficient, and advanced performance shall be an important consideration in all phases of NAEP development and review.

Principle 3
The Governing Board shall have final authority over all NAEP test questions. This authority includes, but is not limited to, the development of items, establishing the criteria for reviewing items, and the process for review.

Principle 4
The Governing Board shall review all NAEP test questions that are to be administered in conjunction with a pilot test, field test, operational assessment, or special study administered as part of NAEP.

Principle 5
NAEP test questions will be accurate in their presentation and free from error. Scoring criteria will be accurate, clear, and explicit.

Principle 6
All NAEP test questions will be free from racial, cultural, gender, or regional bias, and must be secular, neutral, and non-ideological. NAEP will not evaluate or assess personal or family beliefs, feelings, and attitudes, or publicly disclose personally identifiable information.

The test development contractor should build careful review and quality control procedures into the assessment development process. Although large-scale field testing provides critical statistical item-level information for test development, other useful information about the items should be collected before and after field testing. Before field testing, items and scoring rubrics should be reviewed by experts in mathematics and educational measurement, including mathematics teachers and representatives of state education agencies, and by reviewers trained in sensitivity review procedures. After field testing, the items and the assessment as a whole should be reviewed to make sure that they are as free as possible from irrelevant variables that could interfere with students' demonstrating their mathematical knowledge and skills.

Sensitivity reviews are a particularly important part of the assessment development process. Reviewers should include educators and community members who are experts in the schooling or cultural backgrounds of students in the primary demographic groups taking the assessment, including special needs students and English language learners. The reviewers focus on checking that test items are fair for all students, identifying any items that contain offensive or stereotypical subject matter and other irrelevant factors. They provide valuable guidance about the context, wording, and structure of items, and they identify flaws in the items that confound the validity of the inferences for the groups of students they represent.

Classroom tryouts and cognitive labs are two particularly useful procedures for collecting information about how specific items or item prototypes are working. The information collected is valuable for determining whether items are measuring the construct as intended and for refining the items and scoring procedures before field testing. These two techniques can be targeted to specific samples of students (for example, to see if items are appropriately written for English language learners) or to students representing the characteristics of the assessment population. The information that the test development contractor collects from classroom tryouts and cognitive labs should be provided to item writers to help them develop new items and revise existing items before field testing, and it can be used to enhance item writing training and reference materials.

CLASSROOM TRYOUTS

Classroom tryouts are an efficient and cost-effective way to collect information from students and teachers about how items and directions are working. Tryouts allow the test developer to troubleshoot the items and scoring rubrics. Classroom tryouts usually involve a non-random but carefully selected sample; the students should reflect the range of student achievement in the target population and represent the diversity of examinees. For example, the tryout sample should include urban and rural schools; schools in low, middle, and high economic communities;

schools from different regions; and schools with students in all the major racial/ethnic categories in the population. The more the sample represents various groups in the testing population, the more likely the tryout will identify areas that can be improved in the items.

In addition to providing student response data, tryouts can provide various kinds of information about the items, including what students and their teachers think the items are measuring, the appropriateness of the associated test materials, and the clarity of the instructions. Students can be asked to evaluate the items, for example by circling words, phrases, or sentences they find confusing and making suggestions for improvement. Teachers can ask students what they thought each item was asking them to do and why they answered as they did, and provide the information to the test developer. Teachers can be asked to edit items and associated test materials. Item tryouts also are an efficient way to test how accommodations work and to try out new manipulatives and other materials.

Student responses to the items should be reviewed by content and measurement experts to detect any problems in the items and should be used along with the other information gathered to refine the items and scoring rubrics. Using a sample that includes various subgroups in the population will allow reviewers to look for issues that might be specific to these groups. Responses also are useful in developing training materials and notes for item scorers.

COGNITIVE LABS

In cognitive labs, students are interviewed individually while they are taking or shortly after they have completed a set of items. Because cognitive labs highlight measurement considerations in a more in-depth fashion than other administrations can, their use can provide important information for item development and revision. For example, cognitive labs can identify if and why an item is not providing meaningful information about student achievement, provide information about how new formats are working, or identify why an item was flagged by differential item functioning (DIF) analyses.

The student samples used in cognitive labs are much smaller than those used in classroom tryouts. Students are selected purposefully to allow an in-depth understanding of how an item is working and to provide information that will help in revising items or in developing a particular type of item.

CHAPTER FIVE

DESIGN OF THE ASSESSMENT

The NAEP mathematics assessment is complex in its structure. The design of the assessment demands that multiple features stay in balance. A broad range of content should be adequately covered by the test items, balanced at each grade level according to the required distribution for each content area. At the same time, items make differing cognitive demands on students according to how mathematically complex they are. This, too, requires balance. The assessments also need to be balanced according to the three item formats: multiple-choice, short constructed response, and extended constructed response. An additional balance issue involves the mathematical setting of the item, whether it is purely mathematical or set in a real-world context.

There are other features of both the test and the items that are important in the design of a valid and reliable assessment. These include how sampling is used in the design of NAEP, the use of calculators, and the use of manipulatives and other tools. Of critical importance is the issue of accessibility for all students, which is addressed in several different ways on the NAEP mathematics assessments. A final design feature of NAEP is the use of families of items. Each of these features and issues is described in this chapter.

ALIGNMENT OF THE ASSESSMENT AND THE FRAMEWORK

The assessment should be developed so that it is aligned with the content expectations and complexity levels defined by the NAEP *Mathematics Framework for 2009.* Drawing upon Webb and others[7], five interrelated dimensions are considered in structuring the NAEP assessment so that it is aligned with the framework:

1. The match between the content of the assessment and the content of the framework: The assessment as a whole should reflect the breadth of knowledge and skills covered by the topics and objectives in the framework.
2. The match between the complexity of mathematical knowledge and skills on the assessment and in the framework: The assessment should represent the balance of levels of mathematical complexity at each grade as described in the framework.
3. The match between the emphasis of the assessment and the emphasis of topics, objectives, and contextual requirements in the framework: The assessment should represent the balance of content and item formats specified in the framework and give appropriate emphasis to the conditions in which students are expected to demonstrate their mathematics achievement, reflecting the use of calculators, manipulatives, and real-world settings.
4. The match between the assessment and how scores are reported and interpreted: The assessment should be developed so that scores will reflect both the framework and the performance described in the NAEP achievement levels.
5. The match between the assessment design and the characteristics of the targeted assessment population: The assessment should give all students tested a reasonable

[7] For example, Webb, Ely, Cormier, and Vesperman (2006)..

opportunity to demonstrate their knowledge and skills in the topics and objectives covered by the framework.

These five dimensions are the foundation for the NAEP assessment and item specifications. The principles in these dimensions are used in each of the sections that follow.

ACCESSIBILITY

The NAEP mathematics assessment is designed to measure the achievement of students across the nation. Therefore, it should allow students who have learned mathematics in a variety of ways, following different curricula and using different instructional materials; students who have mastered the content to varying degrees; students with disabilities; and students who are English-language learners to demonstrate their content knowledge and skill. The design issue for the assessment is: what is a reasonable way to measure the same mathematics for students who come to the assessment with different experiences, strengths, and challenges, who approach the mathematics from different perspectives, and who have different ways of displaying their knowledge and skill?

Two methods NAEP uses to design an accessible assessment program are (1) developing the standard assessment so that it is accessible and (2) providing accommodations for students with special needs. In Chapter Four each of these design issues is described in greater detail.

REPORTING REQUIREMENTS

The NAEP mathematics assessment reports results for the nation's students in fourth, eighth, and twelfth grades and participating states' students in grades 4 and 8, as well as for subgroups of the population defined by specific demographic characteristics such as gender, type of school attended, and eligibility for free and reduced-priced lunch (see the NAEP Report Card for more information about subgroup scores). Scores are not reported for individual students or schools.

Results are reported in two ways: scale scores and achievement levels. Scale scores can range from 0 to 500. Reports identify the percentage of students who reach three achievement levels—basic, proficient, and advanced. Briefly,

- **Basic** denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.
- **Proficient** represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.
- **Advanced** represents superior performance.

These levels are intended to provide descriptions of what students *should* know and be able to do in mathematics. Established for the 1992 mathematics scale through a broadly inclusive process and adopted by the National Assessment Governing Board, the three levels per grade are the primary means of reporting NAEP data. The updated mathematics framework was developed with these levels in mind to ensure congruence between the levels and the test content. See Appendix A for the NAEP Mathematics Achievement Level Descriptions.

The assessment should be designed so that results can be reliably and validly reported for the population and for subgroups by scale scores and in terms of the achievement levels. Because the results are intended to describe the achievement of all students in the nation, results should provide an accurate picture of achievement across the entire scoring scale.

The test developer should design the assessment so that the content is aligned with the content framework and the knowledge and skills described in the achievement levels. In addition, before new items are developed, the test developer should review the items available for the assessment and the test information functions for the assessment at each grade level to identify the need for new items that measure performance across the achievement scale, including at the upper and lower ends.

BALANCE OF CONTENT

As described in Chapter Two, each NAEP mathematics item is developed to measure one of the objectives, which are organized into the five major content areas of mathematics. The table below shows the distribution of items by grade and content area. See Chapter Two for more details.

Table 2. Percentage Distribution of Items by Grade and Content Area

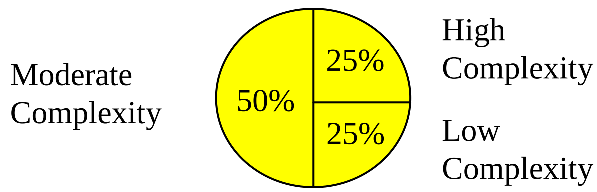
<table>
<tr><th>Content Area</th><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td>Number Properties and Operations</td><td>40%</td><td>20%</td><td>10%</td></tr>
<tr><td>Measurement</td><td>20%</td><td>15%</td><td rowspan="2">30%</td></tr>
<tr><td>Geometry</td><td>15%</td><td>20%</td></tr>
<tr><td>Data Analysis, Statistics, and Probability</td><td>10%</td><td>15%</td><td>25%</td></tr>
<tr><td>Algebra</td><td>15%</td><td>30%</td><td>35%</td></tr>
</table>

BALANCE OF MATHEMATICAL COMPLEXITY

As described in Chapter Three, items are classified according to the level of demands they make on students. This is known as the mathematical complexity of the item. Each item is considered to be at one of three levels of complexity: Low, Moderate, or High.

The ideal balance sought for the 2009 NAEP is not necessarily the balance one would wish for curriculum or instruction in mathematics education. Balance here must be considered in the context of the constraints of an assessment such as NAEP. These constraints include the timed nature of the test and its paper-and-pencil format. Items of high complexity, for example, often take more time to complete. At the same time, some items of all three types are essential to assess the full range of students' mathematical achievement.

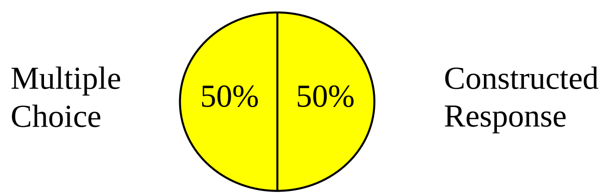
The ideal balance would be that half of the total testing time on the assessment is spent on items of moderate complexity, with the remainder of the total time spent equally on items of low and high complexity. This balance would apply for all three grade levels.



Percent of Testing Time at Each Level of Complexity

BALANCE OF ITEM FORMATS

Items consist of three formats: multiple-choice, short constructed response, and extended constructed response (see Chapter Three for an in-depth discussion of each type). Testing time on NAEP is divided evenly between multiple-choice items and both types of constructed response items, as shown below:



Percent of Testing Time by Item Formats

The design of the assessment, then, must take into account the amount of time students are expected to spend on each of the three formats of items.

BALANCE OF ITEM CONTEXTS

Just as mathematics can be separated into "pure" and "applied" mathematics, NAEP items should seek a balance of items that measure students' knowledge within both realms. Therefore some items will deal with purely mathematical ideas and concepts, while others will be set in the contexts of real-world problems.

In the two pairs of examples below, the first item is purely mathematical, while the second is set in the context of a real-world problem.

EXAMPLE PAIR 1 – PURE MATHEMATICAL SETTING Source: 2005 NAEP 4M4 #1	
Grade 4, Number Properties and Operations: Number operations	Percent Correct: 76% No Calculator

Subtract:

$$\begin{array}{r} 972 \\ \underline{-\ 46} \end{array}$$

Answer: ______________________

Correct Answer: 926

EXAMPLE PAIR 1 – CONTEXTUAL MATHEMATICAL SETTING	
Grade 4 Number Properties and Operations: Number operations	Source: 2005 NAEP 4M12 #4 Percent Correct: 80% No calculator

There are 30 people in the music room. There are 74 people in the cafeteria. How many more people are in the cafeteria than the music room?

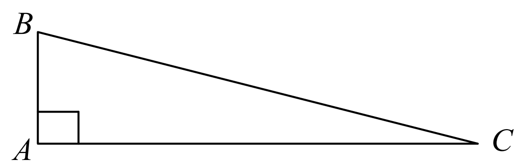
A. 40
B. 44
C. 54
D. 104

Correct Answer: B

Both items involve computation. In the first item the operation is specified. In the other item, the students must interpret the contextual situation, recognizing that it calls for the finding the difference between 74 and 30, and then compute.

EXAMPLE PAIR 2 – PURE MATHEMATICAL SETTING Modified Item
Grade 12 Calculator available Measurement: Measuring physical attributes

In the triangle below $\overline{AC} = 50$ feet and $m\angle C = 40°$.



What is the length of $\overline{AB}$ to the nearest foot?

Answer: ____________________

Correct Answer: 42 feet

EXAMPLE PAIR 2 – CONTEXTUAL MATHEMATICAL PROBLEM Grade 12
Measurement: Measuring physical attributes
Source: 2005 NAEP B3M12 #15
Percent Correct: 41%
Calculator available

A cat lies crouched on level ground 50 feet away from the base of a tree. The cat can see a bird's nest directly above the base of the tree. The angle of elevation from the cat to the bird's nest is 40°. To the nearest foot, how far above the base of the tree is the bird's nest?

A. 32
B 38
C. 42
D. 60
E. 65

Correct Answer: C

NAEP Administration and Student Sampling

As currently planned, the 2009 mathematics assessment will be administered between January and early April, with the administration conducted by trained field staff. Each mathematics assessment booklet will contain two separately timed, 25-minute sections of mathematics items. Three types of items will be used: multiple-choice, short constructed-response, and extended constructed-response. The assessment is designed so that there are multiple forms of the test booklets. The items will be distributed across the booklets using a matrix sampling design so that students taking part in the assessment do not all receive the same items. In addition to the mathematics items, the assessment booklets will include background questionnaires, administered in separately timed sessions[8].

The assessment is designed to measure mathematics achievement of students in the nation's schools in grades 4, 8, and 12 and report the results at the national, regional, and state levels. To implement this goal, schools throughout the country are randomly selected to participate in the assessment. The sampling process is carefully planned to select schools that accurately represent the broad population of U.S. students and the populations of students in each state participating in State NAEP.

The selection process is designed to include schools of various types and sizes from a variety of community and geographical regions, with student populations that represent different levels of economic status, racial, ethnic and cultural backgrounds, and instructional experiences. Students with disabilities and English language learners are included to the extent possible, with accommodations as necessary (see Chapter Four for more information about inclusion criteria and accommodations). The sophisticated sampling strategy helps to ensure that the NAEP program can generalize the assessment findings to the diverse student populations in the nation and participating jurisdictions. This allows the program to present information on the strengths and weaknesses in aggregate student understanding of mathematics and the ability to apply that understanding in problem-solving situations; provide comparative student data according to race/ethnicity, type of community, and geographic region; describe trends in student performance over time; and report relationships between student achievement and certain background variables.

CALCULATORS

The assessment contains blocks (sets of items) for which calculators are not allowed, and calculator blocks, which contain some items that would be difficult to solve without a calculator. At each grade level, approximately two-thirds of the blocks measure students' mathematical knowledge and skills without access to a calculator; the other third of the blocks allow the use of a calculator. The type of calculator students may use on a calculator block varies by grade level, as follows:

- At grade 4, a four-function calculator is supplied to students, with training at the time of administration.
- At grade 8, a scientific calculator is supplied to students, with training at the time of administration.
- At grade 12, students are allowed to bring whatever calculator, graphing or otherwise, they are accustomed to using in the classroom with some restrictions for test security purposes (see below). A scientific calculator is supplied to students who do not bring a calculator to use on the assessment.

[8] See *The Nation's Report Card: Mathematics 2005*, available at http://nces.ed.gov/nationsreportcard/mathematics/,and NAEP technical reports produced by NCES for more information about administration conditions and procedures.

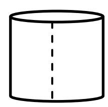
No items on the 2009 NAEP at either grade 8 or grade 12 will be designed to provide an advantage to students with a graphing calculator. Estimated time required for any item should be based on the assumption that students are not using a graphing calculator.

The assessment developer will propose restrictions on calculator use in grades 8 and 12 to (1) help ensure that items in calculator blocks cannot be solved in ways that are inconsistent with the knowledge and skills the items are intended to measure and (2) to maintain the security of NAEP test materials. These restrictions will address issues such as calculators with QWERTY keyboards, communication between students during testing, and the use of stored formulas, algorithms, and other procedures.

Items are categorized according to the degree to which a calculator is useful in responding to the item:

- A calculator inactive item is one whose solution neither requires nor suggests the use of a calculator.

EXAMPLE – CALCULATOR INACTIVE ITEM	Source: 2005 NAEP 8M3 #4
Grade 8	Percent Correct: 86%
Geometry: Transformation of shapes and preservation of properties	Calculator available



The paper tube in the figure above is to be cut along the dotted line and opened up. What will be the shape of the flattened piece of paper?

Answer: ________________________

Correct Answers:
Rectangle or Square

- A calculator is not necessary for solving a calculator neutral item; however, given the option, some students might choose to use one.

EXAMPLE – CALCULATOR NEUTRAL ITEM Grade 8 Algebra: Patterns, relations, and functions	Source: 2005 NAEP 8M3 #12 Percent Correct: 60% Calculator available

$$1 + 3 = 4$$
$$1 + 3 + 5 = 9$$
$$1 + 3 + 5 + 7 = 16$$
$$1 + 3 + 5 + 7 + 9 = 25$$

According to the pattern suggested by the four examples above, how many consecutive odd integers are required to give a sum of 144?

A. 9
B. 12
C. 15
D. 36
E. 72

Correct Answer: B

- A calculator is necessary or very helpful in solving a calculator active item; a student would find it very difficult to solve the problem without the aid of a calculator.

EXAMPLE – CALCULATOR ACTIVE ITEM Grade 12 Measurement: Measuring physical attributes	Source: 2005 NAEP 3M12 #15 Percent Correct: 41% Calculator available

A cat lies crouched on level ground 50 feet away from the base of a tree. The cat can see a bird's nest directly about the base of the tree. The angle of elevation from the cat to the bird's nest is 40°. To the nearest foot, how far above the base of the tree is the bird's nest?

A. 32
B. 38
C. 42
D. 60
E. 65

Correct Answer: C

The assessment uses reasonable manipulative materials, where possible, in measuring students' ability to represent their understandings and to use tools to solve problems. Such manipulative materials and accompanying tasks are carefully chosen to cause minimal disruption of the test administration process. Examples of such materials are number tiles, geometric shapes, rulers, and protractors.

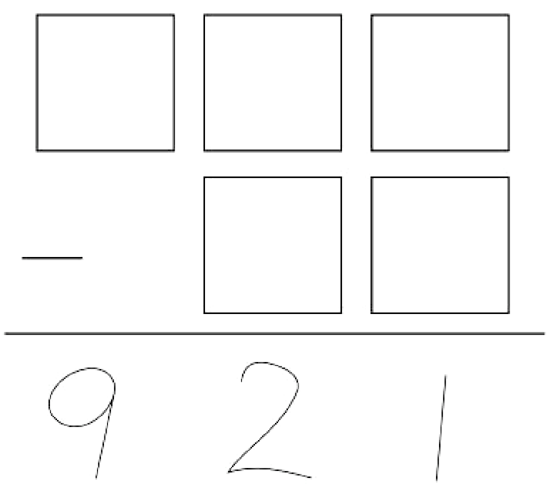
In the following example, number tiles are provided to the student.

EXAMPLE – NUMBER TILES PROVIDED Grade 4 Number Properties and Operations: Number operations	Source: 2005 NAEP 4M4 #11 Percent correct: 47%(Correct) 42% (Partially Correct) No Calculator, tiles provided

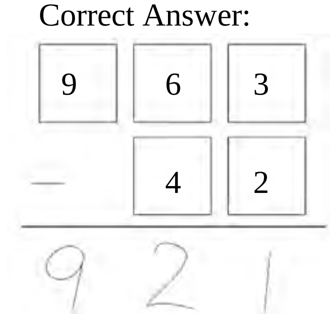
This question refers to the number tiles. Please remove the 10 number tiles and the paper strip from your packet and put them on your desk.

Audrey used only the number tiles with the digits 2, 3, 4, 6, and 9. She placed one tile in each box below so the difference was 921.

Write the numbers in the boxes below to show where Audrey placed the tiles.



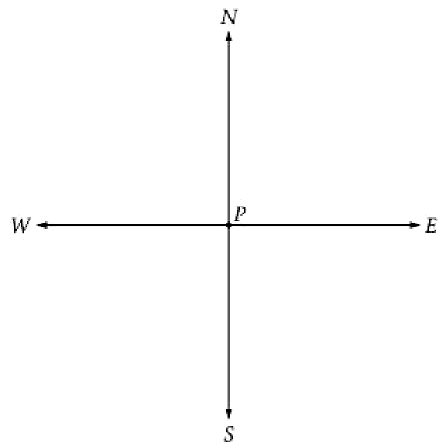
Correct Answer:



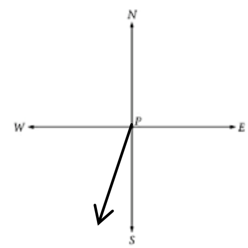
In the next example, students are provided with a protractor.

EXAMPLE – PROTRACTOR PROVIDED Grade 8 Measurement: Measuring physical attributes	Source: 2005 NAEP 8M3 #2 Percent correct: 21% 47% (Partially correct) No calculator, protractor provided

The weather service reported a tornado 75° south of west. On the figure below, use your protractor to draw an arrow from *P* in the direction in which the tornado was sighted.



Correct Answer:



ITEM FAMILIES

Item families are groups of related items designed to measure the depth of student knowledge within a particular content area (vertical item families) or the breadth of student understanding of specific concepts, principles, or procedures across content areas (horizontal item families). Within a family, items may cross content areas, vary in mathematical complexity, and cross grade levels.

Using item families in different ways provides for a more in-depth analysis of student performance than would a collection of discrete, unrelated items. The results might show the degree to which students can solve problems in a given content area at increasing levels of complexity. Or they might give evidence about how a mathematical concept is more accessible to students through one type of model than another. These two types of families are illustrated below.

The first example illustrates a family designed to measure depth of grade 8 students' knowledge of the relationship of area of a square and the length of the sides of the square.
The first item is straightforward, asking the students to determine the length of a side given the area of the square. It would be classified as low complexity. The second item, classified as moderate complexity, requires the students to first determine the area and then find the length of a side. The third item brings together measurement and algebra, but relies on student knowledge

about area of a square and length of its sides. Students must attend to several conditions and write an equation to describe a general situation. It would be classified as high complexity.

The second example illustrates a family designed for grade 4 to measure breath of understanding of the meaning of fractions. Two models, area and sets, are used. For each model, two versions are given. In the first version, the model has the same number of parts or pencils as the denominator. In the second version the number of parts or pencils is a multiple of the denominator.

Other item families might examine the same mathematical idea in different content areas. For example, a family of items might be designed to see how students use proportional thinking in different mathematical contexts such as geometry, algebra, and measurement.

EXAMPLE ITEM FAMILY 1 ILLUSTRATING DEPTH Source: 2003 NAEP 8M6 #10 modified
Grade 8
Measurement: Measuring physical attributes
NOTE: ITEMS WOULD NOT APPEAR TOGETHER, BUT WOULD BE DISTRIBUTED AMONG BLOCKS.

1. What is the length of a side of a square whose area is 36 square yards?

A. 4 yards
B. 6 yards
C. 8 yards
D. 10 yards
E. 12 yards

Correct Answer: B

2. A rectangle that is 2 yards by 18 yards has the same area as a square. What is the length of a side of the square?

A. 4 yards
B. 6 yards
C. 8 yards
D. 10 yards
E. 12 yards

Correct Answer: B

3a.

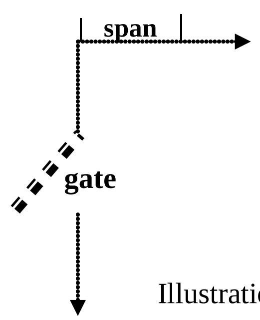


Illustration: Corner of fence showing **span** and **gate.**

Jill had the following requirements for a fence she is building for her dog.

1. The area is 36 square yards.
2. The pen is a square.
3. Fencing comes in 2-yard **spans**.
4. There must be a 2-yards wide gate not made from fencing.

How many **spans** of fencing does she need?

Correct Answer: 11

3b. Use requirements 2-4 given in 3a. Write an equation that expresses the number of spans [use *s*] needed for any area [use *A*]

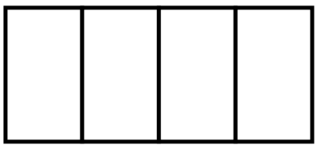
Correct Answer: $s = 2\sqrt{A} - 1$ or equivalent.

EXAMPLE ITEM FAMILY 2 ILLUSTRATING BREADTH Source: Framework authors
Grade 4

Number: Number Sense

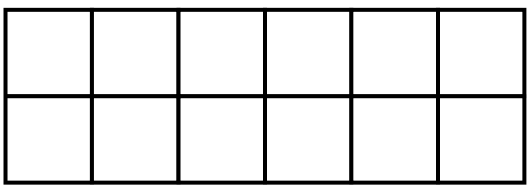
NOTE: ITEMS WOULD NOT APPEAR TOGETHER BUT WOULD BE DISTRIBUTED AMONG BLOCKS.

1. Shade ¾ of the large rectangle.



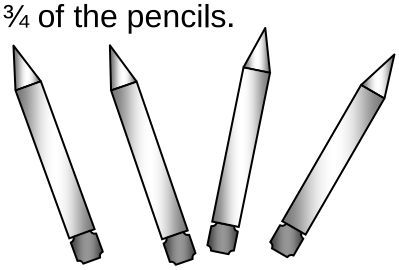
Correct Answer: 3 of the smaller rectangles shaded

2. Shade ¾ of the large rectangle.



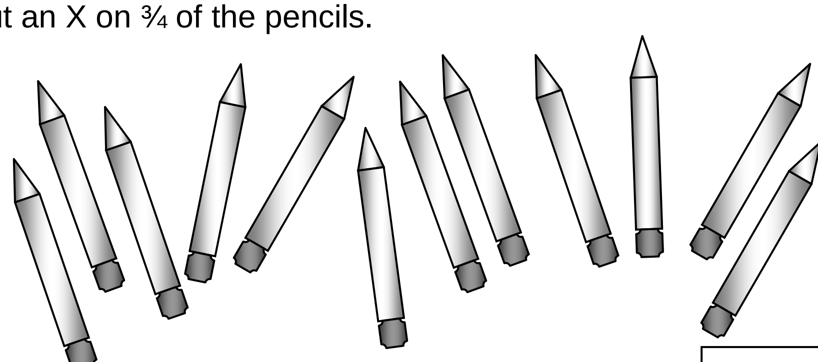
Correct Answer: 9 of the squares shaded

3. Put an X on ¾ of the pencils.



Correct Answer: 3 pencils marked with X

4. Put an X on ¾ of the pencils.



Correct Answer: 9 pencils marked with X

REFERENCES

Abedi, J. (2006). Language Issues in Item-Development. In Downing, S. M. and Haladyna, T. M. (Ed.) *Handbook of Test Development*. New Jersey: Lawrence Erlbaum Associates, Publishers.

Abedi, J. (2004). *Inclusion of students with limited English proficiency in NAEP: Classification and measurement issues* (CSE Report 629). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing, Center for the Study of Evaluation, Graduate School of Education and Information Studies.

Kopriva, R.J., & Carr, T.G. (2006). STELLA: A computerized system for individually assigning test accommodations to ELLs. Presentation at the 2006 National Conference on Large-Scale Assessment, San Francisco, CA.

Kopriva, R.J. (in press). *Improving testing for English language learners: A comprehensive approach to designing, building, implementing, and interpreting better academic assessments.* NJ. Lawrence Erlbaum Publishers.

Kopriva, R., & Saez, S. (1997). *Guide to scoring LEP student responses to open-ended mathematics items*. Washington, DC: Council of Chief State School Officers.

Koretz, D., & Barton, K. (2003-2004). Assessing students with disabilities: Issues and evidence. *Educational Assessment, 9*(1 & 2), 29-60.

National Center for Education Statistics. (2007). NAEP inclusion policy. Retrieved June 6, 2007, from http://nces.ed.gov/nationsreportcard/about/inclusion.asp

National Center for Education Statistics. (2006). NAEP assessment administrator manual. Retrieved June 6, 2007, from http://nces.ed.gov/nationsreportcard/about/inclusion.asp

National Research Council (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments*. Washington, DC: National Academy Press.

National Research Council. (1999). *High stakes: Testing for tracking, promotion, and graduation.* Committee on Appropriate Test Use. J. P. Heubert & R. M. Hauser (Eds.). Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Perie, M., Grigg, W., & Dion, G. (2005). *The nation's report card: Mathematics 2005* (NCES 2006-453). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office.

Solano-Flores, G., Lara., J., Sexton, U., & Navarrete, C. (2001). *Testing English language learners: A sampler of student responses to science and mathematics test items*. Washington, DC: Council of Chief State School Officers.

South Carolina Department of Education (2006). Adding value to accommodations decision-making for English language learners and students with Disabilities (AVAD). Proposal for the US Department of Education Enhanced Assessment Grant Program.

Thompson, S.J., Johnstone, C.J., Thurlow, M. L. (2002). Universal design applied to large-scale assessments. Minneapolis, MN: National Center on Educational Outcomes.

Webb, N.L., Ely, T., Cormier, M., & Vesperman B., (2006). The Web Alignment Tool: Development, refinement, and dissemination. In Council of Chief State School Officers, *Aligning assessment to guide the learning of all students* (pp. 1-30).. Washington, DC: Council of Chief State School Officers.

APPENDIX A

NAEP MATHEMATICS ACHIEVEMENT LEVEL DESCRIPTIONS

The achievement levels are cumulative; therefore, students performing at the *Proficient* level also display the competencies associated with the *Basic* level, and students at the *Advanced* level also demonstrate the skills and knowledge associated with both the *Basic* and the *Proficient* levels. The cut score indicating the lower end of the score range for each level is noted in parentheses.

NAEP MATHEMATICS ACHIEVEMENT LEVELS—GRADE 4

Basic (214)

Fourth-grade students performing at the *Basic* level should show some evidence of understanding the mathematical concepts and procedures in the five NAEP content areas.

Fourth graders performing at the *Basic* level should be able to estimate and use basic facts to perform simple computations with whole numbers; show some understanding of fractions and decimals; and solve some simple real-world problems in all NAEP content areas. Students at this level should be able to use - though not always accurately - four-function calculators, rulers, and geometric shapes. Their written responses are often minimal and presented without supporting information.

Proficient (249)

Fourth-grade students performing at the *Proficient* level should consistently apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP content areas.

Fourth graders performing at the *Proficient* level should be able to use whole numbers to estimate, compute, and determine whether results are reasonable. They should have a conceptual understanding of fractions and decimals; be able to solve real-world problems in all NAEP content areas; and use four-function calculators, rulers, and geometric shapes appropriately. Students performing at the *Proficient* level should employ problem-solving strategies such as identifying and using appropriate information. Their written solutions should be organized and presented both with supporting information and explanations of how they were achieved.

Advanced (282)

Fourth-grade students performing at the *Advanced* level should apply integrated procedural knowledge and conceptual understanding to complex and nonroutine real-world problem solving in the five NAEP content areas.

Fourth graders performing at the *Advanced* level should be able to solve complex nonroutine real-world problems in all NAEP content areas. They should display mastery in the use of four-function calculators, rulers, and geometric shapes. These students are expected to draw logical conclusions and justify answers and solution processes by explaining why, as well as how, they were achieved. They should go beyond the obvious in their interpretations and be able to communicate their thoughts clearly and concisely.

Basic (262)

Eighth-grade students performing at the *Basic* level should exhibit evidence of conceptual and procedural understanding in the five NAEP content areas. This level of performance signifies an understanding of arithmetic operations - including estimation - on whole numbers, decimals, fractions, and percents.

Eighth graders performing at the *Basic* level should complete problems correctly with the help of structural prompts such as diagrams, charts, and graphs. They should be able to solve problems in all NAEP content areas through the appropriate selection and use of strategies and technological tools - including calculators, computers, and geometric shapes. Students at this level also should be able to use fundamental algebraic and informal geometric concepts in problem solving.

As they approach the *Proficient* level, students at the *Basic* level should be able to determine which of the available data are necessary and sufficient for correct solutions and use them in problem solving. However, these eighth graders show limited skill in communicating mathematically.

Proficient (299)

Eighth-grade students performing at the *Proficient* level should apply mathematical concepts and procedures consistently to complex problems in the five NAEP content areas.

Eighth graders performing at the *Proficien*t level should be able to conjecture, defend their ideas, and give supporting examples. They should understand the connections among fractions, percents, decimals, and other mathematical topics such as algebra and functions. Students at this level are expected to have a thorough understanding of *Basic* level arithmetic operations - an understanding sufficient for problem solving in practical situations.

Quantity and spatial relationships in problem solving and reasoning should be familiar to them, and they should be able to convey underlying reasoning skills beyond the level of arithmetic. They should be able to compare and contrast mathematical ideas and generate their own examples. These students should make inferences from data and graphs; apply properties of informal geometry; and accurately use the tools of technology. Students at this level should understand the process of gathering and organizing data and be able to calculate, evaluate, and communicate results within the domain of statistics and probability.

Advanced (333)

Eighth-grade students performing at the *Advanced* level should be able to reach beyond the recognition, identification, and application of mathematical rules in order to generalize and synthesize concepts and principles in the five NAEP content areas.

Eighth graders performing at the *Advanced* level should be able to probe examples and counterexamples in order to shape generalizations from which they can develop models. Eighth graders performing at the *Advanced* level should use number sense and geometric awareness to consider the reasonableness of an answer. They are expected to use abstract thinking to create unique problem-solving techniques and explain the reasoning processes underlying their conclusions.

Basic (141)

Twelfth-grade students performing at the *Basic* level should be able to solve mathematical problems that require the direct application of concepts and procedures in familiar situations.

Students at grade 12 should be able to perform computations with real numbers and estimate the results of numerical calculations. These students should also be able to estimate, calculate, and compare measures and identify and compare properties of two- and three-dimensional figures, and solve simple problems using two-dimensional coordinate geometry.

At this level, students should be able to identify the source of bias in a sample and make inferences from sample results; calculate, interpret, and use measures of central tendency; and compute simple probabilities. They should understand the use of variables, expressions, and equations to represent unknown quantities and relationships among unknown quantities. They should be able to solve problems involving linear relations using tables, graphics, or symbols, and solve linear equations involving one variable.

Proficient (176)

Twelfth-grade students performing at the *Proficient* level should be able to select strategies to solve problems and integrate concepts and procedures.

These students should be able to interpret an argument, justify a mathematical process, and make comparisons dealing with a wide variety of mathematical tasks. They should also be able to perform calculations involving similar figures including right triangle trigonometry. They should understand and apply properties of geometric figures and relationships between figures in two and three dimensions.

Students at this level should select and use appropriate units of measure as they apply formulas to solve problems. Students performing at this level should be able to use measures of central tendency and variability of distributions to make decisions and predictions, calculate combinations and permutations to solve problems, and understand the use of the normal distribution to describe real-world situations. Students performing at the *Proficient* level should be able to identify, manipulate, graph, and apply linear, quadratic, exponential, and inverse functions ($y = k/x$); solve routine and non-routine problems involving functions expressed in algebraic, verbal, tabular, and graphical forms; and solve quadratic and rational equations in one variable and solve systems of linear equations.

Advanced (216)

Twelfth-grade students performing at the *Advanced* level should demonstrate in-depth knowledge of the mathematical concepts and procedures represented in the framework.

Students should be able to integrate knowledge to solve complex problems and justify and explain their thinking. These students should be able to analyze, make and justify mathematical arguments, and communicate their ideas clearly. *Advanced* level students should be able to describe the intersections of geometric figures in two and three dimensions, and use vectors to represent velocity and direction. They should also be able to describe the impact of linear transformations and outliers on measures of central tendency and variability, analyze predictions based on multiple data sets, and apply probability and statistical reasoning in more complex problems. Students performing at the *Advanced* level should be able to solve or interpret systems of inequalities and formulate a model for a complex situation (e.g., exponential growth and decay) and make inferences or predictions using the mathematical model.

APPENDIX B

NAEP ITEM DEVELOPMENT AND REVIEW POLICY STATEMENT

Adopted: May 18, 2002

It is the policy of the National Assessment Governing Board to require the highest standards of fairness, accuracy, and technical quality in the design, construction, and final approval of all test questions and assessments developed and administered under the National Assessment of Educational Progress (NAEP). All NAEP test questions or items must be designed and constructed to reflect carefully the assessment objectives approved by the National Assessment Governing Board. The final assessments shall adhere to the requirements outlined in the following Guiding Principles, Policies and Procedures for NAEP Item Development and Review.

The Governing Board's Assessment Development Committee, with assistance from other Board members as needed, shall be responsible for reviewing and approving NAEP test questions at several stages during the development cycle. In so doing, the Guiding Principles, Policies and Procedures must be adhered to rigorously.

INTRODUCTION

The No Child Left Behind Act of 2001 (P.L. 107-110) contains a number of important provisions regarding item development and review for the National Assessment of Educational Progress (NAEP). The legislation requires that:

- "the purpose [of NAEP] is to provide…a fair and accurate measurement of student academic achievement"
- "[NAEP shall]…use widely accepted professional testing standards, objectively measure academic achievement, knowledge, and skills, and ensure that any academic assessment authorized….be tests that do not evaluate or assess personal or family beliefs and attitudes or publicly disclose personally identifiable information;"
- "[NAEP shall]…only collect information that is directly related to the appraisal of academic achievement, and to the fair and accurate presentation of such information;"
- "the Board shall develop assessment objectives consistent with the requirements of this section and test specifications that produce an assessment that is valid and reliable, and are based on relevant widely accepted professional standards;"
- "the Board shall have final authority on the appropriateness of all assessment items;"
- "the Board shall take steps to ensure that all items selected for use in the National Assessment are free from racial, cultural, gender, or regional bias and are secular, neutral, and non-ideological;" and

- “the Board shall develop a process for review of the assessment which includes the active participation of teachers, curriculum specialists, local school administrators, parents, and concerned members of the public.”

Given the importance of these mandates, it is incumbent upon the Board to ensure that the highest standards of test fairness and technical quality are employed in the design, construction, and final approval of all test questions for the National Assessment. The validity of educational inferences made using NAEP data could be seriously impaired without high standards and rigorous procedures for test item development, review, and selection.

Test questions used in the National Assessment must yield assessment data that are both valid and reliable in order to be appropriate. Consequently, technical acceptability is a necessary, but not a sufficient condition, for judging the appropriateness of items. In addition, the process for item development must be thorough and accurate, with sufficient reviews and checkpoints to ensure that accuracy. The Guiding Principles, Policies, and Procedures governing item development, if fully implemented throughout the development cycle, will result in items that are fair and of the highest technical quality, and which will yield valid and reliable assessment data.

Each of the following Guiding Principles is accompanied by Policies and Procedures. Full implementation of this policy will require supporting documentation from the National Center for Education Statistics (NCES) regarding all aspects of the Policies and Procedures for which they are responsible.

This policy complies with the documents listed below which express acceptable technical and professional standards for item development and use. These standards reflect the current agreement of recognized experts in the field, as well as the policy positions of major professional and technical associations concerned with educational testing.

Standards for educational and psychological testing. (1999). Washington, DC: American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME).

Code of fair testing practices in education. (1988). Washington, DC: Joint Committee on Testing Practices.

National Center for Education Statistics (NCES) Statistical Standards, DRAFT, February 2002.

GUIDING PRINCIPLES – ITEM DEVELOPMENT AND REVIEW POLICY

Principle 1
NAEP test questions selected for a given content area shall be representative of the content domain to which inferences will be made and shall match the NAEP assessment framework and specifications for a particular assessment.

Principle 2
The achievement level descriptions for basic, proficient, and advanced performance shall be an important consideration in all phases of NAEP development and review.

Principle 3
The Governing Board shall have final authority over all NAEP test questions. This authority includes, but is not limited to, the development of items, establishing the criteria for reviewing items, and the process for review.

Principle 4
The Governing Board shall review all NAEP test questions that are to be administered in conjunction with a pilot test, field test, operational assessment, or special study administered as part of NAEP.

Principle 5
NAEP test questions will be accurate in their presentation and free from error. Scoring criteria will be accurate, clear, and explicit.

Principle 6
All NAEP test questions will be free from racial, cultural, gender, or regional bias, and must be secular, neutral, and non-ideological. NAEP will not evaluate or assess personal or family beliefs, feelings, and attitudes, or publicly disclose personally identifiable information.

POLICIES AND PROCEDURES FOR GUIDING PRINCIPLES

Principle 1

NAEP test questions selected for a given content area shall be representative of the content domain to which inferences will be made and shall match the NAEP assessment framework and specifications for a particular assessment.

Policies and Procedures

1. Under the direction of the Board, the framework for each assessment will be developed in a manner that defines the content to be assessed, consistent with NAEP's purpose and the context of a large-scale assessment. The framework

development process shall result in a rationale for each NAEP assessment, which delineates the scope of the assessment relative to the content domain. The framework will consist of a statement of purpose, assessment objectives, format requirements, and other guidelines for developing the assessment and items.

2. In addition to the framework, the Board shall develop assessment and item specifications to define the: a) content and process dimensions for the assessment; b) distribution of items across content and process dimensions at each grade level; c) stimulus and response attributes (or what the test question provides to students and the format for answering the item); d) types of scoring procedures; e) test administration conditions; and f) other specifications pertaining to the particular subject area assessment.
3. The Board will forward the framework and specifications to NCES, in accordance with an appropriate timeline, so that NCES may carry out its responsibilities for assessment development and administration.
4. In order to ensure that valid inferences can be made from the assessment, it is critical that the pool of test questions measures the construct as defined in the framework. Demonstrating that the items selected for the assessment are representative of the subject matter to which inferences will be made is a major type of validity evidence needed to establish the appropriateness of items.
5. A second type of validity evidence is needed to ensure that NAEP test items match the specific objectives of a given assessment. The items must reflect the objectives, and the item pool must match the percentage distribution for the content and cognitive dimensions at each grade level, as stated in the framework. Minor deviations, if any, from the content domain as defined by the framework will be explained in supporting materials.
6. Supporting material submitted with the NAEP items will provide a description of procedures followed by item writers during development of NAEP test questions. This description will include the expertise, training, and demographic characteristics of the groups. This supporting material must show that all item writing and review groups have the required expertise and training in the subject matter, bias, fairness, and assessment development.
7. In submitting items for review by the Board, NCES will provide information on the relationship of the specifications and the content/process elements of the pool of NAEP items. This will include procedures used in classifying each item.
8. The item types used in an assessment must match the content requirements as stated in the framework and specifications, to the extent possible. The match between an objective and the item format must be informed by specifications pertaining to the content, knowledge or skill to be measured, cognitive complexity, overall appropriateness, and efficiency of the item type. NAEP assessments shall use a variety of item types as best fit the requirements stated in the framework and specifications.
9. In order to ensure consistency between the framework and specifications documents and the item pools, NCES will ensure that the development contractor engages a minimum of 20% of the membership of the framework project committees in each subject area to serve on the item writing and review groups as the NAEP test questions are being developed. This overlap between the

framework development committees and the item developers will provide stability throughout the NAEP development process, and ensure that the framework and specifications approved by the Board have been faithfully executed in developing NAEP test questions.

Principle 2

The achievement level descriptions for basic, proficient, and advanced performance shall be an important consideration in all phases of NAEP development and review.

Policies and Procedures

1. During the framework development process, the project committees shall draft preliminary descriptions of the achievement levels for each grade to be assessed. These preliminary descriptions will define what students should know and be able to do at each grade, in terms of the content and process dimensions of the framework at the basic, proficient, and advanced levels. Subsequent to Board adoption, the final achievement level descriptions shall be an important consideration in all future test item development for a given subject area framework.
2. The achievement level descriptions will be used to ensure a match between the descriptions and the resulting NAEP items. The achievement level descriptions will be examined, and appropriate instruction provided to item writers to ensure that the items represent the stated descriptions, while adhering to the content and process requirements of the framework and specifications. The descriptions will be used to evaluate the test questions to make certain that the pool of questions encompasses the range of content and process demands specified in the achievement level descriptions, including items within each achievement level interval, and items that scale below basic.
3. As the NAEP item pool is being constructed, additional questions may need to be written for certain content/skill areas if there appear to be any gaps in the pool, relative to the achievement level descriptions.
4. Supporting materials will show the relationship between the achievement levels descriptions and the pool of NAEP test questions.

Principle 3

The Governing Board shall have final authority over all NAEP test questions. This authority includes, but is not limited to, the development of items, establishing the criteria for reviewing items, and the process for review.

Policies and Procedures

1. Under the No Child Left Behind Act, a primary duty of the Governing Board pertains to “All Cognitive and Noncognitive Assessment Items.” Specifically, the statute states that, “The Board shall have final authority on the appropriateness of

all assessment items.” Under the law, the Board is therefore responsible for all NAEP test questions as well as all NAEP background questions administered as part of the assessment.

2. To meet this statutory requirement, the Board’s Policy on NAEP Item Development and Review shall be adhered to during all phases of NAEP item writing, reviewing, editing, and assessment construction. The National Center for Education Statistic (NCES), which oversees the operational aspects of NAEP, shall ensure that all internal and external groups involved in NAEP item development activities follow the Guiding Principles, Policies and Procedures as set forth in this Board policy.
3. Final review of all NAEP test questions for bias and appropriateness shall be performed by the Board, after all other review procedures have been completed, and prior to administration of the items to students.

Principle 4

The Governing Board shall review all NAEP test questions that are to be administered in conjunction with a pilot test, field test, operational assessment, or special study administered as part of NAEP.

Policies and Procedures

1. To fulfill its statutory responsibility for NAEP item review, the Board shall receive, in a timely manner and with appropriate documentation, all test questions that will be administered to students under the auspices of a NAEP assessment. These items include those slated for pilot testing, field testing, and operational administration.
2. The Board shall review all test items developed for special studies, where the purpose of the special study is to investigate alternate item formats or new technologies for possible future inclusion as part of main NAEP, or as part of a special study to augment main NAEP data collection.
3. The Board shall <u>not</u> review items being administered as part of test development activities, such as small-scale, informal try-outs with limited groups of students designed to refine items prior to large-scale pilot, field, or operational assessment.
4. NCES shall submit NAEP items to the Board for review in accordance with a mutually agreeable timeline. Items will be accompanied by appropriate documentation as required in this policy. Such information shall consist of procedures and personnel involved in item development and review, the match between the item pool and the framework content and process dimensions, and other related information.
5. For its first review, the Board will examine all items prior to the pilot test or field test stage. In the case of the NAEP reading assessment, all reading passages will be reviewed by the Board prior to item development. For each reading passage, NCES will provide the source, author, publication date, passage length, rationale for minor editing to the passage (if any), and notation of such editing applied to

the original passage. NCES will provide information and explanatory material on passages deleted in its fairness review procedures.

6. For its second review, the Board will examine items following pilot or field testing. The items will be accompanied by statistics obtained during the pilot test or field test stage. These statistics shall be provided in a clear format, with definitions for each item analysis statistic collected. Such statistics shall include, but shall not be limited to: p-values for multiple-choice items, number and percentage of students selecting each option for a multiple-choice item, number and percentage not reaching or omitting the item (for multiple-choice and open-ended), number and percentage of students receiving various score points for open-ended questions, mean score point value for open-ended items, appropriate biserial statistics, and other relevant data.
7. At a third stage, for some assessments, the Board will receive a report from the calibration field test stage, which occurs prior to the operational administration. This "exceptions report" will contain information pertaining to any items that were dropped due to differential item functioning (DIF) analysis for bias, other items to be deleted from the operational assessment and the rationale for this decision, and the final match between the framework distribution and the item pool. If the technology becomes available to perform statistically sound item-level substitutions at this point in the cycle (from the initial field test pool), the Board shall be informed of this process as well.
8. All NAEP test items will be reviewed by the Board in a secure manner via in-person meetings, teleconference or videoconference settings, or on-line via a password-protected Internet site. The Board's Assessment Development Committee shall have primary responsibility for item review and approval. However, the Assessment Development Committee, in consultation with the Board Chair, may involve other NAGB members in the item review process on an *ad hoc* basis. The Board may also submit items to external experts, identified by the Board for their subject area expertise, to assist in various duties related to item review. Such experts will follow strict procedures to maintain item security, including signing a Nondisclosure Agreement.
9. Items that are edited between assessments by NCES and/or its item review committees, for potential use in a subsequent assessment, shall be re-examined by the Board prior to a second round of pilot or field testing.
10. Documentation of the Board's final written decision on editing and deleting NAEP items shall be provided to NCES within 10 business days following completion of Board review at each stage in the process.

Principle 5

NAEP test questions will be accurate in their presentation, and free from error. Scoring criteria will be accurate, clear, and explicit.

Policies and Procedures

1. NCES, through its subject area content experts, trained item writers, and item review panels, will examine each item carefully to ensure its accuracy. All materials taken from published sources must be carefully documented by the item writer. Graphics that accompany test items must be clear, correctly labeled, and include the data source where appropriate. Items will be clear, grammatically correct, succinct, and unambiguous, using language appropriate to the grade level being assessed. Item writers will adhere to the specifications document regarding appropriate and inappropriate stimulus materials, terminology, answer choices or distractors, and other requirements for a given subject area. Items will not contain extraneous or irrelevant information that may differentially distract or disadvantage various subgroups of students from the main task of the item.
2. Scoring criteria will accompany each constructed-response item. Such criteria will be clear, accurate, and explicit. Carefully constructed scoring criteria will ensure valid and reliable use of those criteria to evaluate student responses to maximize the accuracy and efficiency of scoring.
3. Constructed-response scoring criteria will be developed initially by the item writers, refined during item review, and finalized during pilot or field test scoring. During pilot or field test scoring, the scoring guides will be expanded to include examples of actual student responses to illustrate each score point. Actual student responses will be used as well, to inform scorers of unacceptable answers.
4. Procedures used to train scorers and to conduct scoring of constructed-response items must be provided to the Board, along with information regarding the reliability and validity of such scoring. If the technology becomes available to score student responses electronically, the Board must be informed of the reliability and validity of such scoring protocol, as compared to human scoring.

Principle 6

All NAEP test questions will be free from racial, cultural, gender, or regional bias, and must be secular, neutral, and non-ideological. NAEP will not evaluate or assess personal or family beliefs, feelings, and attitudes, or publicly disclose personally identifiable information.

Policies and Procedures

1. An item is considered biased if it unfairly disadvantages a particular subgroup of students by requiring knowledge of obscure information unrelated to the construct being assessed. A test question or passage is biased if it contains material derisive or derogatory toward a particular group. For example, a geometry item requiring prior knowledge of the specific dimensions of a basketball court would result in lower scores for students unfamiliar with that sport, even if those students know the geometric concept being measured. Use of a regional term for a soft drink in an item context may provide an unfair advantage to students from that area of the country. Also, an item that refers to a low-achieving student as “slow” would be

unacceptable.

2. In conducting bias reviews, steps should be taken to rid the item pool of questions that, because of their content or format, either appear biased on their face, or yield biased estimates of performance for certain subpopulations based on gender, race, ethnicity, or regional culture. A statistical finding of differential item functioning (DIF) will result in a review aimed at identifying possible explanations for the finding. However, such an item will not automatically be deleted if it is deemed valid for measuring what was intended, based on the NAEP assessment framework. Items in which clear bias is found will be eliminated. This policy acknowledges that there may be real and substantial differences in performance among subgroups of students. Learning about such differences, so that performance may be improved, is part of the value of the National Assessment.
3. Items shall be secular, neutral, and non-ideological. Neither NAEP nor its questions shall advocate a particular religious belief or political stance. Where appropriate, NAEP questions may deal with religious and political issues in a fair and objective way. The following definitions shall apply to the review of all NAEP test questions, reading passages, and supplementary materials used in the assessment of various subject areas:
    - <u>Secular</u> **–** NAEP questions will not contain language that advocates or opposes any particular religious views or beliefs, nor will items compare one religion unfavorably to another. However, items may contain references to religions, religious symbolism, or members of religious groups where appropriate. Examples: The following phrases would be acceptable: “shaped like a Christmas tree”, “religious tolerance is one of the key aspects of a free society,” “Dr. Martin Luther King, Jr. was a Baptist minister,” or “Hinduism is the predominant religion in India.”
    - <u>Neutral and Non-ideological</u> - Items will not advocate for a particular political party or partisan issue, for any specific legislative or electoral result, or for a single perspective on a controversial issue. An item may ask students to explain both sides of a debate, or it may ask them to analyze an issue, or to explain the arguments of proponents or opponents, without requiring students to endorse personally the position they are describing. Item writers should have the flexibility to develop questions that measure important knowledge and skills without requiring both pro and con responses to every item. Examples: Students may be asked to compare and contrast positions on states rights, based on excerpts from speeches by X and Y; to analyze the themes of Franklin D. Roosevelt’s first and second inaugural addresses; to identify the purpose of the Monroe Doctrine; or to select a position on the issue of suburban growth and cite evidence to support this position. Or, students may be asked to provide arguments either for or against Woodrow Wilson’s decision to enter World War I. A NAEP question could ask students to summarize the dissenting opinion in a landmark Supreme Court case. The criteria of neutral and non-ideological also pertain to decisions about the pool of test questions in a subject area, taken as a whole. The Board shall review the

entire item pool for a subject area to ensure that it is balanced in terms of the perspectives and issues presented.

4. The Board shall review both stimulus materials and test items to ensure adherence to the NAEP statute and the polices in this statement. Stimulus materials include reading passages, articles, documents, graphs, maps, photographs, quotations, and all other information provided to students in a NAEP test question.
5. NAEP questions will not ask a student to reveal personal or family beliefs, feelings, or attitudes, or publicly disclose personally identifiable information.

APPENDIX C

NAEP MATHEMATICS PROJECT STAFF AND COMMITTEES

Members of NAGB's Grade 12 Mathematics Panel

Herbert Clemens
Professor, Department of Mathematics
Ohio State University
Columbus, Ohio

Mary Ann Huntley
Assistant Professor, Mathematics
Department of Mathematical Sciences
University of Delaware
Newark, Delaware

Jeremy Kilpatrick
Regents Professor
University of Georgia
Athens, Georgia

Mary Lindquist
Fuller E. Callaway Professor, Emeritus
Columbus State University
Lewisburg, West Virginia

Mary Jo Messenger
Chair, Department of Mathematics (retired)
River Hill High School
Clarksville, Maryland

William Schmidt
University Distinguished Professor
Michigan State University
East Lansing, Michigan

Sharif Shakrani
Director, Education Policy Research Center
Michigan State University
East Lansing, Michigan

Linda Dager Wilson, Chair
Mathematics Consultant
Washington, D.C.

Achieve NAEP Grade 12 Mathematics Panel

Sue Eddins
Mathematics Teacher (retired)
Illinois Mathematics and Science Academy
Aurora, Illinois

William McCallum
University Distinguished Professor of Mathematics
Department of Mathematics
University of Arizona
Tucson, Arizona

Fabio Milner
Professor of Mathematics
Purdue University
West Lafayette, Indiana

William Schmidt
University Distinguished Professor
Michigan State University
East Lansing, Michigan

Lynn Steen
Professor of Mathematics
St. Olaf College
Northfield, Minnesota

Norman Webb
Senior Research Scientist
Wisconsin Center for Education Research
University of Wisconsin
Madison, Wisconsin

Reviews Received on the Draft of NAEP 12th Grade Mathematics Objectives

Achieve, Inc.
American Mathematical Society
Association of State Supervisors of Mathematics
Thomas B. Fordham Institute
State Mathematics Supervisors from various states
National Council of Teachers of Mathematics
State Testing Directors from various states

2009 NAEP Mathematics Specifications Working Group

Mary Lindquist
Fuller E. Callaway Professor, Emeritus
Columbus State University
Lewisburg, West Virginia

Mary Jo Messenger
Chair, Department of Mathematics (retired)
River Hill High School
Clarksville, Maryland

Linda Dager Wilson, Chair
Mathematics Consultant
Washington, D.C.

Phoebe C. Winter, Editor
Measurement Consultant
Richmond, Virginia

2005 NAEP Mathematics Project Steering Committee

Eileen Ahearn
Project Director
National Association of State Directors of Special Education
Alexandria, Virginia

Charles Allan
Mathematics Education Consultant
Michigan Department of Education
Lansing, Michigan

B. Marie Byers
National School Boards Association
Hagerstown, Maryland

Randy DeHoff
Colorado State Board of Education
6th Congressional District–Littleton
Denver, Colorado

M.B. “Sonny” Donaldson
Superintendent
Aldine ISD
Houston, Texas

Janice Earle
Senior Program Director
National Science Foundation
Arlington, Virginia

Lou Fabrizio
Director
Division of Accountability Services
North Carolina Department of Public Instruction
Raleigh, North Carolina

Bettye Forte
Mathematics Consultant
Arlington, Texas

Matt Gandal
Vice President
Achieve, Inc.
Washington, D.C.

Alice Gill
Associate Director
Educational Issues
American Federation of Teachers
Washington, D.C.

M. Kathleen Heid
The Pennsylvania State University
University Park, Pennsylvania

Audrey Jackson
Assistant Principal
Claymont Elementary School
Parkway City Schools
Fenton, Missouri

James M. Landwehr
Director
Data Analysis Research Department
Avaya Labs
Basking Ridge, New Jersey

Sharon Lewis
Research Director
Council of the Great City Schools
Washington, D.C.

Dane Linn
Policy Studies Director
National Governors’ Association
Washington, D.C.

Eddie Lucero
Principal
Griegos Elementary School
Albuquerque, New Mexico

Lee McCaskill
Principal
Brooklyn Technical High School
Brooklyn, New York

Barbara Montalto
Assistant Director of Mathematics
Texas Education Agency
Austin, Texas

Judy Rohde
Mathematics Instructor
John Glenn Middle School
Maplewood, Minnesota

Wilfried Schmid
Professor of Mathematics
Harvard Department of Mathematics
Cambridge, Massachusetts

Sr. Mary Frances Taymans
Associate Executive Director
Secondary Schools Department
National Catholic Education Association
Washington, D.C.

Zalman Usiskin
Professor of Education
Director, University of Chicago School Mathematics Project
Chicago, Illinois

Judy Walter
Association for Supervision & Curriculum Development
Alexandria, Virginia

Diana Wearne
Associate Professor
School of Education
University of Delaware
Newark, Delaware

Hung-Hsi Wu
Professor of Mathematics
Department of Mathematics
University of California–Berkeley
Berkeley, California

2005 NAEP Mathematics Project Planning Committee

Dayo Akinsheye
Mathematics Resource Teacher
Seaton Elementary School
Washington, D.C.

Geri Anderson-Nielsen
Mathematics Specialist
Georgetown Day School
Washington, D.C.

Cindy Chapman
Elementary Teacher
Albuquerque Public Schools
Albuquerque, New Mexico

Herbert Clemens
Professor of Mathematics
Department of Mathematics
University of Utah
Salt Lake City, Utah

Carl Cowen
Professor of Mathematics
Purdue University
West Lafayette, Indiana

Jim Ellingson
Assistant Professor
Concordia College
Moorhead, Minnesota

Joan Ferrini-Mundy
Associate Dean/Director of Science and Mathematics
College of Natural Science
Michigan State University
East Lansing, Michigan

Kim Gattis
Education Program Consultant
Kansas Department of Education
Association of State Supervisors of Mathematics
Topeka, Kansas

Anne Gonzales
Middle School Mathematics Teacher
South Gate Middle School
South Gate, California

Jeremy Kilpatrick
Professor of Mathematics Education
University of Georgia
Athens, Georgia

Gerald Kulm
Curtis D. Robert Professor
of Mathematics Education
Texas A & M University
College Station, Texas

Mary Lindquist
Fuller E. Callaway Professor of Mathematics Education
Columbus State University
Columbus, Georgia

Mary Jo Messenger
Chair, Department of Mathematics
River Hill High School
Clarksville, Maryland

Marjorie Petit
Senior Associate
National Center for the Improvement of Educational Assessment
(The Center for Assessment)
Portsmouth, New Hampshire

Edward Silver
Professor
School of Education
University of Michigan
Ann Arbor, Michigan

Debra Vitale
Mathematics Specialist
Arlington Public Schools
Fairfax, Virginia

Frank Wang
President/CEO
Saxon Publishing, Inc.
Norman, Oklahoma

Norman Webb
Senior Research Scientist
Wisconsin Center for Education Research
Madison, Wisconsin

John Wisthoff
Member, Maryland State Board of
Education and Mathematics Professor
Anne Arundel Community College
Pasadena, Maryland

2005 NAEP Mathematics Project Technical Advisory Panel

Fen Chou
Psychometrician
Louisiana Department of Education
Baton Rouge, Louisiana

Eugene Johnson
Chief Psychometrician
American Institutes for Research
Washington, D.C.

Edward Kifer
Professor and Chairperson
Department of Educational Policy Studies
and Evaluation
College of Education
University of Kentucky
Lexington, Kentucky

Ina Mullis
Co-Director
International Study Center
Boston College
Chestnut Hill, Massachusetts

Barbara Plake
Director
Buros Center for Testing
University of Nebraska–Lincoln
Lincoln, Nebraska

Roger Trent
Ohio State Assessment Director (Emeritus)
Ohio Department of Education
Columbus, Ohio

CCSSO Staff

Rolf Blank
Director of Indicators Project
State Education Assessment Center
Council of Chief State School Officers
Washington, D.C.

Wayne Martin
Director
State Education Assessment Center
Council of Chief State School Officers
Washington, D.C.

John Olson
Director of Assessment
State Education Assessment Center
Council of Chief State School Officers
Washington, D.C.

Frank Philip
Senior Project Associate
State Education Assessment Center
Council of Chief State School Officers
Washington, D.C.

Linda Dager Wilson
Consensus Coordinator Consultant
Council of Chief State School Officers
Washington, D.C.

Phoebe Winter
Project Director
State Education Assessment Center
Council of Chief State School Officers
Richmond, Virginia

Subcontractors and Consultants

Patricia Kenney
Senior Research Associate
University of Michigan
Ann Arbor, Michigan

Rebecca Kopriva
Director
Center for the Study of Assessment Validity in Education
Department of Measurement & Statistics
University of Maryland
College Park, Maryland

Christopher Cross
President (Former)
Council for Basic Education
Washington, DC

Kim Gattis
President, Association of State Supervisors of Mathematics
Education Program Consultant, Kansas Department of Education
Topeka, Kansas

Linda Plattner
Director of Policy, Standards & Instruction
Council for Basic Education
Washington, DC

Committee Representation

Policy Organizations

Achieve, Inc.
American Association of School Administrators (AASA)
American Federation of Teachers (AFT)
American Mathematical Society (AMA)
American Statistical Association (ASA)
Association for Supervision and Curriculum Development (ASCD)
Association of State Assessment Programs (ASAP)
Association of State Supervisors of Mathematics (ASSM)
Business Roundtable/National Alliance of Business
Council of the Great City Schools
Education Leaders Council (ELC)
National Association of Elementary School Principals (NAESP)
National Association of Secondary School Principals (NASSP)
National Association of State Boards of Education (NASBE)
National Association of State Directors of Special Education (NASDE)
National Catholic Education Association (NCEA)
National Education Association (NEA)
National Governors' Association (NGA)
National Science Foundation (NSF)
National School Boards Association (NSBA)
Representative from national textbook publisher

Mathematical Associations and Groups

Mathematically Correct
Mathematics Association of America (MAA)
National Council of Teachers of Mathematics (NCTM)
Third International Mathematics and Science Study (TIMSS)

Educators

Classroom mathematics teachers from public and non-public schools
Principals
District and state mathematics specialists
Mathematics and mathematics education professors from public and private universities, colleges, and community colleges

Technical Experts

University professors
State testing specialists
Representatives from private research organizations

ACKNOWLEDGMENTS

The following people were the primary authors of the introductions to the content areas:

Roger Howe, Yale University (Number Properties and Operations, Geometry, and Algebra)

Richard Scheaffer, University of Florida (Data Analysis, Statistics, and Probability)

Mary Lindquist, Columbus State University (Measurement)