Assessment Development Committee

March 6, 2020

Voir Dire Room

AGENDA

10:30 – 10:45 am	**Update: 2025 NAEP Mathematics Assessment and Item Specifications** ***(SV #5)*** *Michelle Blair, Assistant Director for Assessment Development*	*Attachment A*
10:45 – 11:20 am	**Update: 2025 NAEP Reading Framework** ***(SV #5)*** *Mark Loveland and Cynthia Greenleaf, WestEd*	*Attachment B*
11:20 am – 11:55 pm	**Revision of the Item Development and Review Policy** *Dana Boyd, Chair* *Michelle Blair*	*Attachment C*
11:55 – 12:25 pm	**Changes to NAEP Mathematics, Reading, Civics, and U.S. History Questionnaires** *James Deaton, NCES*	*Under separate cover*
12:25 – 12:30 pm	**Information Items and Next Steps** *Dana Boyd*	
Information Item	**ADC Activities in the Strategic Vision**	*Attachment D*

The National Assessment Governing Board's Strategic Vision (SV) below-listed goals are annotated with "SV" and a number, which will be included on relevant agenda items for Board and Committee meetings during implementation.

Inform

The National Assessment Governing Board will promote The Nation's Report Card's wealth of information to facilitate the awareness and uses of NAEP in appropriate, timely, new, and meaningful ways. Examples of NAEP resources include: results; trends; test questions and tasks; studies; measurement innovations; frameworks that specify the content and design of NAEP assessments; and contextual variables about student demographics and educational experiences collected from students, teachers, and schools. The Governing Board will:

- Strengthen and expand partnerships by broadening stakeholders' awareness of NAEP and facilitating their use of NAEP resources. ***(SV #1)***
- Increase opportunities to connect NAEP to administrative data and state, national, and international student assessments. ***(SV #2)***
- Expand the availability, utility, and use of NAEP resources, in part by creating new resources to inform education policy and practice. ***(SV #3)***
- Promote sustained dissemination and use of NAEP information beyond Report Card releases with consideration for multiple audiences and ever-changing multi-media technologies. ***(SV #4)***

Innovate

The National Assessment Governing Board will revise the design, form, and content of The Nation's Report Card using advances in technology to keep NAEP at the forefront of measuring and reporting student achievement.
The Governing Board will:

- Develop new approaches to update NAEP subject area frameworks to support the Board's responsibility to measure evolving expectations for students, while maintaining rigorous methods that support reporting student achievement trends. ***(SV #5)***
- Continue improving the content, analysis, and reporting of NAEP contextual variables by considering the questions' relevance, sensitivity, and potential to provide meaningful context and insights for policy and practice. ***(SV #6)***
- Research policy and technical implications related to the future of NAEP Long-Term Trend assessments in reading and mathematics. ***(SV #7)***
- Research assessments used in other countries to identify new possibilities to innovate the content, design, and reporting of NAEP. ***(SV #8)***
- Develop policy approaches to revise the NAEP assessment subjects and schedule based on the nation's evolving needs, the Board's priorities, and NAEP funding. ***(SV #9)***
- Develop new approaches to measure the complex skills required for transition to postsecondary education and career. ***(SV #10)***

This Strategic Vision will focus the work of the Governing Board through the year 2020. By pursuing these priorities, the Governing Board will ensure that The Nation's Report Card provides the country with valuable data that measure and contribute to the improvement of student progress in achieving important knowledge and skills necessary for success as citizens in our democratic society. *The Strategic Vision was unanimously approved November 18, 2016*

2025 NAEP Mathematics Assessment and Item Specifications Executive Summary

2025 NAEP Mathematics Framework Background

Each NAEP Assessment is guided by a framework that defines the knowledge and skills to be assessed at each grade. Through active participation of NAEP stakeholders, each framework is developed through a comprehensive process that considers various factors, such as state and local curricula and assessments, widely accepted professional standards, international standards, and exemplary research.

Framework development and update processes are overseen by the Assessment Development Committee (ADC). The ADC conducted a review of the 2017 NAEP Mathematics Framework (last updated in 2006), which included a discussion with external experts as well as a Board-commissioned inventory of state standards. Based on the ADC review, the Governing Board initiated an update of the framework. The Board awarded a contract to WestEd for implementation of the update project. WestEd convened subject matter experts, practitioners, policy makers, administrators, researchers, business representatives, and members of the general public – serving as the Visioning and Development Panels in accordance with their Board-adopted Charge. The Charge calls for recommendations that balance necessary changes with the Board's desire for stable trend reporting, continued breadth of content coverage, and innovation.

Determining the content and format of each NAEP assessment is one of the Governing Board's Congressionally-mandated responsibilities. Using recommendations that reflect Visioning and Development Panel deliberations and public comment, the framework process concludes when the Governing Board adopts a framework that reflects its priorities.

The Development Panel submitted a draft framework responsive to Board deliberations, which the Board adopted in November 2019. Following a progress update, the Board will be asked for comments on the related 2025 NAEP Mathematics Assessment and Item Specifications (attached) at the March 2020 Board meeting.

Milestones

Board Review of State Standards	August 2017 – May 2018
ADC Framework Review	May 2018
Board Adoption of Charge to Framework Panels	August 2018
Board Contract Award to Launch Project	September 2018
Visioning and Development Panel Meetings	November 2018 – September 2019
Public Comment Period	April – June 2019
Board Review and Discussion	May 2019 Board Meeting August 2019 Board Meeting
Board Adoption of Framework	November 2019 Board Meeting
Board Discussion on Specifications	March 2020 Board Meeting
Board Action on Specifications	May 2020 Board Meeting

2025 NAEP Mathematics Specifications Background

The Board adopted the 2025 NAEP Mathematics Framework in November 2019. In January 2020, the Assessment Development Committee (ADC) and the Committee on Standards, Design and Methodology (COSDAM) met to discuss the related Assessment and Item Specifications (attached). This specifications document elaborates on the Board-adopted framework with details for NCES's assessment and item development. Governing Board contractor WestEd has been compiling these details, gathered through the framework process. These details are presented through:

- Demonstrations of how content objectives can be paired with the NAEP Mathematical Practices
- Illustrative items
- Annotations for content objectives in the Framework
- Annotations of 2025 achievement levels descriptors
- Descriptions of special studies to support reporting goals

ADC and COSDAM's January 2020 review focused on two questions:

- Are there any concerns about how this document will support NCES assessment development?
- Are there any aspects out of step with Board priorities?

ADC and full Board discussion on the Specifications document is slated for this March 2020 Board meeting. By May 2020, the Specifications document will be updated to reflect Board discussion along with comments raised in a review by NCES. Below is a summary of how issues raised in the January 2020 joint committee review and the NCES review are addressed in the attached draft.

Issues

[1] Additional Guidance for Initial Item Development

<u>Board discussion</u>*:* The Board discussed the importance of conducting early reviews of draft NAEP items to evaluate whether the new items match the intended construct, as newly defined for the 2025 NAEP Mathematics Framework.

<u>Response</u>*:* To provide this kind of guidance, the Governing Board's current Item Development and Review Policy requires that NCES subject-area item development committees include external experts, and that some of these experts be recruited from the panelists who helped to draft the original framework. The ADC is also beginning discussions to revise this policy, and other potential strategies for achieving the Board's goals will be considered.

[2] Additional Guidance for Sampling the Content Domain

<u>Board discussion</u>*:* The breadth of NAEP assessments often means that content objectives are sampled, i.e., there may be instances where some objectives are not covered in a particular administration year for the assessment. Board members discussed whether there is an opportunity for the 2025 NAEP Mathematics Specifications document to provide more direction to NCES about how to prioritize emphasis, e.g., using information about the nation's future workforce needs.

Response: Currently, NAEP frameworks and specifications provide this guidance for the sub-content domains and the cognitive process dimensions of NAEP assessments, per the Board's Framework Development Policy. More detailed guidance is only noted in a few places. For example, for the 2025 NAEP Mathematics Specifications, the Framework Development Panel explicitly recommended that fractions should be emphasized within grade 4 sub-area of Number Properties and Operations. Accordingly, the draft specifications indicate that "at least one third" of the items for this area will address fractions (see pages 21 and 206). As the ADC begins to update the Board's Item Development and Review Policy, the ADC will also consider how similarly detailed coverage goals should be considered in assessment reviews that the committee conducts.

[3] Additional Guidance for Integrating New Item Types Over Time

Board discussion: Board members discussed whether there is an opportunity for the 2025 NAEP Mathematics Specifications document to provide more direction to NCES about how to integrate new item types over time.

Response: The Governing Board framework development process currently does not develop this level of detailed guidance for NCES. As the ADC begins to update the Board's Item Development and Review Policy, the ADC will also consider how new item types should be considered in assessment reviews conducted by the Committee.

[4] Excessive Elaboration of the Achievement Level Descriptors

Board discussion: The January 2020 draft of the specifications included an appendix with annotations about the differences between the 2025 framework's achievement level descriptors for NAEP Basic versus NAEP Proficient. A similar annotation was provided for differences between the 2025 framework's descriptors for NAEP Proficient and NAEP Advanced. These annotations were provided for each grade level. Board members noted that the annotations could be interpreted as borderline performance descriptors, but borderline performance descriptors are typically generated as part of NAEP Achievement Level Setting processes. To avoid potential confusion, the Board asked for these particular annotations to be removed from the specifications.

Response: WestEd has updated the draft accordingly by removing the annotations on the achievement level descriptors that could be construed as borderline performance descriptors. See the updated Appendix A (attached).

[5] Clarifying How Practices are Being Defined

NCES feedback: NCES reviews of the draft Specifications document have noted opportunities to clarify how the practices will be assessed.

Response: WestEd has provided a number of clarifications throughout the document along these lines. See Chapter 3 for additional details.

[6] Clarifying Ancillary Tools Can Be Used During the Assessment

NCES feedback: NCES reviews of the draft Specifications document have also noted opportunities to clarify how digital tools and off-screen peripherals can be implemented to ensure fidelity to the intended construct. For example, the updated draft specifications further clarify when calculator emulators should be available to students. See page 199 for major clarifications along these lines.

Response: WestEd has updated the specifications to address these clarifications.

[7] Ensuring Overall Accuracy of the Document

NCES feedback: NCES reviews of the draft Specifications document have noted opportunities to update labeling of illustrative items, to confirm scoring criteria, and to address various fixes, such as making grade 4 Number Properties and Operations objective 3.b consistent with itself by deleting the word "whole."

Response: WestEd has updated the specifications to address these corrections.

Assessment and Item Specifications for the 2025 NAEP Mathematics Assessment

24 February 2020 DRAFT

National Assessment Governing Board
U.S. Department of Education

National Assessment Governing Board

Tonya Matthews
Acting Chair

Lesley Muldoon
Executive Director

Lisa Stooksberry
Deputy Executive Director

National Assessment of Educational Progress
Mathematics Project

Steven Schneider
Project Director

Mark Loveland
Deputy Project Director

Ann Edwards & Shandy Hauk
Mathematics Content Leads

Angela Bowzer
Assessment Development Specialist

Developed for the National Assessment Governing Board under contract number 91995918C0001 by WestEd, with a subcontract to the Council of Chief State School Officers.

For further information, contact:

National Assessment Governing Board
800 N. Capitol St., NW, Suite 825
Washington, DC 20002–4233
www.nagb.gov

February 2020

TABLE OF CONTENTS

LIST OF EXHIBITS

LIST OF ILLUSTRATIONS

CHAPTER 1

OVERVIEW

What Is an Assessment Specifications Document?

This document is a companion to the *Mathematics Framework for the 2025 National Assessment of Educational Progress.* The 2025 Mathematics Framework informs NAEP assessment development, describing the subject matter to be assessed, the questions to be asked, as well as the assessment's design and administration. This Assessment and Item Specifications extends the Framework, providing greater detail about development of the items and conditions for the 2025 NAEP Mathematics Assessment. The Framework and these accompanying assessment and item specifications are directed toward the National Center for Education Statistics (NCES) and their contractors, critical NAEP partners, who will use both documents to develop the 2025 NAEP Mathematics Assessment.

Background on NAEP

The National Assessment of Educational Progress (NAEP) has measured student achievement nationally since 1973, and state-by-state since the early 1990s, providing the nation with a snapshot of what students in this country know and can do in mathematics. Starting in 2002, urban school districts that meet certain selection criteria could volunteer to participate in the Trial Urban District NAEP Assessment.

There are two distinct components to the NAEP Mathematics Assessment which differ in purpose. The NAEP Long-Term Trend assessment has measured trends in achievement among 9-, 13-, and 17-year-old students nationally since 1973, and the assessment's content has been essentially unchanged ever since. The second assessment, referred to as "main NAEP," is adjusted over time to reflect shifts in research, policy, and practice. The content and format of the main NAEP Mathematics Assessment are the focus of the Framework.

The main NAEP Mathematics Assessment is administered at the national, state, and selected urban district levels every two years, by Congressional mandate. In mathematics, NAEP 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 level and for large urban districts that volunteer to participate.

Taken together, the NAEP assessments provide a rich and broad picture of patterns in U.S. student mathematics achievement. National and state level results are reported in terms of scale scores, achievement levels, and percentiles. These reports provide comprehensive information about what U.S. students know and can do in mathematics. In addition, NAEP provides comparative subgroup data according to gender, race/ethnicity, socio-economic status, and geographic region; describes trends in performance over time; and reports on relationships between student achievement and certain contextual variables.

The main NAEP assessment is administered to a nationally representative sample of students and reports on student achievement in the aggregate. The assessment is not designed to measure the performance of any individual student or school. To obtain reliable estimates across the population that is assessed, a large pool of assessment items is developed. Subsets of items are administered to each student in the sample. Student results on the main NAEP assessments are reported for three achievement levels established and defined by the National Assessment Governing Board (Governing Board), which oversees NAEP:

- *NAEP Basic* denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the *NAEP Proficient* level.
- *NAEP Proficient* represents solid academic performance for each NAEP assessment. 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.
- *NAEP Advanced* signifies superior performance beyond *NAEP Proficient*.

These policy definitions can be found in the Governing Board's (2018b) *Developing Student Achievement Levels for the National Assessment of Educational Progress Policy Statement*. Descriptions of student performance at these levels of achievement at grades 4, 8, and 12 for the 2025 NAEP Mathematics Framework are provided in Appendix A1. Example items illustrating each achievement level for each grade level are provided in Appendix A2. Chapter 5 includes further discussion of the achievement levels.

This document describes specifications for an assessment framework, not a curriculum framework. The 2025 NAEP Mathematics Framework lays out the basic design of the assessment by describing the mathematics content and mathematical practices that should be assessed and the types of questions that should be included. The specifications in this document extend and illustrate these ideas. The Framework also describes how various assessment design factors should be balanced across the assessment. In broad terms, the Framework attempts to answer the question: What mathematics knowledge, skills, and practices are to be assessed on NAEP at grades 4, 8, and 12? The Framework does not cover all relevant content for each grade level; some concepts, practices, and activities in school mathematics are not suitable to be assessed on NAEP, although they may well be important components of a school curriculum. For example, the practice of extended investigation would not be possible in the NAEP assessment, although it would be quite reasonable for teachers to have multi-day investigations of some important mathematical ideas. This document also does not attempt to answer the question: How should mathematics be taught?

The Visioning and Development Process

The process for updating the mathematics assessment framework started with a review of existing frameworks by experts in mathematics education research, policy, and practice representing key stakeholder groups. This process – which is described in the Governing Board's (2018c) *Framework Development Policy Statement* and elaborated in the 2025 NAEP Mathematics Framework – involved visioning for the update, and then development. For more on this process, see Chapter 1 of the Framework. Complementary to the Visioning and Development Panels, a Technical Advisory Committee (TAC) of eight recognized measurement experts advised the panels about technical issues and provided feedback on drafts of this specifications document as it was developed.

Overview of Assessment Design and Item Specifications

The Assessment and Item Specifications that guided the development and implementation of the NAEP Mathematics Assessments administered since 2009 were established more than 10 years ago, and significant updates to the Framework for the 2025 assessment have been made. The update included revisions to the mathematics content objectives, the description of new NAEP Mathematical Practices, attention to the evolving role of technology in students' in-school and out-of-school experiences, and consideration of new item formats. These changes required a parallel update to the Assessment and Item Specifications.

The proposed design for the 2025 assessment aims to provide a fair and valid measure of how well all students have achieved the depth and breadth of the mathematics content and practice articulated by the Framework. To do this, the design:

- incorporates a mix of traditional and innovative item types that reflect recent research on the science of learning, to capture both the process and outcomes of student learning, and emphasizes authentic applications of mathematics knowledge and skill;
- capitalizes on the use of technology to assure accessibility, promote engagement for all students, and explore new options for task design and scoring, including the use of multimedia;
- encourages continuing prototyping and research to capitalize on the capacities of current and emerging technology to assess students at deeper levels, while still ensuring validity and fairness of scores; and
- recognizes the potential of technology and new task designs while also acknowledging limitations and potential negative unintended consequences. The design plan is a careful balance to promote more valid assessment of mathematics content and practices without compromising fairness or reliability (e.g., fairness for students who have less access to technology, scenarios that avoid construct-irrelevant barriers of language, and innovative task types that reduce the number of items).

Text and sample items that support and help to clarify the description of the assessment design in the Framework have been included in this Assessment and Item Specifications document. Illustrations include both examples and non-examples to assist in the development and implementation of updates for the NAEP Mathematics Assessment.

Introduction to the Assessment and Item Specifications

The Assessment and Item Specifications consists of five chapters and several appendices. Throughout this document, figures have been included to illustrate a particular point of emphasis from the Framework. Note that "illustration" is used throughout the Assessment and Item Specifications to indicate figures not included in the Framework. Illustrations in this document include non-examples – "anti-exemplar" items to support item writers in avoiding items that "function as simpler item types, incorporate superficial complexity that does not improve fidelity to the construct, introduces construct-irrelevant variance, or any combination of the three" (Martineau, Dadey, & Marion, 2018, p.1). Each exhibit carried from the Framework into the Assessment and Item Specifications remains labeled as an "exhibit." Therefore, in this document, illustrations are numbered consecutively and separately from exhibits.

Chapter 2 describes the content domains: Number Properties and Operations (including computation and understanding of number concepts); Measurement (including use of instruments and concepts of area and volume); Geometry (including spatial reasoning and applying geometric properties); Data Analysis, Statistics, and Probability (including graphical displays and statistical measures); and Algebra (including representations and relationships). Each content area is broken into subtopics (e.g., for Number Properties and Operations, these are number sense, estimation, number operations, ratios and proportional reasoning, and properties of number and operations) identifying what should be measured on NAEP at grades 4, 8, and 12. Further specifications are added to some content areas and most objectives to clarify the measurement intent for item writers.

Chapter 3 describes the NAEP Mathematical Practices that play a role in measuring student knowledge and skills in mathematics. These are Representing, Abstracting and Generalizing, Justifying and Proving, Mathematical Modeling, and Collaborative Mathematics. The chapter argues that content and practices are interwoven and interdependent: one cannot demonstrate mathematics achievement without knowing content and being able to think mathematically. Chapter 3 also offers example items across grades 4, 8, and 12 that illustrate how NAEP Mathematical Practices can be assessed with particular content. Illustrations in this chapter include non-examples – “anti-exemplar” items to support item writers in avoiding potential barriers to NAEP Mathematical Practice alignment.

Chapter 4 focuses on issues of technology and accessibility, assessment design, and item format. The chapter argues for the need to ground the NAEP Mathematics Assessment in tasks in familiar contexts to foster student engagement. Second, by expanding item types and thoughtfully using technology, the NAEP Mathematics Assessment can provide greater access to all students, diversify the ways in which student achievement can be recognized and measured, and more robustly assess both what students know and can do. This will involve expanding to include scenario-based tasks (which involve clusters of related items within one task) along with continued use of existing discrete NAEP items that capture student understanding of content and mathematical practices. As the technology of assessment evolves, alternative formats might also be considered. Illustrations in this chapter include examples and non-examples to clarify less familiar item types and best practices in item development.

Chapter 5 addresses how NAEP results are reported. The chapter describes the three NAEP achievement levels and the development of the mathematics achievement level descriptions (see Appendix A1). The chapter builds on an expansive conception of “opportunity to learn” as called for by the Visioning Panel Guidelines (see Appendix C). The chapter also discusses how research on student diversity and schooling informs mathematics-specific contextual variables.

Opportunity to Learn and an Expansive Understanding of Contextual Variables

What students learn is inseparable from the conditions of their learning and broader social aspects of mathematics learning. Hence, interpreting differences in what students can do on NAEP requires an understanding of the range of factors that affect student learning. In particular, the Framework articulates an expansive conception of opportunities to learn informed by educational research on students and their in- and out-of-school learning and experiences, as well as research on the variations in human, material, and social resources that shape what students have an opportunity to learn about mathematics in the U.S. (e.g., Cohen, Raudenbush, & Ball, 2003; Tatto et al., 2012).

Opportunity to learn (OTL) is generally understood to refer to inputs and processes that shape student achievement, including the school conditions, curriculum, instruction, and resources to which students have access. When opportunity to learn was first used as a concept, Carroll (1963, 1989) emphasized the time allowed for learning. For the past 50 years, the concept of opportunity to learn has continued to evolve, as have efforts to measure in-school opportunities to learn, with the majority of scholars focusing on the classroom as the unit of analysis and instruction as central. Research, for example, has documented the negative effects on achievement of policies and practices that are often found in schools serving children who live in poverty or have special needs, including an inadequate supply of mathematics teachers with strong knowledge and skills, a tendency to offer few advanced mathematics courses, and a common practice of tracking these students disproportionately into low-level courses that restrict their learning opportunities (e.g., Husén, 1967; Tan & Kastberg, 2017), all of which can be understood as instructional resources that shape what students learn.

Important to note is the sociopolitical turn that has taken place in research on school mathematics (Gutiérrez, 2013), which positions mathematics as a "dynamic, political, historical, relational, and cultural subject" (TODOS & NCSM, 2016, p. 3) in which identity and power both play central roles. This turn has led scholars and educators to explore how school mathematics marginalizes and alienates students who do not see connections to their own lives and experiences. It raises questions about how school mathematics might be reformed to engage all students and their communities. This includes students with disabilities who are often relegated to classrooms where learning differences are conceptualized as a deficit rather than a potential strength, and that focus on procedural approaches rather than leveraging students' own particular strategies to engage in mathematical reasoning and sense making (e.g., Lambert, Tan, Hunt, & Candella, 2018).

Another noteworthy development in mathematics education research is acknowledgment that students themselves are a resource in learning, including their interests, abilities, and in- and out-of-school experiences. Research, for example, suggests that students' experiences out-of-school can be directly relevant to the ways they think mathematically and use mathematics (e.g., Martin, 2000; Nasir & Hand, 2008). Some scholars refer to this as students' "funds of knowledge," defined as the skills, knowledge, habits of mind, practices, and experiences acquired through historical and cultural interactions of an individual in their community, family life, and culture through everyday living as well as in school (e.g., Aguirre et al., 2013; Civil, 2016; de Freitas & Sinclair, 2016; González, Moll, & Amanti, 2005; Moll, Amanti, Neff, & González, 1992). Students' funds of knowledge include what has often been referred to as students' prior knowledge, but expands that idea to include cultural, linguistic, and social traditions that characterize students' out of school lives. While these funds of knowledge might differ from those of the teacher or the traditional curriculum, the broad experiences of students can be used to make powerful connections that enable learning and thus can be understood as an additional resource in instruction and assessment. Therefore, the Framework's conception of OTL includes "students' experiences, out-of-school learning, and funds of knowledge" as an instructional resource.

Relevant OTL indicators have been clustered in various ways (e.g., Abedi & Herman, 2010; Elliott & Bartlett, 2016; Herman, Klein, & Abedi, 2000; Husén, 1967; Schmidt, Burroughs, Zoido, & Houang, 2015; Wang, 1998). These can be grouped into five strands: time, content, instructional strategies, teachers, and instructional resources. Examples of indicators that have been used in research are provided in Exhibit 1.1.

Exhibit 1.1. Opportunity to Learn (OTL) Strands

OTL Strand	Example indicators
Time (OTL-T)	time scheduled for instruction proportion of allocated time used for instruction time students are engaged in learning time students are experiencing a high success rate of learning
Content and Practices (OTL-C/P)	content and practices exposure content and practices emphasis content and practices coverage
Instructional Strategies (OTL-IS)	instructional approaches (e.g., strategies that facilitate student thinking and understanding, instruction that promotes student engagement) classroom climate instructional group size
Teacher Factors (OTL-TF)	teacher preparation and professional development teacher knowledge, including mathematical knowledge for teaching teaching experience teacher attitudes about themselves, students, learning, and mathematics
Instruction Relevant Resources (OTL-IR)	material resources (e.g., textbooks, manipulatives) school policies (e.g., tracking) school community and climate; school and instructional leadership students' experiences, out-of-school learning, and funds of knowledge student access to technological tools

To support audiences in interpreting NAEP results, information about contextual variables is collected through student, teacher, and administrator surveys. The Framework development process drew broadly on the literature to create an ambitious conception of OTL as the basis for recommendations about mathematics-specific contextual variables on NAEP surveys. As is the case with mathematics content, it is neither possible nor appropriate to measure all potentially relevant contextual variables on NAEP. For example, questions that ask students about their home or out-of-school experiences can be experienced as intrusive. Priorities for the selection of mathematics-relevant OTL variables are described in Chapter 5.

Major Changes in the 2025 NAEP Mathematics Assessment and Item Specifications

This Assessment and Item Specifications reflects several major changes, both those made to the Framework and those made to support item development. The changes are summarized below and elaborated in Exhibit 1.2 at the end of this section.

Mathematics Content

Chapter 2 presents an updated set of content objectives for the 2025 NAEP Mathematics Assessment at grades 4, 8, and 12. The updates reflect the last decade of changes in state standards for mathematics curriculum, instruction, and assessment. State standards shape what students have had an opportunity to learn by the time they take a NAEP assessment. To ensure the updates reflect current state-level emphases for mathematics content, the Framework incorporates findings from several reports that compared NAEP and state standards (e.g., Achieve, 2016; Johnston, Stephens, & Ratway, 2018), as well as reports on the mathematics content taught in leading countries around the world (e.g., as assessed in the *Trends in International Mathematics and Science Study* [TIMSS], National Center for Education Statistics [NCES], 2019 and *Programme for International Student Assessment* [PISA], Organisation for Economic Co-operation and Development [OECD], 2019). Because the Framework has been written for an assessment in 2025 and beyond, it is also informed by national policy that foreshadows likely changes in state policy (e.g., Franklin et al., in press; Garfunkel & Montgomery, 2016).

Mathematical Literacy

In every state, all high school graduates are required to study mathematics whether or not their future plans involve college or a field in which high school mathematics is heavily involved. The purpose of this universal practice is to ensure that the U.S. citizenry is mathematically literate. Recent policy developments have included attention to mathematical literacy, for example, in mathematical modeling of real-world problems and interpreting reports of data.

Mathematical literacy is the ability to apply mathematical concepts to everyday situations. It has been recognized worldwide as important. In 2015, the PISA assessments, given to 15-year-olds every three years, were conducted in 70 countries, more countries than any other mathematics assessment (OECD, 2018). The PISA assessments emphasize mathematical literacy and define it as the application of numerical, spatial, or symbolic mathematical information to situations in a person's life as a consumer, employee, or citizen. The definition for the Framework is based on the PISA definition, given its extensive, worldwide use and given the availability of assessment items that have been created following that definition:

> *Mathematical literacy is the application of numerical, spatial, or symbolic mathematical information to situations in a person's life as a community member, citizen, worker, or consumer.*

A large body of experiences can be viewed as requiring mathematical literacy, including: fluency in the broad range of mathematics of personal finances; understanding statistical information and displays found in print and visual media; household tasks such as cooking, cleaning, and furnishing that require a variety of measurements. For example, mathematical literacy affects how one critically evaluates reports on environmental issues, estimates how many bricks are needed to build a walkway, or compares interest rates for a loan. Mathematical literacy is part of the everyday experiences that occur in community, civic, professional, and personal contexts of adults in the United States, regardless of career.

At grades 4 and 8, instances of mathematical literacy are found in the standard content taught in schools, have been in previous NAEP frameworks, and remain in the objectives enumerated here. At grade 12, historically, instances of mathematical literacy have been given less attention. In the 2025 NAEP Mathematics Framework, throughout grade 12, objectives that provide opportunities for assessment of mathematical literacy are identified by the number/hashtag sign (#). See Chapter 2 for more on the issue of mathematical literacy.

NAEP Mathematical Practices

Since the late 1980s, there have been ongoing efforts to more clearly specify mathematical processes like "higher-order thinking" or "mathematical reasoning." Current conceptions of mathematical knowledge and skill have shifted to specify mathematical practices and processes. At the turn of the 21st century, in *Adding It Up*, the National Research Council (2001) enumerated five strands of mathematical proficiency, including

- *conceptual understanding*: comprehension of mathematical concepts, operations, and relations
- *procedural fluency*: skill in carrying out procedures flexibly, accurately, efficiently, and appropriately
- *strategic competence*: ability to formulate, represent, and solve mathematical problems
- *adaptive reasoning*: capacity for logical thought, reflection, explanation, and justification
- *productive disposition*: habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficacy.

For decades, the National Council of Teachers of Mathematics (NCTM) has discussed five "mathematical processes standards": problem solving, reasoning and proof, communication, connections, and representation (NCTM, 2000). Processes like these have been central to NAEP frameworks for the last 20 years and state standards have re-iterated the important role of practices. The language of "practice" has become increasingly popular, establishing a foothold through various state standards, as well as in discussions of teaching with and through practices (NCTM, 2014). The Framework provides the following definition:

> *NAEP Mathematical Practices are the routines, norms, and processes needed to do the work of mathematics.*

Based on the current state of the field, the Framework identifies five NAEP Mathematical Practices for the NAEP Mathematics Assessment:

NAEP Mathematical Practice 1: Representing
NAEP Mathematical Practice 2: Abstracting and Generalizing
NAEP Mathematical Practice 3: Justifying and Proving
NAEP Mathematical Practice 4: Mathematical Modeling
NAEP Mathematical Practice 5: Collaborative Mathematics

These mathematical practices are described in depth in Chapter 3. Note that these mathematical practices are not instructional practices used by teachers. They are the actions necessary to do mathematics. This list of NAEP Mathematical Practices also does not endorse one particular view of mathematical practices (an issue further discussed in Chapter 3).

Item Formats and Technology in Assessment

A fourth major change involves item formats and the role of technology in assessment. As noted above and as further explained in Chapter 4, technological innovation is relevant to NAEP because it allows both for more authentic assessments and for a broader range of accommodations to meet students' needs.

Since 1992, the NAEP Mathematics Assessment has used three formats for the items (questions): multiple choice, short constructed response, and extended constructed response. In 2017, the NAEP assessment began to include these item formats in a digital platform as part of the NAEP transition to digitally based assessment. The transition to digital administration provided opportunities to expand the range of formats used for items.

In advancing the expansion of item types and formats, three themes emerged. One theme concerns how research on student knowledge and experience can be used to design assessments that capture their capacity to do mathematics. This includes the use of interactive, multimedia scenario-based tasks to assess what students know and can do. Scenario-based tasks currently exist in other NAEP assessments including NAEP Science and NAEP Technology and Engineering Literacy.

By expanding item formats, to include scenario-based tasks (and new item formats that emerge in the future), and to thoughtfully use technology, the aim is to provide greater access to all students, as well as to diversify the ways in which student achievement can be recognized and measured. Note that technological innovation is not just limited to enhancing assessment accommodations. Technology is a part of every student's life and learning, and mathematical thinking can be enhanced by its judicious use.

A second theme concerns the use of technology to enable assessment of the NAEP Mathematical Practices, including an expanded range of response types leveraging object-based and discourse responses within a scenario-based task. Less often noted but equally important is a third theme concerning the intended or unintended negative consequences of technology, which include inequitable access to technologies. That is, while technology may have the potential to increase access and opportunities to demonstrate learning, students unfamiliar with technologies used in the assessment could be at a disadvantage. With the introduction of scenario-based tasks it is critical to ensure that students have ample time to understand how to engage with assessment items along with opportunities to experience the task type.

Changes from the 2009–2017 Assessment and Item Specifications

Exhibit 1.2 compares the Framework and Assessment and Item Specifications for the 2025 NAEP Mathematics Assessment and those used for the 2009–2017 NAEP Mathematics Assessments. The focus here is on major changes. Many of the points summarized below are expanded in Chapters 2, 3, and 4. Justifications for these changes are briefly described below, with more details in the relevant chapters.

Exhibit 1.2 Summary of Changes in the 2025 NAEP Mathematics Framework and Assessment and Item Specifications

Topic	Change	Rationale
Mathematics Content	Many objectives edited to increase clarity and specificity.	Objectives and balance of topics were updated to reflect shifts in expectations evident from reviews of state and national standards, policy documents from leading professional organizations, and expectations for mathematical literacy on U.S. and international assessments. For more details on changes, see Chapter 2.
	Additional clarifications and limitations were included with the content objectives to further guide item development.	Suggestions were included to reflect content descriptions from the previous Assessment and Item Specifications (2009), 2025 Framework authors, state standards and supporting documents, and public-facing information from current state and national assessments.
	The objectives in the mathematical reasoning subtopics have been removed–this subtopic was introduced in 2009 for Number Properties and Operations; Geometry; Data Analysis, Statistics, and Probability; Algebra.	With the introduction of the NAEP Mathematical Practices (see Chapter 3), mathematical reasoning was no longer needed as a subtopic. To preserve attention to the content that was uniquely present in some of the mathematical reasoning objectives, objectives in other subtopics were revised. For more details on changes, see Chapter 2.
	Distribution of items for grade 12 remains the same. The proportion of Data Analysis, Statistics, and Probability items has increased for grade 8 and decreased for grade 4. Concurrently, the proportion of items in Measurement in grade 8 decreased and the proportion in Number Properties and Operations in grade 4 increased.	Adjustments to the proportion of items on the assessment in Data Analysis, Statistics, and Probability at grades 4 and 8 reflect changes in opportunity to learn common across state standards. The distribution of attention to content topics in state standards informed the related decisions to increase the proportion of items at grade 4 in Number Properties and Operations and decrease the proportion in Measurement at grade 8. For more details on changes, see Chapter 2.

	Illustrations containing items and associated text providing clarification for specific text from the Framework were included.	Illustrations containing example and non-example items as well as discussions of these were included to assist item writers in developing a richer understanding of what was (and was not) intended by the Framework.
Mathematical Complexity (2009-2017 Framework)	This was a chapter that defined mathematical complexity as "the demands on thinking that an item expects" (National Assessment Governing Board, 2017, p. 37). The chapter was removed.	"Mathematical complexity" aimed to address the process dimension, the "doing" of knowing and doing mathematics. It was a mixing of cognitive demands (e.g., on working memory, reading comprehension, and attention) and the challenges inherent in developing mathematical understanding. However, it was not supportive of score interpretation. Many decades of research and development have shown that assessing students' knowledge and use of mathematics is more nuanced than was accounted for in the "mathematical complexity" approach used in the 2009 to 2017 frameworks.
NAEP Mathematical Practices (NEW)	A new chapter, Chapter 3–NAEP Mathematical Practices, has been added describing and illustrating the assessment of five mathematical practices through which students engage in knowing and doing mathematics.	Since the 1990s, the field of mathematics education has seen increasing focus on mathematical processes and the interacting social and mental activities of knowing and doing mathematics. This chapter reflects the field's attention to mathematical activity by describing five NAEP Mathematical Practices. These are assessable aspects of activity at work across mathematics content when students do mathematics.
	Distribution of items for each mathematical practice were developed.	Most NAEP Mathematics Assessment items will feature at least one of the five NAEP Mathematical Practices (55 to 85 percent). This range allows flexibility in assessment and item development across grades 4, 8 and 12 while also ensuring that the majority of the assessment is designed to capture information on student knowledge while engaging in mathematical practices. The balance of

		items (15 to 45 percent), will assess knowledge of content without calling on a particular mathematical practice (e.g., procedural or computational skill).
	Items illustrative of a NAEP Mathematical Practice or serving as non-examples of a practice were introduced within the text for each practice.	These items were included to provide additional support for item writer conceptualization of the NAEP Mathematical Practices.
Item Formats and Assessment Design	Two chapters in the previous framework (Item Formats and Design of Test and Items) were merged into a single chapter, Overview of the Assessment Design, and updated.	The combination of chapters on assessment and item design allowed addressing interrelationships among: (1) the new digital format of NAEP administration, and (2) developments in technology for assessment, including scenario-based tasks.
	A new format, scenario-based task, was introduced.	With the addition of scenario-based tasks, the NAEP Mathematics Assessment continues to provide greater access to all students, diversifies the ways in which student achievement can be recognized and measured, and more robustly assesses both what students know and what they can do.
Calculator Policy	Continuing the policy established for the 2017 digital administration of NAEP, students will have access to a calculator emulator in blocks of items designated as “calculator blocks”: four-function for grade 4, scientific for grade 8. The one change in 2025 will be that the grade 12 calculator will include a graphing emulator.	High school students typically use graphing calculators or online emulators and not scientific calculators (Crowe & Ma, 2010).

Item Types	Chapter 4 includes updates to reflect current and future digital platform use and the new format option of scenario-based tasks.	To better assess the diversity of ways of doing mathematics, technology available now and in the near future allows scenario-based tasks. Scenario-based item collections can be used to assess aspects of mathematical activity that have been difficult (if not impossible) to assess in the past. Building on the work in the last five years to use scenario-based tasks in NAEP Science and NAEP Technology and Engineering Literacy assessments, Chapter 4 details the ways scenario-based and traditional items can be combined to assess achievement in mathematics content and NAEP Mathematical Practices.
	Items illustrative of an item type or best practice in development of items for the NAEP Mathematics Assessment were introduced. Illustrations serving as non-examples of best practices in development of items for the NAEP Mathematics Assessment were also included.	Illustrations containing example and non-example items and discussions of these items were included to provide additional support for application of best practices in item writing for the NAEP Mathematics Assessment and actualization of potential NAEP mathematics item types.
Tools and Manipulatives	Students will continue to have the tools and manipulatives used in the digital administration of the 2017 NAEP Mathematics Assessment. Chapter 4 also explores the potential of behind-the-scenes technology to capture and use process data for assessment; these are data generated by students as they work with the assessment.	The existing digital system tools and mathematics-specific tools have proven worthwhile since the 2017 administration. Additionally, in acknowledgement of the continuing evolution and use of technology in mathematics, Chapter 4 includes examples of other tools (e.g., simulations, dynamic geometry software, and “smart” physical objects) that may be common in 2025 and beyond.

Aligning the Assessment with the Framework and Assessment and Item Specifications

The assessment should be developed so that it is aligned with the guidelines defined by the intersection of content objectives and NAEP Mathematical Practices, as set forth in the Framework and the Assessment and Item Specifications. More specifically:

- The content of the assessment should be matched with the content of the Framework and Assessment and Item Specifications. The assessment as a whole should reflect the breadth of knowledge covered by content objectives, clarifications, and limitations in the Framework and Assessment and Item Specifications. The content of the assessment should not go beyond the content boundaries as defined in these documents. The assessment should represent the balance of mathematics content at each grade as described in Chapter 4 of the Framework and Assessment and Item Specifications.
- The mathematical practices reflected in items on the assessment should be matched to those in the Framework and Assessment and Item Specifications. The assessment should represent the balance of NAEP Mathematical Practices at each grade as described in Chapter 4 of the Framework and Assessment and Item Specifications.
- While it is not possible to cover all possible combinations of content objectives and practices for each achievement level on one assessment, appropriate alignment between the assessment and the Framework and Assessment and Item Specifications at each grade should be maintained in the item pools. The assessment should be built so that the constructs represented by the objectives for each content area are adequately represented. The breadth and relative emphasis of mathematics knowledge covered in each content area, as presented in the Framework and Assessment and Item Specifications, should be represented on the assessment as a whole. The developer should avoid under- or overemphasizing particular content objectives, NAEP Mathematical Practices, or achievement expectations, the goal being to ensure broad coverage in any given year's item pool and coverage of all content objectives over time.
- The assessment should represent the balance of response types specified in Chapter 4 of the Framework and Assessment and Item Specifications and give appropriate emphasis to the testing time allocated for scenario-based tasks.
- The assessment should report and interpret scores based on the Framework, the Assessment and Item Specifications, and the NAEP Achievement Level Descriptions (ALDs). That is, the assessment should be developed so that scores will reflect both the guidelines in the Framework and Assessment and Item Specifications and the range of performances illustrated in the NAEP Mathematics ALDs.
- The assessment design should match the characteristics of the targeted assessment population. That is, the assessment should give all students tested a reasonable opportunity to demonstrate their knowledge and skills in the content areas and mathematical practices covered by the Framework and Assessment and Item Specifications.

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 criteria, such as grade level and content area. Information about the items includes student performance data 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 or identified by the question ID from the NAEP Questions Tool, found on the website.

CHAPTER 2

MATHEMATICS CONTENT

The NAEP Mathematics Assessment measures what mathematics students know and are able to do, which involves understanding of particular mathematical ideas (content) and of how to use those ideas in mathematical activity (practices). The content of mathematics can be described by nouns: numbers, data, variables, functions, graphs, geometric figures of various kinds, and the like. In contrast, mathematical practices can be described by verbs: recognize, generalize, deduce, justify, and other processes of mathematical reasoning; represent, use, symbolize, and other actions involved in applying mathematics; describe, explain, model, and other activities inherent in mathematics being a discipline that is socially constructed by, and communicated among, individuals and societies.

This chapter focuses on the mathematics content objectives; Chapter 3 focuses on the NAEP Mathematical Practices. Mathematical proficiency involves knowing both.

Content Areas

NAEP has regularly gathered data on students' understanding of five broad areas of mathematics content:

- **Number Properties and Operations** (including computation and 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)
- **Algebra** (including expressions, equations, representations, and relationships)

Classification of an item into one primary content area is not always clear-cut, but it helps to ensure that the indicated mathematical concepts and skills are assessed in a balanced way.

Certain aspects of mathematics occur in all content areas. For example, there is no single objective for computation. Instead, computation is embedded in many content objectives. In the Framework, computation appears in the Number Properties and Operations objectives, which encompass a wide range of concepts about the numeration system and explicitly include a variety of computational skills, ranging from operations with whole numbers to work with decimals, fractions, percents, and real and complex numbers. Computation is also critical in Measurement and Geometry in determining, for example, the perimeter of a rectangle, estimating the height of a building, or finding the hypotenuse of a right triangle. Data analysis often involves computation in calculating a mean, or other statistics describing a collection of values, or in calculating probabilities. Solving algebraic equations also frequently involves numerical computation.

The objectives describe what is to be assessed on NAEP given operational limitations. As noted in Chapter 1, the NAEP content objectives are not a complete description of mathematics that should be taught at these grade levels.

NAEP Mathematics Assessment Objectives Terminology

Some terms that are broadly used in mathematics education must take on narrower meanings in order to clearly describe measurable mathematics objectives. To support item development aligned with the objectives given in this document, several points bear mention:

- The phrase "solve problems" means to complete tasks where the task contexts may range from the purely mathematical to those that are experientially concrete or real to students.
- When the word "or" is used in an objective, it means that an item may assess one or more of the concepts included and the full collection of items will include assessment of each listed concept. The table in Illustration 2.1 provides example objectives to further clarify this intention.
- Specific to grade 12 are three distinctions in NAEP content objectives:
    - Some grade 12 objectives are marked with an asterisk (*). This denotes objectives that describe mathematics content beyond what is typically taught in a 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra, with statistics and probability included). These objectives will be selected less often than the others for inclusion on the assessment.
    - Some objectives in grade 12 are marked with the number and hashtag sign (#). This designates objectives that most closely reflect opportunities to assess mathematical literacy. However, not all items associated with an objective that has the # sign will assess mathematical literacy.
    - At grade 12, geometry and measurement are combined as one content area. This reflects the fact that the majority of measurement topics suitable for high school students are geometric in nature.
- Although every assessment item will be assigned a primary classification, some items could potentially fall under more than one objective.

As mentioned in Chapter 1, "Illustration" is used throughout the Assessment and Item Specifications to indicate exhibits and figures not included in the Framework. These include examples and non-examples intended to further clarify particular points of emphasis in the Framework. Each exhibit carried from the Framework into the Assessment and Item Specifications remains labeled as an "Exhibit."

The table in Illustration 2.1 provides examples to further clarify the development of items when the content objective uses the word "or."

Illustration 2.1. Example: Multi-Verb Objectives and the Use of "Or"

Grade Level	Number Properties and Operations Objective	Clarifications
4	3e) Interpret, explain, or justify whole number operations and explain the relationships between them.	The item pool will include items that measure each of the four targets of this objective: (1) interpreting whole number operations, (2) explaining whole number operations, (3) justifying whole number operations, and (4) explaining the relationships between whole number operations.
8	3e) Interpret, explain, or justify rational number operations and explain the relationships between them.	The item pool will include items that measure each of the four targets of this objective: (1) interpreting rational number operations, (2) explaining rational number operations, (3) justifying whole number operations, and (4) explaining the relationships between rational number operations.
12	3e) *Analyze or interpret a proof by mathematical induction of a simple numerical relationship.	The item pool will include items that measure each of the two targets of this objective: (1) analyzing a proof of a simple numerical relationship by mathematical induction, and (2) interpreting a proof of a simple numerical relationship by mathematical induction.

Mathematical Literacy

As noted in Chapter 1, mathematical literacy is related to an individual's capacity to "understand the role that mathematics plays in the world, to make well-founded judgments and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned citizen" (PISA, 2003, p. 3). It includes the ability to formulate and interpret problems, and to use mathematical knowledge and skill in creative ways across a range of situations – complex and simple, routine and unusual. These situations can occur in one's private life (measuring cloth for a project), one's occupational and professional life (using proportions to make sense of a situation), one's social life with friends or family (paying in a restaurant), and in one's life as a citizen (processing information relevant to voting).

Some objectives at grade 12 are identified with the theme of mathematical literacy. If there are everyday applications of the objective to situations in a person's life as a community member, citizen, worker, or consumer, then the number/hashtag sign (#) precedes the objective. For example, for an objective that calls for students to analyze situations, develop mathematical models, or solve problems using a particular form of equation or inequality, mathematical literacy items might be given in real-world contexts such as solving a problem about tax implications of a workplace policy change, or, in the context of community decisions, analyzing or modeling with an inequality the upper bounds for safe levels of lead in water from a local water treatment facility. Other items *not* focused on mathematical literacy might ask the student to solve a problem by graphing the consequences of doubling the value of a variable in a linear relationship.

As another example, a mathematical literacy assessment item might provide information about a seismic magnitude scale (used to measure the intensity of earthquakes), indicate that on the scale a Magnitude 5 earthquake is ten times stronger than a Magnitude 4 earthquake, and ask grade 12 students to make sense of, model, or draw conclusions in a problem situation that uses that information. An alternate assessment item for the same objective that would not be focused on mathematical literacy might ask students to apply and justify the use of logarithms to determine the seismic magnitude measurement in a given situation. The goal of the identification of objectives with # is to support exploration of NAEP reporting on mathematical literacy. See Appendix E for a description of a special study on measuring and reporting on mathematical literacy.

Item Distribution

The distribution of items among the various mathematics content areas is a critical feature of the assessment design because it reflects the relative importance given to each area in the assessment. As has been the case with past NAEP assessments, the categories have different emphases at each grade. Exhibit 2.1 provides the balance of items in the assessment by content area for each grade (4, 8, and 12). The percentages refer to the proportion of items, not the amount of testing time.

For the 2025 NAEP Mathematics Assessment, a greater number of items assessing fraction concepts will be sampled than have been in past administrations. This increase not only reflects the focus on fraction instruction in the early grades, but also the importance of understanding student early knowledge of and skills with fraction concepts, as they are a predictor of success in high school mathematics courses (Siegler, et al., 2012).

Exhibit 2.1. 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>45*</td><td>20</td><td>10</td></tr>
<tr><td>Measurement</td><td>20</td><td>10</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>5</td><td>20</td><td>25</td></tr>
<tr><td>Algebra</td><td>15</td><td>30</td><td>35</td></tr>
</table>

*Note: Increased attention to assessing fraction content – at least one-third of grade 4 Number Properties and Operations items should assess fraction content.

NAEP Mathematics Objectives Organization

Mathematical ideas in different content areas are often interconnected. Organizing the Framework by content areas has the potential for obscuring these connections and leading to fragmentation. However, the intent here is that the objectives and the assessment of those objectives will, in many cases, cross content area boundaries.

To provide clarity and specificity in grade level objectives, the Framework matrix (Exhibits 2.2, 2.3, 2.4, 2.5, and 2.6) depicts the objectives appropriate for assessment under each subtopic. For example, within the Number Properties and Operations subtopic of Number Sense, specific objectives are listed for assessment at grades 4, 8, and 12. In general, objectives within content areas are different across the grades. Occasionally, the same objective may appear at more than one grade level; this suggests an implicit developmental sequence for that concept or skill. An empty cell in the matrix conveys that an objective is not appropriate or not deemed as important as other areas for assessment at that grade level. Explanations of changes in the mathematics objectives are elaborated in the final section of this chapter.

Objective Alignment and Illustrations

Throughout this Assessment and Item Specifications document, assessment items have been included to illuminate particular text in the Framework. The items used in exhibits and illustrations come from a variety of sources, including released items from the NAEP Questions Tool (NCES, n.d.), suppliers of state assessments – including the Smarter Balanced Assessment Consortium (SBAC, 2018) and the Partnership for Assessment of Readiness for College and Careers (PARCC, 2015) – and international mathematics assessments, such as TIMSS, PISA, and England's *Key Stage* tests. Sources are named with the description of each item, and a note is included when the item has been modified for the purposes of this document.

At the top of each illustration is a metadata table with key information about the item used. These metadata are specific to the 2025 NAEP administration and identify five pieces of information (see Illustration 2.2a).

- **Grade Level**: identifies the 2025 grade level
- **Content Area**: identifies the 2025 primary content area. Abbreviations for each content area used throughout this document are included in parentheses.
    - Number Properties and Operations (Num)
    - Measurement (Meas)
    - Geometry (Geom)
    - Data Analysis, Statistics, and Probability (Data)
    - Algebra (Alg)
- **Assessed Practice(s)**: identifies the assessed NAEP Mathematical Practice(s)
- **Objective ID**: identifies the 2025 NAEP content objective alignment
- **Item Format**: identifies the 2025 item format. Abbreviations used for item formats are listed below. See Chapter 4 for a description of each.
    - SR: selected response
        - SR – MC (multiple choice)
        - SR – MS (multiple select)
        - SR – matching
        - SR – zone
        - SR – grid
        - SR – IC (in-line choice)
        - SR – composite
    - SCR: short constructed response
        - SCR – FIB (fill in the blank)
        - SCR – composite
    - ECR: extended constructed response

- ECR – ET (extended text)
- ECR – composite

Illustration 2.2a. Example: Item Metadata

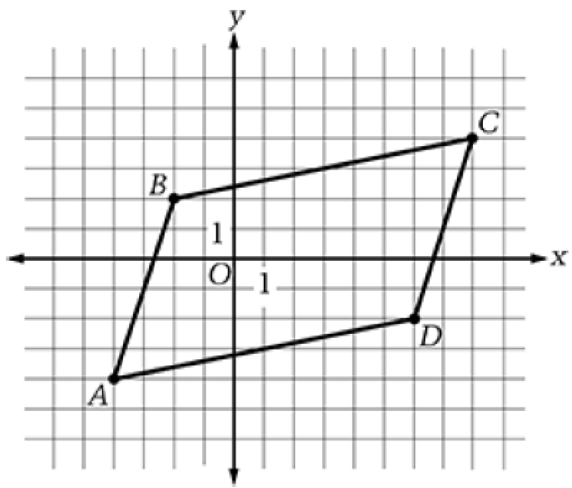
Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Geometry	Other	Geom – 3.h	SCR

As noted in Chapter 1 of this document, for the 2025 assessment, the "Mathematical Reasoning" subtopics in the 2009 Framework were removed. The intent of objectives in the 2009 Mathematical Reasoning subtopics was addressed in the 2025 Framework through additions to other subtopics and/or through the NAEP Mathematical Practices (see Chapter 3 for more on the NAEP Mathematical Practices). Therefore, the Objective ID for a 2025 item may differ from the original item Objective ID.

The item whose metadata are shown in Illustration 2.2a was administered on the 2009 NAEP Mathematics Assessment with Objective ID Geom – 5.a. (in 2009, the Framework included Mathematical Reasoning as the fifth subtopic). However, the 2025 Objective ID is Geom – 3.h. The wording of these objectives is the same across the 2009 and 2025 Frameworks (see Exhibit 2.4, p. 61, for the text of the objective).

Illustration 2.2b. Example: Original Objective ID and 2025 Objective ID Differ

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Geometry	Other	Geom – 3.h	SCR



In the figure above, the vertices of $ABCD$ are $A(-4, -4)$, $B(-2, 2)$, $C(8, 4)$, and $D(6, -2)$.

Give a mathematical justification that $ABCD$ is a parallelogram.

Scoring Information	
Key	[Scoring Rubric – for more information on scoring information, see Chapter 4]

This item was revised for the purpose of this document. The original version of this item appeared in the 2009 NAEP Mathematics Assessment with NAEP Item ID 2009-12M2 #12 M195001.

Another difference worth noting is the adjusting from grade 4 to grade 8 of some objectives in probability. In grade 4, a review of state and national mathematics standards indicated an absence of student opportunity to learn the content of probability objectives. Therefore, probability items originally developed for grade 4 may now be aligned to objectives that appear at grade 8 in the 2025 Framework. Illustration 2.3 gives an example.

Illustration 2.3. Example: Probability Objective Moved from Grade 4 to Grade 8

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Other	Data – 4.e	SCR – composite

Al, Bev, and Carmen are going on a ride at the park. Only 2 people can go on the ride at a time. They can pair up 3 different ways, as shown below.

Al and Bev

Al and Carmen

Bev and Carmen

Derek decides to join the group. How many different ways can the 4 students pair up?

Answer: ___________________

Show your work or explain how you got your answer.

Scoring Information	
Key	[Scoring Rubric – for more information on scoring information, see Chapter 4]

This item was revised for the purpose of this document. The original version of this item appeared in the 2013 NAEP Mathematics Assessment with NAEP Item ID 2013-4M6 #14 M136901.

Similar Objectives Across Multiple Grade Levels

Several concepts included in NAEP objectives span multiple grade levels. In this document, through the language used in the objectives or in additional notes for item development, the content is differentiated at each grade level. For example, Number Properties and Operations objective 1.i at each grade level involves ordering and comparing numbers. These objectives are shown in Illustration 2.4.

Illustration 2.4. Number Properties and Operations Objectives 1.i

Grade 4	Grade 8	Grade 12
i) Order or compare whole numbers, decimals, or fractions using common denominators or benchmarks.	i) Order or compare rational numbers including very large and small integers, and decimals and fractions close to zero.	i) Order or compare rational or irrational numbers, including very large and very small real numbers.

The objectives are worded similarly. The differences are in the types of numbers being compared. At grade 4, students compare whole numbers, decimals, or fractions; at grade 8 the sets of numbers are expanded to include rational numbers; and at grade 12 irrational numbers are included.

Specifications Added to Content Objective Exhibits

Exhibits for the content objectives from the *Mathematics Framework for the 2025 National Assessment of Educational Progress* have been augmented in this document (Exhibits 2.2, 2.3, 2.4, 2.5, and 2.6). Clarification on the presentation of these specifications is given here (e.g., meaning of frequently occurring phrases, the use of italics, the inclusion of indicator symbols like carat (^) and plus (+)).

Italicized text provides clarifications or limitations to inform item development. All such text is from the 2009 Assessment and Item Specifications document except for text that includes a leading symbol.

- leading carat (^) indicates edited text from the 2009 Assessment and Item Specifications
- leading plus (+) indicates text new to the 2025 Assessment and Item Specifications

Specifications related to wording of statements included as italicized text are described below.

- "Items should" and "Items should not" statements provide constraints and limitations for the assessment of the associated objectives.
- "Emphasis should be on" statements indicate characteristics of a majority of the items in the item pool for the associated objectives.
- Statements that indicate an item or other object of interest "can" be, do, or contain something indicate allowance for the described action or description. These include "For example" statements that provide examples of ways objectives might be assessed.
- "Include items that" statements indicate characteristics of at least some of the items in the item pool for the associated objectives.
- "See" statements refer the reader to a specific location in the chapter for additional information.

Many objectives and clarifications indicate developed items should have a context. At times the word "context" is modified by an adjective to provide specific information regarding the type of context required.

- "Real-world context" refers to situations that are concrete or have specific details related to human perception, activities, or relationships with the physical world. These specific details are necessary in order to understand or complete the item.
- "Mathematical context" refers to purely mathematical or abstract item settings that are not connected with students' everyday life experiences. In these cases, the mathematics is central to the item; the context may provide a setting for the mathematics but is often thin and does not need to be interpreted to solve the problem.
- "Familiar context" and "meaningful context" may be either a real-world context, a mathematical context, or a combination of the two. In these cases, students have experience with the context, or the context has meaning for the students.

The sources of these suggestions include the previous Assessment and Item Specifications (2009), 2025 NAEP Mathematics Framework authors, public-facing information from current state and national assessments (e.g., state assessment websites; SBAC, 2018; PARCC, 2015), related *Progressions* documents (https://www.math.arizona.edu/~ime/progressions/), mathematical modeling guidelines (Garfunkel & Montgomery, 2016), and preK–12 statistics guidelines (Franklin et al., 2007).

Mathematics Areas

Number Properties and Operations

Numbers (used as counts, measures, ratio comparisons, and scale values) are tools for describing the world quantitatively. It is thus not surprising that *Number* constitutes a major content focus of school mathematics, especially through grade 8. This focus includes facility with different notational forms (as whole numbers, fractions, decimals, percents, powers, and radicals), an understanding of number systems (e.g., integers, rational numbers, real numbers) and their properties, and calculational proficiency with these forms within systems.

Ancient cultures around the world had names for numbers and ways of doing arithmetic. The accessibility and usefulness of arithmetic today is greatly enhanced by the worldwide use of the Hindu-Arabic decimal place value system. In its full development, this remarkable system includes finite and infinite decimals that allow approximating any real number as closely as desired. Decimal notation simplifies arithmetic by means of routine algorithms; it makes size comparisons straightforward and estimation simple.

Numbers are not simply labels for quantities; they form systems with their own internal structure. For instance, at times problems can be more easily solved by considering what numbers add up to a certain value (e.g., 100 – 98 can be thought of as "98 plus what adds up to 100?"). Multiplication is connected to the idea of repeated addition just as division is connected to the idea of repeated subtraction and the relationship between multiplication and division can be used to simplify computation (e.g., instead of multiplying a number by 25, a number can be multiplied by 100 and then divided by 4, perhaps by halving and halving again). Arithmetic operations (addition, subtraction, multiplication, and division) and the relationships among them help students determine the mathematics that corresponds to basic real-world actions. For example, joining two collections or laying two lengths end-to-end can be described by addition, while comparing two collections can be described by subtraction, and 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. Another basic structure of real numbers is ordering, as in which is greater and lesser. Attention to the relative size of quantities provides a basis for making sensible estimates.

Number is not an isolated mathematics domain; it is intimately interwoven with other content strands. In their study of measurement, students use numbers to describe continuous quantities such as length, area, volume, weight, and time, and even to describe more complicated derived quantities such as rates of speed, density, inflation, interest, and so on. With numbers, students can count collections of discrete objects or describe fractional parts of data sets, allowing for statistical analysis. As elementary grade students generalize number relationships and properties they engage in algebraic thinking. In pursuit of graphical depictions of algebraic relationships, students use Cartesian coordinates—ordered pairs of numbers to identify points in a plane and ordered triples of numbers to label points in space. Numbers allow precise communication about anything that can be counted, measured, or located in space.

Comfort in dealing with numbers effectively is called *number sense*. It includes intuition about what numbers mean; understanding the ways to represent numbers symbolically (including facility with converting between different representations); the ability to calculate, either exactly or approximately, and by several methods (e.g., mentally, with paper and pencil, or calculator, as appropriate); and the ability to estimate. Skill in working with proportions (including percents) is another important part of number sense.

Number sense is a major expectation of the NAEP Mathematics Assessment. In grade 4, students are expected to have a solid grasp of whole numbers as represented in the base-10 system and to begin understanding fractions. By grade 8, students should be comfortable with rational numbers, represented either as decimal fractions or as common fractions, and should be able to use them to solve problems involving proportionality, percentages, and rates. At this level, numbers should also begin to coalesce with geometry by extending students' understanding of the number line. This concept is connected with approximation and the use of scientific notation. Grade 8 students should also have some acquaintance with naturally occurring irrational numbers, such as square roots and π (pi). By grade 12, students should be comfortable dealing with all types of real numbers and various representations, for example, as powers. Students in grade 12 should be able to establish the validity of numerical properties using mathematical arguments.

The 2025 Number Properties and Operations objectives are shown in Exhibit 2.2. Included with many of the objectives is italicized text providing clarifications or limitations for use during item development.

Exhibit 2.2. Number Properties and Operations (Num)

Num – 1. Number sense		
Grade 4	**Grade 8**	**Grade 12**
a) Identify place value and actual value of digits in whole numbers, and think flexibly about place value notions (e.g., there are 2 hundreds in 253, there are 25 tens in 253, there are 253 ones in 253). *+Items should limit numbers to whole numbers through 999,999.* *+Emphasis should be on numbers through 999.*	a) Use place value to represent and describe integers and decimals.	
b) Represent numbers using base 10, number line, and other representations. *+Items should limit numbers to whole numbers through 999.* *+Items should involve representations that students can use intuitively, without formal instruction or explanation of purpose or use (e.g., number lines, dots, tallies, base 10 blocks).*	b) Represent or describe rational numbers or numerical relationships using number lines and diagrams. *+For example, an item might require completion of a representation to show that a number and the opposite of the number are the same distance from 0 on a number line.*	

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 1. Number sense (continued)		
Grade 4	**Grade 8**	**Grade 12**
c) Compose or decompose whole quantities either by place value (e.g., write whole numbers in expanded notation using place value: 342 = 300 + 40 + 2 or 3 × 100 + 4 × 10 + 2 × 1) or convenience (e.g., to compute 4 × 27 decompose 27 into 25 + 2 because 4 × 25 is 100, and 4 × 2 is 8 so 4 × 27 is 108). *^Items should limit numbers to whole numbers through 999,999.* *+Emphasis should be on numbers through 999.* *+Emphasis should be on application of place value concepts as a way to express quantities.*		
d) Write or rename whole numbers (e.g., 10: 5 + 5, 12 – 2, 2 × 5). *+Items should limit numbers to whole numbers through 999,999.* *+Emphasis should be on numbers through 999.* *+Emphasis should be on multiple representations of a number using different operations.*	d) Write or rename rational numbers. *+For example, an item might involve writing a fraction as a decimal or a decimal as a fraction.* *+Decimals can be terminating or repeating.*	# d) Represent, interpret, or compare expressions for real numbers, including expressions using exponents and *logarithms. *^For example, an item might include expressions containing π or the square root of 2, or numerical relationships represented on a number line or with a diagram.* *^Exponents can be negative or fractional.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).
# Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 1. Number sense (continued)		
Grade 4	**Grade 8**	**Grade 12**
e) Connect across various representations for whole numbers, fractions, and decimals (e.g., number word, number symbol, visual representations). *+Items should involve representations that students can use intuitively, without formal instruction or explanation of purpose or use (e.g., number lines, dots, tallies, base 10 blocks).* *+For example, an item might include representation of a number on a number line or with an area diagram.*	e) Recognize, translate or apply multiple representations of rational numbers (fractions, decimals, and percents) in meaningful contexts. *+Items should avoid renaming of rational numbers as described in Number Properties and Operations objective 1.d.* *+For example, an item might situate a representation or multiple representations in context, such as a thermometer in a temperature-related item or a fuel gauge in a gas-related item.*	
	f) Express or interpret large numbers using scientific notation from real-life contexts. *+Items should present a number as a quantity or measurement.*	# f) Represent or interpret expressions involving very large or very small numbers in scientific notation. *^Exponents can be negative.* *^Include items that require interpreting calculator or computer displays given in scientific notation.*
	g) Find absolute values or apply them to problem situations. *+For example, an item might ask for the locations of a number and the absolute value of the number on a number line.* *+Include items that use absolute value to represent distance.*	g) Represent, interpret, or compare expressions or problem situations involving absolute values.

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 1. Number sense (continued)		
Grade 4	**Grade 8**	**Grade 12**
h) Recognize and generate simple equivalent (equal) fractions and explain why they are equivalent (e.g., by using drawings). *+Items should limit denominators of fractions to 2, 3, 4, 5, 6, 8, 10, 12, or 100.*	h) Order or compare rational numbers (fractions, decimals, percents, or integers) using various representations (e.g., number line). *+Include items that present values to be ordered or compared as quantities in familiar contexts.*	
i) Order or compare whole numbers, decimals, or fractions using common denominators or benchmarks. *+Items should involve ordering or comparing numbers of the same type (i.e., whole numbers, decimals, fractions), and limit numbers to:* • *whole numbers through 999,999;* • *fractions with denominators 2, 3, 4, 5, 6, 8, 10, 12, or 100; or* • *decimals to hundredths.*	i) Order or compare rational numbers including very large and small integers, and decimals and fractions close to zero. *+Include items that present one or more numbers in scientific notation.*	# i) Order or compare rational or irrational numbers, including very large and very small real numbers.

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 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., ½ and 0.5 may be used as benchmarks for fractions and decimals between 0 and 1.00). *+Items should limit benchmarks to numbers of the same type, using fraction benchmarks for fractions and decimal benchmarks for decimals.*	a) Establish or apply benchmarks for rational numbers and common irrational numbers (e.g., π) in contexts. *+Items can involve minimal context provided for the purpose of determining an appropriate benchmark.*	
b) Make estimates appropriate to a given situation with whole numbers, fractions, or decimals. *+Items should limit numbers to* • *whole numbers through 999,999;* • *fractions with denominators 2, 3, 4, 5, 6, 8, 10, 12, or 100; or* • *decimals to hundredths.*	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. *+Items should avoid estimation of square and cube roots as described in Number Properties and Operations objective 2.d.*	# 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. *+ Items should avoid estimation of square and cube roots as described in Number Properties and Operations objective 2.d.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 2. Estimation (continued)		
Grade 4	**Grade 8**	**Grade 12**
c) Verify and defend solutions or determine the reasonableness of results in meaningful contexts. *+Items should avoid estimation as described in Number Properties and Operations objective 2.b.* *+For example, an item might require justification for a whole number response based on the context used in division involving a remainder.*	c) Verify solutions or determine the reasonableness of results in a variety of situations, including calculator or computer results. *+Item should focus on solutions to and results from real-world and mathematical situations appropriate for grade 8 (e.g., determining the reasonableness of a calculation involving a whole number exponent).* *+Items should avoid estimation as described in Number Properties and Operations objectives 2.b and 2.d.*	# c) Verify solutions or determine the reasonableness of results in a variety of situations. *+Items should avoid estimation as described in Number Properties and Operations objectives 2.b and 2.d.* *^Include items that involve using estimation and order of magnitude to determine the reasonableness of technology-aided computations and interpreting results in terms of the context (e.g. verifying a computation involving numbers written in scientific notation).*
	d) Estimate square or cube roots of numbers less than 150 between two whole numbers. *^Items should limit numbers to whole numbers between perfect squares 1 through 144 or perfect cubes 1 through 125.* *+Items should not allow use of a calculator.*	d) Estimate square or cube roots of numbers less than 1,000 between two whole numbers. *+Items should limit numbers to whole numbers between perfect squares 1 through 900 or perfect cubes 1 through 729.* *+Items should not allow use of a calculator.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 3. Number operations		
Grade 4	**Grade 8**	**Grade 12**
a) Add and subtract using conventional or unconventional procedures (e.g., strategic decomposing and composing): • Whole numbers, or • Fractions and mixed numbers with like denominators. *+Items should limit numbers to whole numbers through 9,999 or fractions with denominators 2, 3, 4, 5, 6, 8, 10, 12, or 100.* *+Items that use a mathematical context should not allow use of a calculator.* *^Include items using a mathematical context that require computation with common and decimal fractions (decimals that can be written as a fraction with whole number numerators and denominators).*	a) Perform computations with rational numbers. *+Items that use a mathematical context should not allow use of a calculator.* *Include items that* • *^use a mathematical context and require computation with common and decimal fractions (decimals that can be written as a fraction with integer numerators and denominators).* • *^use a real-world context.* • *+require recognition of a numerical expression equivalent to a given numerical expression that allows for a friendlier computation (e.g., adding up to solve fraction subtraction problems, doubling and halving to solve fraction multiplication problems).* • *+require selection or creation of representations of a rational number computation (e.g., representing rational number division when the quotient is not a whole number).*	a) Find integer or simple rational powers of real numbers. *+Items that use a mathematical context should not allow use of a calculator.* *^For example, an item might require the evaluation of $27^{1/3}$.* *^Include items that involve numbers expressed with negative exponents.*

<table>
<tr><th colspan="3">Exhibit 2.2. Number Properties and Operations (continued)</th></tr>
<tr><th colspan="3">Num – 3. Number operations (continued)</th></tr>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td>b) Multiply numbers using conventional or unconventional procedures (e.g., strategic decomposing and composing):
Whole numbers no larger than two digits by two digits with paper and pencil computation, or
Larger whole numbers using a calculator, or
Multiplying a fraction by a whole number.
+Items presenting unconventional procedures should focus on an efficient procedure for multiplying based on the given factors.

+Items should limit denominators of fractions to 1, 2, 3, 4, 5, 6, 8, 10, 12, or 100.

^Multiplication problems involving decimal fractions (e.g., money) can be included on calculator blocks.</td><td></td><td>b) Perform arithmetic operations with real numbers, including common irrational numbers.

^Items should not include absolute value, which is addressed in Number Properties and Operations objective 3.c.

+Items that use a mathematical context should not allow use of a calculator.

^Include items that
use a mathematical context and require computation with common and decimal fractions (decimals that can be written as a fraction with integer numerators and denominators).
use a real-world context.
require application of order of operations.</td></tr>
<tr><td>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.</td><td></td><td>c) Perform arithmetic operations with expressions involving absolute value.</td></tr>
</table>

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 3. Number operations (continued)		
Grade 4	**Grade 8**	**Grade 12**
	d) Describe the effect of operations on size, including the effect of attempts to multiply or divide a rational number by: • Zero, or • A number less than zero, or • A number between zero and one, or • One, or • A number greater than one. *^For example, an item might ask about the effect of multiplying a fraction by a fraction less than one, or a fraction by a fraction greater than one.*	d) Describe the effect of multiplying and dividing by numbers including the effect of attempts to multiply or divide 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. *^For example, an item might ask about the effect of multiplying 2√3 by 1/2.*
e) Interpret, explain, or justify whole number operations and explain the relationships between them. *^Emphasis should be on interpreting, explaining, or justifying* • *subtracting a number as the inverse operation to adding a number, or* • *dividing by a number as the inverse operation to multiplying a number.*	e) Interpret, explain, or justify rational number operations and explain the relationships between them. *^Emphasis should be on interpreting, explaining, or justifying* • *the four operations (including additive and multiplicative inverses),* • *whole number square roots,* • *whole number cube roots, or* • *integer exponents.*	e) *Analyze or interpret a proof by mathematical induction of a simple numerical relationship. *+For example, an item might require proving that the sum of consecutive whole numbers from 0 to n can be determined using the expression n(n+1)/2.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 3. Number operations (continued)		
Grade 4	**Grade 8**	**Grade 12**
f) Solve problems involving whole numbers and fractions with like denominators. *+Items should avoid concepts assessed by Measurement objectives, such as determining the perimeter of a rectangle.* *+Include items that present contexts using a variety of addition/ subtraction problem structures (e.g., add to, take from, put together/ take apart, compare) and multiplication/ division problem structures (e.g., equal groups, arrays, area, compare).* *+Include items that require no more than three unique mathematical operations (addition, subtraction, multiplication, or division).* *^See Number Properties and Operations objectives 3.a, 3.b, and 3.c for number limitations and computation specifications.*	f) Solve problems involving rational numbers and operations using exact answers or estimates as appropriate. *+Items should avoid concepts assessed by Measurement or Geometry objectives, such as determining the volume of a cube.*	# f) Solve problems involving numbers, including rational and common irrationals. *+Items should avoid concepts assessed by Measurement or Geometry objectives, such as application of the Pythagorean Theorem or determining the volume of a cylinder.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 4. Ratios and proportional reasoning		
Grade 4	**Grade 8**	**Grade 12**
	a) Use ratios to describe problem situations. *+A ratio can be written a/b, a : b, or a to b.*	
	b) Use fractions to represent and express ratios and proportions. *+Include items that involve* • *ratios of whole numbers.* • *ratios of fractions.*	
	c) Use proportional reasoning to model and solve problems (including rates and scaling). *+Items should avoid scale drawings, which are addressed in Measurement objective 2.f.*	# c) Use proportions to solve problems (including rates of change and per capita problems). *^Items should avoid scale drawings, which are addressed in Measurement objective 2.f.*
	d) Solve problems involving percentages (including percent increase and decrease, interest rates, tax, discount, tips, or part/whole relationships).	# d) Solve multistep problems involving percentages, including compound percentages.

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 5. Properties of number and operations		
Grade 4	**Grade 8**	**Grade 12**
a) Identify odd and even numbers. *+Include items that involve determining whether the number of objects in a given set is even or odd.* *+Include items that involve writing an even number as the sum of two equal addends or as a sum of twos.*		
b) Identify factors of whole numbers. *^Items should involve identification of single-digit factors of whole numbers through 100.*	b) Recognize, find, or use factors, multiples, or prime factorization. *^Items should involve lowest common multiple, greatest common factor, or common multiples.* *^Items written for non-calculator blocks should use numbers less than 400.* *^Items written for calculator blocks should use numbers less than 1,000.*	
	c) Recognize or use prime and composite numbers to solve problems. *+Items can use a mathematical context or a real-world context.*	c) Solve problems using factors, multiples, or prime factorization. *+Items can use a mathematical context or a real-world context.* *^Include items that involve prime numbers.*
	d) Use divisibility or remainders in problem settings. *+Items should use a real-world context.* *+Items at grade 8 should be less complex than those developed at grade 12 (e.g., involve rational numbers).*	# d) Use divisibility or remainders in problem settings. *+Items should use a real-world context.* *+Items at grade 12 should be relevant to older students and may be more complex than those at grade 8 (e.g., involve irrational numbers).*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.2. Number Properties and Operations (continued)		
Num – 5. Properties of number and operations (continued)		
Grade 4	**Grade 8**	**Grade 12**
e) Apply basic properties of operations. *^Items should involve the commutative and associative properties of addition and multiplication, the distributive property of multiplication across addition, the identity property of addition, and multiplication by zero.* *+Items should not assess naming of properties.* *+Emphasis should be on properties rather than computation.* *+See Number Properties and Operations objectives 3.a and 3.b for number limitations and computation specifications.*	e) Apply basic properties of operations, including conventions about the order of operations as applied to integers and rational numbers. *^Items should involve the commutative and associative properties of addition and multiplication, the distributive property of multiplication across addition, the identity and inverse properties of addition and multiplication, and multiplication by zero.* *+Items should not assess naming of properties.* *+Emphasis should be on properties rather than computation with rational numbers.*	e) Apply basic properties of operations, including conventions about the order of operations as applied to real numbers. *^Items should involve the commutative and associative properties of addition and multiplication, the distributive property of multiplication across addition, the identity and inverse properties of addition and multiplication, and multiplication by zero.* *+Items should not assess naming of properties.* *^Emphasis should be on properties rather than computation with real numbers, including irrational numbers.*
		f) Recognize properties of the number system (whole numbers, integers, rational numbers, real numbers, and *complex numbers) and how they are related to each other and identify examples of each type of number. *^Items can include questions about identifying irrational numbers (e.g., Which number is irrational: 0.333, 0.333…, 3.14, √3?).*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Measurement

Measuring is the process by which numbers are assigned 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 banner, one may select the attribute of length and the inch as a unit for the comparison. In comparing lengths to the nearest inch, it may be that a length is *about* 42 inches. If considering only the domain of whole numbers, one would report that the banner is 42 inches long. However, because length is a continuous attribute, in the domain of rational numbers the length of the banner might be reported as $41^{13}/_{16}$ inches (to the nearest 16^{th} of an inch) or 41.8 inches (to the nearest 0.1 inch).

The connection between measuring and number makes measurement a vital part of school mathematics. Measurement is an important setting for negative and irrational numbers as well as positive numbers, since negative numbers arise naturally from situations with two directions and irrational numbers are commonplace in geometry. Measurement representations and tools are often used when students are learning about number properties and operations. For example, area grids and representations of volume using unit cubes can help students understand multiplication and its properties. The number line can help students understand ordering and rounding numbers. Measurement also has a strong connection to other areas of school mathematics and other subjects. Problems in algebra are often drawn from measurement situations and functions are used to relate measures to each other. Geometry regularly focuses on measurement aspects of geometric figures. Probability and statistics provide ways to measure chance and to compare sets of data. The measurement of time, values of goods and services, physical properties of objects, distances, and various kinds of rates exemplify the importance of measurement in everyday activities.

In the Framework, attributes such as capacity, weight, mass, time, and temperature are included, as are the geometric attributes of length, area, and volume. Many of these attributes appear in grade 4, where the emphasis is on length, including perimeter, distance, and height. At grade 4, students do not use formulas to determine area. Instead, they use informal or physical understandings (e.g., grids or blocks). More emphasis is placed on area and angle measure in grade 8. By grade 12, measurement in everyday life – as well as in the study of volumes and rates constructed from other attributes, such as speed – are emphasized.

The 2025 NAEP Mathematics Assessment includes nonstandard, customary, and metric units. At grade 4, common customary units such as inch, quart, pound, hour, and degree (for measuring angles) are included; and 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, continued use of 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. These are described later in this section.

Items classified in this content area depend on some knowledge of measurement. For example, an item comparing a 2-foot segment with an 8-inch line segment is classified as a measurement item; whereas an item that asks for the difference between a 3-inch and a 1¾ -inch line segment would be classified as a number item. In many secondary schools, measurement becomes an integral part of geometry and this is reflected in the proportion of items recommended for these two areas (see Exhibit 2.1).

The items in Illustrations 2.5 and 2.6 demonstrate the difference between a number item that involves units of measure and a measurement item. In the grade 4 item in Illustration 2.5, the context of weight is not necessary to determine the two consecutive whole numbers between which 12.4 lies. Since the focus of the item is comparing values, the item assesses a Number Properties and Operations objective.

Illustration 2.5. Example: A Number Properties and Operations Item Involving Units of Measure

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Other	Num – 1.i	SR – MC

A bag of potatoes weighs 12.4 pounds. Which of the following statements is true?

A. There are between 1 and 2 pounds of potatoes in the bag.

B. There are between 12 and 13 pounds of potatoes in the bag.

C. There are between 124 and 125 pounds of potatoes in the bag.

D. There are between 1,246 and 1,247 pounds of potatoes in the bag.

Scoring Information	
Key	B. There are between 12 and 13 pounds of potatoes in the bag.

This item was revised for the purpose of this document. The original version of this item appeared in the 2013 NAEP Mathematics Assessment with NAEP Item ID 2013-4M6 #3 M135801.

In the grade 12 NAEP released item in Illustration 2.6, a measurement context is the focus of the item. That is, the accuracy of the measurements used forms the foundation of the item and must be considered when determining the range of measurements for the area of the room. Therefore, this item assesses a Measurement objective.

Illustration 2.6. Example: An Item with a Measurement Focus

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Measurement	Other	Meas – 2.e	SR – MC

Carlene told Kyle that a rectangular room measured 16 feet by 12 feet, to the nearest foot. This means that the length could measure between 15.5 feet and 16.5 feet and the width could measure between 11.5 feet and 12.5 feet.

Kyle performed the following calculations.

Dimensions (feet)	Area (square feet)
15 by 11	165
15.5 by 11.5	178.25
16 by 12	192
16.5 by 12.5	206.25
17 by 13	221

Of the following intervals, which is the smallest interval that contains all possible values of the area of the room?

A. Between 191.5 and 192.5 square feet

B. Between 191 and 193 square feet

C. Between 179 and 206 square feet

D. Between 178 and 207 square feet

E. Between 165 and 221 square feet

Scoring Information	
Key	D. Between 178 and 207 square feet

This item was revised for the purpose of this document. The original version of this item appeared in the 2009 NAEP Mathematics Assessment with NAEP Item ID 2009-12M2 #10 M176801.

General Guidelines for Measurement

This section describes specifications common to many of the measurement objectives. *Any attribute, unit, instrument, conversion factor, or formula included in a list at a lower grade level is also appropriate for a higher grade level.*

Attributes

Attributes used in items are cumulative and listed below.

- Grade 4: perimeter, height, and distance, time, temperature, capacity, weight or mass, area, and angle measure. Item content should emphasize **length** (measures of length include measures of perimeter, height, and distance).
- Grade 8: all attributes listed for grade 4, surface area, and volume. Item content should emphasize **area**. Attributes such as speed, measured in terms of the attributes of time and distance, are also appropriate.
- Grade 12: all attributes listed for grades 4 and 8. Item content should emphasize area, surface area, and volume. Rates constructed from other attributes such as speed or flow rate are appropriate.

<u>Units</u>

Units used in items are cumulative and listed below.

- Grade 4: non-standard units, common customary units (inch, foot, mile, cup, quart, gallon, pound, hour, minute, day, year, degrees of measured angles, degrees Fahrenheit) and metric units (centimeter, millimeter, meter, liter, gram, degrees Celsius) for the allowed attributes at this grade level.
- Grade 8: all units listed for grade 4 and square units, cubic units, and constructed units such as miles per hour; additional customary units (yard, fluid ounce, pint, ounce, ton) and additional metric units (kilometer, kilogram) for the attributes at this grade level.
- Grade 12: all units listed for grades 4 and 8 for the attributes at this grade level.

<u>Instruments</u>

The following measurement instruments are commonly found in curricula. Variations based on the same principles could be used during item development (e.g., graduated cup measures).

- All grades: ruler, clock, thermometer, graduated cylinder, balance scales, scales, protractor.

<u>Conversions</u>

Equivalencies that should be known by students and not provided in items are cumulative and listed below. All other conversions should be provided.

- Grade 4: feet/inches, hours/minutes, and meters/centimeters.
- Grade 8: square and cubic unit conversions, common time equivalences, and all common metric equivalences.
- Grade 12: conversions involving 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 at least one method for determining the perimeter and at least one method for determining the area of a rectangle. That is, students are expected to know that the perimeter of a rectangle can be determined by adding the lengths of all of its sides, but they do not need to know the formula $P = 2l + 2w$. Additionally, students can determine the area of a rectangle by tiling it with unit squares, without gaps or overlap, then counting the number of unit squares or by multiplying the length and the width, but they do not need to know the formula $A = l \cdot w$.

Both grade 8 and grade 12 students should know formulas for the areas of a rectangle, triangle, and circle, the circumference of a circle, and the volumes of a cylinder and rectangular solid. When other formulas are needed to complete an item, they should be given. See General Guidelines for Geometry for more information about formulas for area, circumference, and volume.

The 2025 Measurement objectives are shown in Exhibit 2.3. Included with many of the objectives is italicized text providing clarifications or limitations for use during item development.

Exhibit 2.3. Measurement (Meas)

Meas – 1. Measuring physical attributes		
Grade 4	**Grade 8**	**Grade 12**
a) Identify the attribute that is appropriate to measure in a given situation. *+See the General Guidelines for Measurement for clarifications and limitations on attributes used in items.*		
b) Compare objects with respect to a given attribute, such as length, area, capacity, time, or temperature. *+Items involving area should avoid computing areas as described by Measurement objective 1.g.* *+See the General Guidelines for Measurement for clarifications and limitations on attributes used in items.*	b) Compare objects with respect to length, area, volume, angle measurement, weight, or mass. *+See the General Guidelines for Measurement for clarifications and limitations on attributes used in items.*	# b) Determine the effect of proportions and scaling on length, area, and volume. *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*
c) Estimate the size of an object with respect to a given measurement attribute (e.g., length, perimeter, or area using a grid). *+For example, an item might require estimating the area of an irregular shape presented on a grid.* *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*	c) Estimate the size of an object with respect to a given measurement attribute (e.g., area). *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*	# c) Estimate or compare perimeters or areas of two-dimensional geometric figures. *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.3. Measurement (continued)		
Meas – 1. Measuring physical attributes (continued)		
Grade 4	**Grade 8**	**Grade 12**
		d) Solve problems of angle measure, including those involving triangles or other polygons or parallel lines cut by a transversal. *^Items should assume that students know* • *the sum of the measures of the interior angles of a triangle is 180º, and* • *the relationships among the measures of 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. *+"Other scaled instruments" may include a protractor.* *+See the General Guidelines for Measurement for clarifications on measurement instruments used in items.*	e) Select or use appropriate measurement instruments to determine or create a given length, area, volume, angle, weight, or mass. *+See the General Guidelines for Measurement for clarifications on measurement instruments used in items.*	
f) Solve problems involving perimeter of plane figures. *+Plane figures can be polygons but cannot be circles.* *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.*	f) Solve mathematical or real-world problems involving perimeter or area of plane figures such as triangles, rectangles, circles, or composite figures. *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.*	f) Solve problems involving perimeter or area of plane figures such as polygons, circles, or composite figures. *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.*

Exhibit 2.3. Measurement (continued)		
Meas – 1. Measuring physical attributes (continued)		
Grade 4	**Grade 8**	**Grade 12**
g) Solve problems involving area of squares and rectangles. *+Items should use measurements and right-angle markings, as appropriate, when art includes squares or rectangles.* *+Items should not require a formula but should assume that students know at least one method for determining the area of a square or rectangle.* *+Include items that relate area to the operations of multiplication and addition, such as tiling a rectangle with whole-number side lengths and showing that the area is the same as would be found by multiplying the side lengths.* *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.*		
	h) Solve problems involving volume or surface area of rectangular solids, and volume of right cylinders and prisms, or composite shapes. *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.* *+See the General Guidelines for Geometry for clarifications and limitations on the formulas items should assume students know or can use.*	# h) Solve problems by determining, estimating, or comparing volumes or surface areas of three-dimensional figures. *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.* *+See the General Guidelines for Geometry for clarifications and limitations on the formulas items should assume students know or can use.*
	i) Solve problems involving rates such as speed or ratios such as population density. *+See the General Guidelines for Measurement for clarifications and limitations on units used in items.*	# i) Solve problems involving rates and ratios such as speed, density, population density, or flow rates. *+ See the General Guidelines for Measurement for clarifications and limitations on units used in items.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.3. Measurement (continued)		
Meas – 2. Systems of measurement		
Grade 4	**Grade 8**	**Grade 12**
a) Select or use an appropriate type of unit for the attribute being measured such as length, angle size, time, or temperature. *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*	a) Select or use an appropriate type of unit for the attribute being measured such as length, area, angle, time, or volume. *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*	# a) Choose appropriate units for geometric measurements (length, area, perimeter, volume) and apply units in expressions, equations, and problem solutions. *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*
b) Solve problems involving conversions within the same measurement system such as conversions involving inches and feet or hours and minutes. *+Emphasis should be on conversions of measurements from a larger unit to a smaller unit.* *^Items can include additional conversions given the conversion information (e.g., 1 quart = 2 pints).* *+See the General Guidelines for Measurement for conversions that should be known and not provided.*	b) Solve problems involving conversions within the same measurement system such as conversions involving square inches and square feet. *+See the General Guidelines for Measurement for conversions that should be known and not provided.*	# b) Solve problems involving conversions within or between measurement systems, given a relationship between the units. *^Conversions can include cubic units and compound rates such as miles per hour to feet per second.* *^See the General Guidelines for Measurement for conversions that should be known and not provided.*
	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: 1 kilometer is approximately 0.6 mile. • Money: U.S. dollars to Canadian dollars. • Temperature: Fahrenheit to Celsius.	

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.3. Measurement (continued)		
Meas – 2. Systems of measurement (continued)		
Grade 4	**Grade 8**	**Grade 12**
d) Determine appropriate unit of measurement in problem situations involving such attributes as length, time, capacity, or weight. *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*	d) Determine appropriate unit of measurement in problem situations involving such attributes as length, area, or volume. *+See the General Guidelines for Measurement for clarifications and limitations on attributes and units used in items.*	# d) Understand that numerical values associated with measurements of physical quantities are approximate, subject to variation, and must be assigned units of measurement. *+See the General Guidelines for Measurement for limitations on units used in items.*
		# 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, an item might ask for the range within which the actual area of a rectangle could be if the side lengths of the rectangle measured to the nearest inch are 3 inches and 5 inches.*
	f) Construct or solve problems (e.g., floor area of a room) involving scale drawings. *+Include items that involve* • *computing actual lengths and areas from a scale drawing.* • *reproducing a scale drawing at a different scale.*	# f) Construct or solve problems involving scale drawings. *^For example, an item might require determination of the number of rolls of insulation needed for insulating a house.* *^A scale drawing can be excluded from the item stem.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.3. Measurement (continued)		
Meas – 3. Measurement in triangles		
Grade 4	**Grade 8**	**Grade 12**
		# a) Solve problems involving indirect measurement. *^For example, an item might require determining the height of a building using the distance to the base of the building and the angle of elevation to the top of the building.*
		b) Solve problems using the fact that trigonometric ratios (sine, cosine, and tangent) stay constant in similar triangles. *^For example, an item might ask 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. *^Items should assume that students know* • *the definitions of sine, cosine, and tangent, and* • *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. *^Items should assume that students know that $\sin^2\theta + \cos^2\theta = 1$.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).
# Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.3. Measurement (continued)		
Meas – 3. Measurement in triangles (continued)		
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. *^Items should limit angle measures 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. *^Items should provide relevant trigonometric formulas (e.g., law of cosines, double-angle formula).* *^For example, an item might require an explanation for whether or not sin20º and 2sin10º are equivalent.*
		g) * Use the law of cosines and the law of sines to find unknown sides and angles of a triangle. *^Items should provide relevant trigonometric formulas (e.g., law of cosines, double-angle formula).*
		h) * Interpret the graphs of the sine, cosine, and tangent functions with respect to periodicity and values of these functions for multiples of π/6 and π/4.

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Geometry

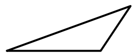
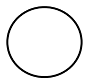
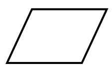
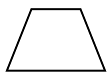
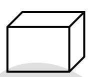
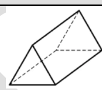
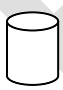
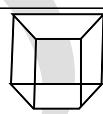
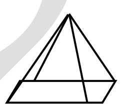
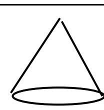
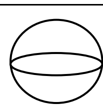
Geometry began thousands of years ago in many lands as sets of practical rules related to describing and predicting locations of astronomical objects, calculating land areas, and building structures. More than 2200 years ago, the Greek mathematician Euclid organized the geometry known at that time into a coherent collection of results, all deduced using logic from a small number of postulates assumed to be true. Euclid's work was fundamental in establishing mathematical truth as dependent on valid deductive reasoning rather than reliant on educated guesses from several specific examples. The theorems obtained via deduction by Euclid remain fundamental to the study of geometry, and for this reason the geometry studied in school is called Euclidean geometry.

The fundamental concepts of Euclidean geometry are congruence, similarity, and symmetry. 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, squares, and rectangles) and in space (cubes, spheres, and cylinders).

By grade 8, 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. Reflections, translations, and rotations (mathematical models of the physical phenomena of reflecting, sliding, and turning) are introduced as distance-preserving transformations that map a figure onto a congruent image. Dilatations (expansions and contractions) map figures onto similar images. Properties of congruent and similar figures involve angle measures and lengths, so geometry becomes more and more mixed with measurement in later grades. Placing figures on a coordinate plane provides the beginnings of the connections among algebra, geometry, and analytic geometry.

In secondary school, the content of plane geometry is logically ordered and students are expected to make, test, and validate conjectures. Students see that most of the commonly-studied plane figures – triangles (scalene, isosceles, equilateral) and quadrilaterals (parallelogram, rectangle, rhombus, square, trapezoid) – may possess reflection or rotation symmetry, or both, and can use triangle congruence and similarity theorems as well as symmetry to establish properties of figures. By grade 12, students may also gain insight into systematic structure, such as the classification of distance-preserving transformations of the plane (that is, reflections, rotations, translations, or glide reflections), and what happens when two or more isometries are performed in succession (composition). In analytic geometry, the key areas of geometry and algebra merge into a powerful tool that provides a basis for calculus and much of applied mathematics.

General Guidelines for Geometry
This table provides expectations for knowledge of geometric formulas at each grade level.

Shape		Formulas for Area and Circumference		
		Grade 4	Grade 8	Grade 12
Rectangle		find area without a formula	expected to know the formula	expected to know the formula
Triangle		not tested	expected to know the formula	expected to know the formula
Circle		not tested	expected to know the formula	expected to know the formula
Parallelogram		not tested	formula should be provided	formula should be provided
Trapezoid		not tested	formula should be provided	formula should be provided
Figure		**Formulas for Volume and Surface Area**		
		Grade 4	**Grade 8**	**Grade 12**
Rectangular Prism		not tested	expected to know the formula	expected to know the formula
Triangular Prism		not tested	expected to know the formula	expected to know the formula
Right Circular Cylinder		not tested	expected to know the formula for volume only	expected to know the formula
General Prisms		not tested	formula should be provided	formula should be provided
Square Pyramid		not tested	formula should be provided	formula should be provided
Right Circular Cone		not tested	formula should be provided	formula should be provided
Sphere		not tested	formula should be provided	formula should be provided

The 2025 Geometry objectives are shown in Exhibit 2.4. Included with many of the objectives is italicized text providing clarifications or limitations for use during item development.

Exhibit 2.4. Geometry (Geom)

Geom – 1. Dimension and shape		
Grade 4	**Grade 8**	**Grade 12**
a) 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, an item might require identification of rectangles in a picture of a room.*	a) Identify a geometric object given a written description of its properties. *+Items should include geometric objects appropriate to grade 8, such as polygons, composite shapes, and right pyramids, prisms, and cones.*	
b) Identify or draw angles and other geometric figures in the plane. *+Geometric figures can include points, lines, line segments, rays, polygons, and circles.*	b) Identify, define, or describe geometric shapes in the plane and in three-dimensional space given a visual representation. *^Items should be more complex than those presented at grade 4, such as those involving geometric shapes and figures composed of triangles, quadrilaterals, polygons, cubes, or right prisms.*	b) Give precise mathematical descriptions or definitions of geometric shapes in the plane and in three-dimensional space. *^Three-dimensional shapes can include the full set of Platonic solids (e.g., cube, regular tetrahedron).*
	c) Draw or sketch from a written description polygons, circles, or semicircles.	c) 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.*

Exhibit 2.4. Geometry (continued)		
Geom – 1. Dimension and shape (continued)		
Grade 4	**Grade 8**	**Grade 12**
		# d) Use two-dimensional representations of three-dimensional objects to visualize and solve problems. *+Items should involve three-dimensional objects composed of triangles, rectangles, and/or circles (e.g., net of a cylinder in a context about packages of oatmeal).*
e) Describe or distinguish among attributes of two- and three-dimensional shapes. *+Items should focus on countable or defining attributes, such as number of sides or number of right angles, and should avoid concepts assessed by Measurement objectives, such as determining perimeter or area.* *+For example, an item might require identification of characteristics that all rectangles have in common.*	e) Demonstrate an understanding of two- and three-dimensional shapes in the world through identifying, drawing, reasoning from visual representations, composing, or decomposing. *+For example, an item might involve use of a cylinder to represent a construction barrel, or recognition that a cube can be decomposed into four same-sized pyramids or three non-congruent pyramids having equal volumes.*	# e) Analyze properties of three-dimensional figures including prisms, pyramids, cylinders, cones, spheres and hemispheres. *+Items should avoid explicitly requiring the volume or surface area of a prism, pyramid, cylinder, cone, sphere, or hemisphere, but may require analysis of a familiar object to determine it has properties similar to one of the named figures.* *+For example, an item might require an informal argument for the formula for the volume of a cylinder, the volume of a pyramid, or the volume of a cone.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.4. Geometry (continued)		
Geom – 2. Transformation of figures and preservation of properties		
Grade 4	**Grade 8**	**Grade 12**
	a) Identify lines of symmetry in plane figures or recognize and classify types of symmetries of plane figures. *^Items should involve point, line, and rotational symmetry.*	a) Recognize or identify types of symmetries (e.g., translation, reflection, rotation) 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) Recognize or informally describe the effect of a transformation (reflection, rotation, translation, or dilation) on two-dimensional figures. *+For example, an item might require recognition that any transformation takes a line segment to a line segment, but that the type of transformation determines whether the line segments have the same length.*	c) Perform or describe the effect of a single transformation (reflection, rotation, translation, or dilation) on two- or three-dimensional geometric figures. *+Items can involve more than one application of a single type of transformation (e.g., viewing of the image of a reflection of an image in a mirror).*
d) Recognize attributes (such as shape and area) that do not change when plane figures are subdivided and rearranged. *+Items should limit plane figures to those composed of triangles and rectangles.* *+Items can involve subdividing while maintaining the original shape.*	d) Predict results of combining, subdividing, and recombining shapes of plane figures and solids (e.g., paper folding, tiling, subdividing and rearranging the pieces).	d) Identify transformations 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.*

Exhibit 2.4. Geometry (continued)		
Geom – 2. Transformation of figures and preservation of properties (continued)		
Grade 4	**Grade 8**	**Grade 12**
	e) Justify relationships of congruence and similarity and apply these relationships using scaling and proportional reasoning. *^Items should limit figures to those in two-dimensions.*	e) Justify relationships of congruence and similarity and apply these relationships using scaling, proportional reasoning, and established theorems. *^Items should allow for a variety of forms of proof (e.g., flow diagrams, paragraph, two-column proofs).* *^Proofs can include standard SAS, SSS, or ASA congruence proofs with corresponding parts.* *^Include items that* • *apply scaling and proportional reasoning to two-dimensional figures.* • *apply scaling and proportional reasoning to three-dimensional figures.* • *ask for justifications less formal than proofs of established theorems (e.g., giving reasons why figures are congruent or similar).*
	f) Apply the relationships among angle measures, lengths, and perimeters among similar figures. *^Emphasis should be on right triangles and quadrilaterals.*	f) Apply the relationships among angle measures, lengths, perimeters and volumes among similar figures. *+For example, an item might present two similar triangles with the necessary measures and require determining a missing angle measure or side length in one of the triangles.*

Exhibit 2.4. Geometry (continued)		
Geom – 2. Transformation of figures and preservation of properties (continued)		
Grade 4	**Grade 8**	**Grade 12**
		g) Perform or describe the effects of successive (composites of) isometries and/or similarity transformations. *+Items should be limited to transformations on one-dimensional geometric objects, two-dimensional geometric shapes, or three-dimensional geometric figures.* *+Items should avoid transformations on algebraic representations as described in Algebra objective 2.d.* *+For example, an item might require the selection of a different set of transformations that have the same result as a series of three reflections over three parallel lines.*

Exhibit 2.4. Geometry (continued)		
Geom – 3. Relationships between geometric figures		
Grade 4	**Grade 8**	**Grade 12**
a) Analyze or describe patterns in polygons when the number of sides increases, or the size or orientation changes.		
b) Combine simple plane shapes to construct a given shape. *+Include items that involve combining two-dimensional shapes to construct a three-dimensional figure.*	b) Apply geometric properties and relationships in solving problems in two and three dimensions. *^Items should limit figures to parallel and perpendicular lines, triangles, quadrilaterals, circles, cylinders, and cones.* *^Include items that involve properties of geometric similarity, congruence, and angle sum.* *^Include items that involve angle relationships and transversal properties of quadrilateral angles.*	b) Apply geometric properties and relationships to solve problems in two and three dimensions. *+Items should avoid concepts assessed by Measurement objectives, such as determining the volume of a composite figure.* *^Emphasis should be on solving problems.* *Problems can involve multiple steps.* *^Figures can include parallel and perpendicular lines, triangles (including triangles with angle measures of 45-45-90 and 30-60-90), cylinders, cones, prisms, and pyramids.*
c) Recognize two-dimensional faces of three-dimensional shapes.	c) Represent problem situations with geometric figures to solve mathematical or real-world problems. *+Emphasis should be on grade-level appropriate representations or figures.*	# c) Represent problem situations with geometric figures to solve mathematical or real-world problems. *+Items should be more complex than grade 8 items. For example, grade 12 items might involve more figures, or more properties than grade 8 items.* *^Emphasis should be on representations or figures.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.4. Geometry (continued)		
Geom – 3. Relationships between geometric figures (continued)		
Grade 4	**Grade 8**	**Grade 12**
	d) Use the Pythagorean theorem to solve problems in two-dimensional situations. *^Items should assume that students know the Pythagorean theorem.* *+Items can use a real-world context.* *+Include items that involve application of the Pythagorean Theorem to determine the distance between two points.*	# d) Use the Pythagorean theorem to solve problems in two- or three-dimensional situations. *^Items should assume that students know the Pythagorean theorem.*
		e) Recall and interpret or use definitions and basic properties of congruent and similar triangles, circles, quadrilaterals, other polygons, parallel, perpendicular and intersecting lines, and associated angle relationships (e.g., in solving problems or creating proofs). *^Emphasis should be on direct application of definitions or defining properties of lines, angles, and shapes.*
f) Describe and compare properties of simple and compound figures composed of triangles, squares, and rectangles. *+For example, an item might provide a rectangular prism, and require identification of the faces that have the same area.*	f) Describe, compare or analyze attributes of, or relationships between, triangles, quadrilaterals, and other polygonal plane figures. *+Items should avoid situations in which the definition of a trapezoid must be assumed.*	f) Analyze attributes or relationships of triangles, quadrilaterals, and other polygonal plane figures. *^Items should avoid situations in which the definition of a trapezoid must be assumed.* *^Emphasis should be on examining figures, identifying their properties, and applying identified properties.* *^Figures can include rhombi, parallelograms, and trapezoids.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.4. Geometry (continued)		
Geom – 3. Relationships between geometric figures (continued)		
Grade 4	**Grade 8**	**Grade 12**
	g) Describe or analyze properties and relationships of parallel or intersecting lines. *+For example, an item might present a pair of parallel lines cut by a transversal and require identification of the angles that have the same measure.*	g) Analyze properties and relationships of parallel, perpendicular, or intersecting lines including the angle relationships that arise in these cases. *^Emphasis should be on examining lines and angles, identifying their properties, and applying identified properties.*
		h) Make, test, and validate geometric conjectures using a variety of methods, including deductive reasoning and counterexamples.
		i) * Analyze properties of circles and the intersections of lines and circles (inscribed angles, central angles, tangents, secants, and chords). *^For example, an item might ask about measures of angles inscribed in a semicircle, or the relationship between tangents, secants, chords, and radii.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.4. Geometry (continued)		
Geom – 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 using distance between two points, the midpoint of a segment, or slopes of perpendicular or parallel lines. *+Items should avoid application of the Pythagorean Theorem as described in Geometry objective 3.d.* *+Items should avoid concepts assessed by Algebra objectives, such as determining the equation of a line through two points.* *+For example, an item might involve determining the slope of a line given two points or given the slope of a line to which it is perpendicular.*
	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. *^Items should involve cross-sections of standard, familiar solids such as a sphere, cylinder, or rectangular solid.*	c) Describe or identify conic sections and other cross sections of solids. *^Items should involve cross-sections of standard, familiar solids such as a cone, sphere, or cylinder, and of Platonic solids such as a cube or regular tetrahedron.*
	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.

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.4. Geometry (continued)		
Geom – 4. Position, direction, and coordinate geometry (continued)		
Grade 4	**Grade 8**	**Grade 12**
		f) Find an equation of a circle given its center and radius and, given an equation of a circle, find its center and radius. *^Items should assume that students know the equation of a circle.* *^Include items that require the derivation of the center or radius of a circle.*
		g) * Graph or determine equations for images of lines, circles, parabolas, and other curves under translations and reflections in the coordinate plane. *^Items should provide the formulas for ellipses and hyperbolas in standard form.* *^Items should not require knowledge of technical characteristics of these functions (e.g., equations of asymptotes or foci).* *^Items can require knowledge of general characteristics of these functions (e.g., drawing a graph).*
		h) * Represent situations and solve problems involving polar coordinates.

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Data Analysis, Statistics, and Probability

Data analysis and statistics refers to the entire process of collecting, organizing, summarizing, and interpreting data. This is the heart of statistics and is in evidence whenever quantitative information is used to determine a course of action. Data analysis normally begins with a question to be answered. Statistical questions can arise prior to data collection, or from existing data sets. Beginning at an early age, students should grasp the fundamental principle that exploratory data analysis of an existing data set is far different from the scientific method of collecting data to verify or refute a well-posed question. Data can be useful when collected with a specific question in mind and when there is a plan (usually called a *design*) for using the data to answer the question. However, contemporary uses of data-mining techniques associated with "big data" suggest that data sets may subsequently be useful in answering questions that were not envisioned when the data collection was initiated.

A probability is a measure of uncertainty. This measure may be determined from a theoretical model that makes assumptions about equally likely or weighted outcomes for an event (as when one says that the probability of a coin landing head-side up is one-half) or it may be determined in some way from past experience, as when forecasters say the probability of rain tomorrow is 40 percent. Statistical analysis often involves studying whether assumptions about theoretical probability match observed relative frequencies. For instance, if a coin tossed 100 times turned up heads 80 times, one might suspect that the probability of heads for that coin is not ½ (the theoretical probability of heads for a fair coin). Under random sampling, patterns for outcomes of designed studies can be anticipated and used as a basis for making decisions. The probability distribution of all possible outcomes is important in most statistical decision-making because the key is to decide whether or not a particular observed outcome is typical or unusual (located in a tail of a probability distribution). For example, 4.0 as a grade-point average is unusually high among most student groups, 4 as the weight in pounds of a human baby is unusually low, and 4 as the number of floors in a building is not unusual in either direction.

By grade 4, students are expected to apply their understanding of number and quantity to consider questions that can be answered by examining appropriate data. Building on the principles of describing data distributions through minimum, maximum, and clusters of values, grade 8 students are expected to use a wider variety of organizing and summarizing techniques for center, spread, and shape. They can identify and construct a statistical question, one that needs data in order to be addressed. They can also begin to analyze statistical claims through designed surveys and experiments that involve randomization. Also by grade 8, students are expected to begin to use more formal terminology related to probability and data analysis. They can identify associations between two numerical variables in scatterplots, as well as the relative strength of those associations.

Grade 12 students are expected to use a wide variety of statistical techniques for all phases of data analysis, including a more formal understanding of statistical inference, and simulation as an inferential analysis tool. In addition to comparing univariate data sets, students at this level can recognize and describe possible associations between two variables by looking at two-way tables for categorical variables or scatterplots for measurement variables. By grade 12, students should be able to use linear equations to describe possible associations between measurement variables and should be familiar with techniques for fitting functions to data.

Implications of Updates to Data Analysis, Statistics, and Probability Objectives

As mentioned in Chapter 1, a re-examination of statistics, data analysis, and probability concepts and skills in light of current scholarship and content of standards documents led to significant changes in the objectives for this content area at grade 4. Along with the decrease in the number of Data Analysis, Statistics, and Probability objectives, the phrasing of objectives has changed. Illustration 2.7 compares wording for an objective in grade 4 that was revised.

Illustration 2.7. Grade 4 Data Analysis, Statistics, and Probability Objective 2.b

Objective	2009 Wording	2025 Wording
2.b	Given a set of data or a graph, describe the distribution of data using median, range, or mode.	Given a distribution of whole number data in a context, identify and explain the meaning of the greatest value, the least value, or of any clustering or grouping of data in the distribution.

The composite item in Illustration 2.8 shows two ways in which objective 2.b can be assessed in grade 4. The item is adapted from England's *Key Stage 2, Paper 3: Reasoning* and contains material developed by the Standards and Testing Agency for 2019 national curriculum assessment, licensed under Open Government Licence v3.0. (Key Stage 2 students are 7 to 11 years old.)

Illustration 2.8. Example: Item Aligning to Grade 4 Objective 2.b

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Data Analysis, Statistics, and Probability	Other	Data – 2.b	SCR – composite

This chart shows the masses of eight kittens in grams (g).

305 g	375 g	310 g	255 g
275 g	410 g	360 g	345 g

A. Complete these sentences about the masses of the kittens.

The mass of the lightest kitten is [305 / 255 / 275 / 345] grams.

The mass of the heaviest kitten is [375 / 310 / 410 / 360] grams.

B. Complete this table to put the masses of the kittens into four groups.

Mass (grams)	Number of kittens
250 – 299	
300 – 349	
350 – 399	
400 – 449	

<table>
<tr><th colspan="2">Scoring Information</th></tr>
<tr><td>Part A</td><td>255; 410</td></tr>
<tr><td>Part B</td><td>
<table>
<tr><th>Mass (grams)</th><th>Number of kittens</th></tr>
<tr><td>250 – 299</td><td>2</td></tr>
<tr><td>300 – 349</td><td>3</td></tr>
<tr><td>350 – 399</td><td>2</td></tr>
<tr><td>400 – 449</td><td>1</td></tr>
</table>
</td></tr>
</table>

This item was revised for the purpose of this document. The original version of this item appeared in the 2019 England *Key Stage 2* as Item 7.

General Guidelines for Data Analysis, Statistics, and Probability

This section describes additional specifications for data representations used in items at each grade level.

- Limitations on representations of data are indicated for each grade level in Exhibit 2.5. Objectives in which only a subset of these representations is applicable are indicated in the parenthesis associated with the objective.
- Items should include interpretation of a variety of less common representations of data, such as those found in newspapers and magazines.
- Bar graphs and plots over time (line graphs) should increase in complexity (e.g., through using more complex scales and greater number of categories) from grade to grade.
- Descriptions of data sets at grade 4 may be informal.

The 2025 Data Analysis, Statistics, and Probability objectives are shown in Exhibit 2.5. Included with many of the objectives is italicized text providing clarifications or limitations for use during item development.

Exhibit 2.5. Data Analysis, Statistics, and Probability (Data)

Data – 1. Data representation		
Grade 4	**Grade 8**	**Grade 12**
Representations of data are indicated for each grade level. Objectives in which only a subset of the representations is applicable are indicated in the parenthesis in the objective.		
Pictographs, bar graphs, dot plots, tables, and tallies.	Histograms, plots over time, dot plots, scatterplots, box plots, bar graphs, circle graphs, stem and leaf plots, frequency distributions, and tables.	Histograms, plots over time, dot plots, scatterplots, box plots, bar graphs, circle graphs, stem and leaf plots, frequency distributions, and tables, including two-way tables.
a) Read or interpret a single distribution of data. *+Representations of data can be graphical or tabular.*	a) Read or interpret data, including interpolating or extrapolating from data. *+Representations of data can be graphical or tabular.*	# a) Read or interpret graphical or tabular representations of data.
b) For a given distribution of data, complete a graph (limits of time make it difficult to construct graphs completely).	b) For a given distribution of data, complete a graph and solve a problem using the data in the graph (histograms, plots over time, dot plots, scatterplots, bar graphs, circle graphs).	# b) For a given set of data, complete a graph and solve a problem using the data in the graph (histograms, plots over time, dot plots, scatterplots). *+Items should involve a single data set and a single data representation.*
c) Answer statistical questions by estimating and computing within a single distribution of data.	c) Answer statistical questions by estimating and computing with data from a single distribution or across distributions of data.	c) Answer statistical questions involving univariate or bivariate distributions of data. *+Items can utilize any of the representations listed for grade 12.* *^Include items that require using multiple sets of data. For example, an item might require construction and comparison of three box plots based on given data sets.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 1. Data representation		
Grade 4	**Grade 8**	**Grade 12**
	d) Given a graphical or tabular representation of a distribution of data, determine whether the information is represented effectively and appropriately (histograms, plots over time, dot plots, scatterplots, box plots, bar graphs, circle graphs).	# d) Analyze, compare and contrast different graphical representations of univariate and bivariate data (e.g., identify misleading uses of data in real-world settings and critique different ways of presenting and using information). *^For example, an item might ask for a comparison of the effects of scale changes on the representation of data in a graph.*
		# e) * Organize and display data in a spreadsheet in order to recognize patterns and solve problems. *^Items can ask for the manipulation of spreadsheets, the recognition of patterns displayed in a spreadsheet, or the use of data to solve problems.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).
# Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 2. Characteristics of data sets		
Grade 4	**Grade 8**	**Grade 12**
	a) Calculate, use, or interpret mean, median, mode, range or shape of a distribution of data.	# a) Calculate, interpret, or use summary statistics for distributions of data including measures of center (mean, median), position (quartiles, percentiles), spread (range, interquartile range, variance, and standard deviation) or shape (skew, uniform, uni/bi-modal). *+Items involving shape should focus on interpreting and using.*
b) Given a distribution of whole number data in a context, identify and explain the meaning of the greatest value, the least value, or of any clustering or grouping of data in the distribution. *+The terms "clustering" and "grouping" can be used interchangeably but should not both be used in the same item.* *+Include items that allow students to describe clustering/ grouping of data within a distribution.*	b) Describe a distribution of data using its mean, median, mode, range, interquartile range, and shape.	b) Recognize how linear transformations of one-variable data affect mean, median, mode, range, interquartile range, and standard deviation. *^For example, an item might ask about the effect on the mean when a constant is added to each data point in a set.*
	c) Identify outliers and determine their effect on the mean, median, mode, or range.	# c) Determine the effect of outliers on the mean, median, mode, range, interquartile range, or standard deviation.
d) Compare two sets of related data using the greatest value, least value, or any clustering or grouping of data. *+The terms "clustering" and "grouping" can be used interchangeably but should not both be used in the same item.* *+Include items that allow students to describe clustering/ grouping of data within a distribution.*	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 should limit statistical measures to mean, median, mode, range, and interquartile range.*	# d) Compare data sets using summary statistics (mean, median, mode, range, interquartile range, shape, or standard deviation) describing the same characteristic for two different populations or subsets of the same population.

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 2. Characteristics of data sets (continued)		
Grade 4	**Grade 8**	**Grade 12**
	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. *+Items should present a scatterplot but not require drawing a line of best fit on the scatterplot.*	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 predictions. *^Items can require the use of technology to construct a least-squares regression line from a small data set.*
		# f) Recognize or explain how an argument based on data might confuse correlation with causation. *+For example, an item might require the critique of an argument about one of two strongly correlated variables causing change in the other.*
		g) * Identify and interpret the key characteristics of a normal distribution such as shape, center (mean), and spread (standard deviation).
		# h) * Recognize and explain the potential errors that can arise when extrapolating from data. *^For example, an item might require an explanation of the danger of using a line of best fit to make predictions for values well beyond the range of the given data.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 3. Experiments and samples		
Grade 4	**Grade 8**	**Grade 12**
	a) Given a sample, identify possible sources of bias in sampling. *+For example, an item might require identification of whether the members of a sample are representative of the population of interest.*	# a) Identify possible sources of bias in sample survey populations or questions 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. *+Items should focus on ways to select a random sample where every element of the population has the same likelihood of being selected.* *+Items should not assess the impact of random sampling on bias as described in Data Analysis, Statistics, and Probability objective 3.a.* *+For example, an item might involve using a random number generator to model a population.*
		# 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. *+For example, an item might require recognition of whether a given technique is appropriate to address a particular research question, or a comparison of the appropriateness of different sampling designs.*	d) Identify or evaluate the characteristics of a good survey or of a well-designed experiment. *+For example, an item might require reasoning about whether a sample is of sufficient size to draw conclusions about the population of interest.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 3. Experiments and samples (continued)		
Grade 4	**Grade 8**	**Grade 12**
		e) * Recognize the differences in design and in conclusions between randomized experiments and observational studies. *^For example, an item might ask 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.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 4. Probability		
Grade 4	**Grade 8**	**Grade 12**
		# a) Determine whether two events are independent or dependent.
	b) Using assumption of randomness, determine the theoretical probability of simple or compound events in familiar contexts. *^Items should use familiar contexts such as rolling a number cube, flipping a coin, or spinning the arrow of a spinner.*	# b) Using assumptions such as randomness, determine the theoretical probability of simple or compound events in familiar or unfamiliar contexts. *^Items should use* • *simple events that are independent or dependent, or* • *compound events that are independent.*
	c) Given the results of an experiment or simulation, estimate the probability of simple and compound events in familiar contexts. *^Items should use familiar contexts such as rolling a number cube, flipping a coin, or spinning the arrow of a spinner.*	# c) Given the results of an experiment or simulation, estimate the probability of simple or compound events in familiar or unfamiliar contexts. *^For example, an item might require an explanation involving how the relative frequency of occurrences of a specified outcome is not the same as its probability but can be used to estimate the probability of the outcome (e.g., Anita flipped a coin 10 times and got 7 heads, but the theoretical probability of a head is not 0.7).*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 4. Probability (continued)		
Grade 4	**Grade 8**	**Grade 12**
	d) Use theoretical probability to evaluate or predict experimental outcomes in familiar contexts. *+Items should use familiar contexts such as rolling a number cube, flipping a coin, or spinning the arrow of a spinner.*	# d) Use theoretical probability to evaluate or predict experimental outcomes in familiar or unfamiliar contexts. *^Items should be more complex than those at grade 8 (e.g., involve more events).* *+Item should present contexts of interest to a large cross-section of students. To increase the likelihood of capturing interests of the assessed students, the item pool should include a variety of student-relevant contexts.*
	e) Determine the sample space for a given situation. *+Include items that allow students to determine the number of different ways objects can be grouped (e.g., given three shirts and two pairs of pants, show how to determine the number of ways the shirts and pants can be paired).*	e) Determine the number of ways an event can occur using tree diagrams, formulas for combinations and permutations, or other counting techniques. *^Items should assess understanding of how to generate sample spaces.*
	f) Use a sample space to determine the probability of possible outcomes for an event.	
	g) Represent the probability of a given outcome using fractions, decimals, and percents. *+Items should involve writing a description of an outcome as a probability and should not involve calculating probabilities.*	

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 4. Probability (continued)		
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. *^Items should use simple events that are independent or dependent, or compound events that are dependent.*
		i) Determine conditional probability using two-way tables.
	j) Interpret and apply probability concepts to practical situations, and simple games of chance. *+Items should* • *assume that students are not familiar with specifics regarding playing cards, such as the number of cards in a deck, the suits represented in a deck of cards, or the number of cards of each suit;* • *use "number cube" instead of "dice" and assume that students are not familiar with the specifics of a number cube, such as the numbers presented on each face; and* • *avoid references to gambling.* *+For example, an item might state that 10% of the population is left-handed and require an estimate of the number of students that are left-handed in a school with 825 students.*	# j) Interpret and apply probability concepts to practical situations, including odds of success or failure in simple lotteries or games of chance. *+Items should* • *assume that students are not familiar with specifics regarding playing cards, such as the number of cards in a deck, the suits represented in a deck of cards, or the number of cards of each suit;* • *use "number cube" instead of "dice" and assume that students are not familiar with the specifics of a number cube, such as the numbers presented on each face; and* • *avoid references to gambling.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.5. Data Analysis, Statistics, and Probability (continued)		
Data – 4. Probability (continued)		
Grade 4	**Grade 8**	**Grade 12**
		k) * Use the binomial theorem to solve problems. *^Items should provide the binomial theorem.* *^For example, an item might present a binomial problem situation with the probability of an event being 0.1 and require determination of the probability of that event occurring 3 out of 11 times.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Algebra

Algebra began in the use of systematic methods for solving problems and numerical puzzles by mathematicians in the Middle East, South Asia, and China, and made its way to Europe in the late Middle Ages. The modern symbolic notation, with letters to stand for unknowns and constants, was developed in the 16^{th} century. The notation so greatly enhanced the power of the algebraic method that the basic ideas of both analytic geometry and calculus were developed within a century.

The increased use of algebra led to study of its formal structure. Gradually, the "rules of algebra" were distilled into a compact summary of the principles behind algebraic manipulation. In the 19^{th} century, these principles (e.g., commutativity, distributivity) were codified into a deductive system parallel to that of Euclidean geometry. A corresponding 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 taken broadly as including these ideas, the study and uses of algebra reach from the foundations of mathematics to the frontiers of current research.

The notion of variable – a symbol that can stand for any member of an identified set – has multiple facets (e.g., as an unknown, parameter, varying quantity); variables are used in many ways in school mathematics. Variables are used to express structural generalizations such as the commutativity of addition. In formulas such as $d = rt$ or $c = \sqrt{a^2 + b^2}$, variables stand for quantities that may take on a variety of values. In problem solving, a variable may represent an unknown quantity. The study of functions includes attention to independent variables, dependent variables and parameters.

When students make abstractions and generalizations about numbers and operations in early arithmetic by attending to underlying structure, they are engaging in algebraic thinking even though the formalism of algebraic notation may not be evident. As students progress through the grades, they continue to engage in algebraic thinking and they add more algebraic formalism to their repertoire.

By grade 4, students are expected to recognize and extend simple numeric patterns as a foundation for a later understanding of function. They begin to understand the meaning of equality and some of its properties, as well as the idea of an as-yet-unknown quantity as a precursor to the concept of variable. They also begin to informally explore properties of operations, including how inverse operations can be used to simplify a computation or how numbers can be decomposed and recomposed for more efficient computational strategies.

As students move into grade 8, the ideas of variable, covariation (two or more quantities varying simultaneously), and function become more important. By using variables to describe patterns and solve simple equations, students become familiar with manipulating them. Representations of covariation in 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, ratio, 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 ways of finding solutions, including graphing by hand or with technology.

By grade 12, students are expected to be skillful at manipulating and interpreting more complex expressions. Nonlinear functions, especially quadratic, power, and exponential functions whose graphs are accessible using graphing technology are used by students to solve real-world problems. Grade 12 students are also expected to be accomplished at translating verbal descriptions of problem situations into symbolic form. Also, by grade 12, students should understand expressions involving several variables, systems of linear equations, and solutions to inequalities.

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, determining solutions of higher degree polynomial equations or systems of linear or non-linear equations are not expected at grade 8 but are expected at grade 12.

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 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.

The 2025 Algebra objectives are shown in Exhibit 2.6. Included with many of the objectives is italicized text providing clarifications or limitations for use during item development.

Exhibit 2.6. Algebra (Alg)

Alg – 1. Patterns, relations, and functions		
Grade 4	**Grade 8**	**Grade 12**
a) Recognize, describe (in words or symbols), or extend simple numerical and visual patterns. *+Items should assess extensions of patterns in mathematically appropriate ways. For example, patterns should either be presented in ways that are transferable to a larger set or allow for multiple correct responses when not transferable to a larger set (e.g., when the first six elements of a pattern do not necessarily indicate the next six elements).* *+Pattern types can include whole numbers or shapes.*	a) Recognize, describe, or extend numerical and visual patterns using tables, graphs, words, or symbols. *^Items should involve more complex patterns than those presented at grade 4.* *+Items should assess extensions of patterns in mathematically appropriate ways. For example, patterns should either be presented in ways that are transferable to a larger set or allow for multiple correct responses when not transferable to a larger set (e.g., when the first six elements of a pattern do not necessarily indicate the next six elements).* *+Items should avoid linear patterns, which are addressed by other Algebra objectives.* *^Pattern types can include rational numbers, powers, simple recursive patterns, regular polygons, and three-dimensional shapes.*	a) Recognize, describe, or extend numerical patterns, including arithmetic and geometric sequences (progressions). *^Items should clearly define the nature of the pattern in the problem.* *+Items should assess extensions of patterns in mathematically appropriate ways. For example, patterns should either be presented in ways that are transferable to a larger set or allow for multiple correct responses when not transferable to a larger set (e.g., when the first six elements of a pattern do not necessarily indicate the next six elements).* *+Items should avoid linear patterns, which are addressed by other Algebra objectives.* *^Items can use patterns with multiple solutions when students are asked to explain their answers.* *^Pattern types can include those from grade 8, along with quadratic patterns and exponential patterns.* *Responses can include verbal descriptions or equations.*
		b) Express linear and exponential functions in recursive and explicit form given a verbal description, table, or some terms of a sequence. *^Include items that require* • *the explicit form of a function, given a recursive form.* • *the equation of a line, given a table of points.*

Exhibit 2.6. Algebra (continued)		
Alg – 1. Patterns, relations, and functions (continued)		
Grade 4	**Grade 8**	**Grade 12**
c) Given a description, extend or find a missing term in a pattern or sequence. *+Items should involve rules that follow the clarifications and limitations of numbers and operations identified in the Number Properties and Operations objectives.*	c) Examine or create patterns, sequences, or linear functions expressed as a rule numerically, verbally, or symbolically.	
d) Create a different representation of a pattern or sequence given a verbal description.		
	e) Identify functions as linear or nonlinear or contrast distinguishing properties of functions from tables, graphs, or equations. *^Items can ask about properties of lines or curves, including slopes and intercepts, but determination of the value of the slope of a curve is not required.*	e) Identify or analyze distinguishing properties of linear, quadratic, rational, exponential, or *trigonometric functions from tables, graphs, or equations. *+Items should avoid inverses as described in Algebra objective 2.j.* *^Items can include properties such as rate of change, intercepts, periodicity or symmetry.*
	f) Interpret the meaning of slope or intercepts, or determine the rate of change between two points, on a graph of a linear function. *+Items can use a real-world context.*	
		g) Determine whether a relation, given in verbal, symbolic, tabular, or graphical form, is a function.

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.6. Algebra (continued)		
Alg – 1. Patterns, relations, and functions (continued)		
Grade 4	**Grade 8**	**Grade 12**
		h) Recognize and analyze the general forms of linear, quadratic, rational, exponential, or *trigonometric functions. *+Items should avoid inverses as described in Algebra objective 2.j.* *^Items can include examining parameters and their effect on the graph of linear and quadratic functions (e.g., in y = ax + b, recognize the roles of a and b).*
		i) Determine the domain and range of functions given in various forms and contexts. *^Items should limit functions to linear, quadratic, inverse proportionality (y=k/x), absolute value, exponential, and trigonometric functions.* *^Items can include characteristics of domain and range in real-life 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, an item might ask: When f(t) represents a population in year t, what is the meaning of* $f^{-1}(3000) = 1965$*?*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.6. Algebra (continued)		
Alg – 2. Algebraic representations		
Grade 4	**Grade 8**	**Grade 12**
a) Translate between different representational forms (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). *+Items should involve whole number relationships that follow the clarifications and limitations of numbers and operations identified in the Number Properties and Operations objectives.*	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 should require either* • *translating between two different forms of representation, or* • *given one form of representation, creating a different form of representation.* *Items can include those that require the construction of graphs.* *The stimulus can include symbols, graphs, tables, diagrams, or written descriptions.*
	b) Interpret and compare representations of linear relationships expressed in symbols, graphs, tables, diagrams, or written descriptions. *+Representations are limited to linear relationships.* *^Items can include identification of strengths and weaknesses of different representations for different purposes.*	# b) Interpret and compare representations of relationships expressed in symbols, graphs, tables, diagrams (including Venn diagrams), or written descriptions. *+Representations can include any linear or nonlinear relationship appropriate to grade 12.* *+Items can include identification of strengths and weaknesses of different representations for different purposes.*
	c) Graph or interpret points represented by ordered pairs of numbers on a rectangular coordinate system. *^Items should limit coordinates to rational numbers.*	

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.6. Algebra (continued)		
Alg – 2. Algebraic representations (continued)		
Grade 4	**Grade 8**	**Grade 12**
	d) Solve problems involving coordinate pairs on the rectangular coordinate system. *^Items can include determining areas of simple geometric figures.*	d) Perform or interpret transformations on the graphs of linear, quadratic, exponential, and *trigonometric functions. *^Items should present the graph of the function in the stem.* *^For example, an item might ask for the vertex of the parabola resulting from $y = x^2$ being 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 relationships (e.g., compound interest, bacterial growth) in tables, graphs, words, or symbols. *^Items involving 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. *^Examples of real-world situations can be projectile motion, half-life, bacterial growth, Richter scale for earthquakes, and logarithmic scales in graphs.*
		# g) Solve problems involving exponential growth and decay. *+Items can involve science or finance contexts that will be familiar to students. For example, an item might involve modeling the effect of remediation of exponential growth of the bacteria in spinach production.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).
# Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.6. Algebra (continued)		
Alg – 2. Algebraic representations (continued)		
Grade 4	**Grade 8**	**Grade 12**
		h) *Identify distinguishing characteristics of exponential, logarithmic, and rational functions (e.g., discontinuity, asymptotes, concavity). *^Items should not require determining domains and ranges, which are addressed in Algebra objective 1.i.* *^Items can involve functions with points of discontinuity or asymptotes (vertical and horizontal).*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.6. Algebra (continued)		
Alg – 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. *+Items that involve numbers and operations should follow the clarifications and limitations identified in the Number Properties and Operations objectives.*		
b) Express simple mathematical relationships using expressions, equations or inequalities.	b) Write algebraic expressions, equations, or inequalities to represent a situation. *^Items should limit expressions, equations, or inequalities to those with first degree terms.*	b) Write algebraic expressions, equations, or inequalities to represent a situation. *^ Items can include determining the equation of a line given the slope and a point or given two points.* *^Expressions, equations, or inequalities can have terms of degree greater than one.*
	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.*

Exhibit 2.6. Algebra (continued)		
Alg – 3. Variables, expressions, and operations (continued)		
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. *^Items should provide formulas for the sum of a finite or infinite series.* *^For example, an item might ask for a range of possible total distances traveled by a ball when it is dropped from 20 feet above ground and makes three bounces, each up to 75% of its previous height.*
		h) Use basic properties of exponents and *logarithms to solve problems.

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).
# Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.6. Algebra (continued)		
Alg – 4. Equations and inequalities		
Grade 4	**Grade 8**	**Grade 12**
a) Find the unknown(s) in a whole number sentence (e.g., in an equation or simple inequality like [_] + 3 > 7). *^Items should present equations and inequalities that involve no more than one operation in the process of determining an unknown or a set of unknowns.*	a) Solve linear equations or inequalities (e.g., Solve for x in ax + b = c or ax + b = cx + d or ax + b > c). *^Items in a non-calculator block should limit coefficients to rational numbers.*	a) Solve linear, rational, or quadratic equations or inequalities, including those involving absolute value. *^Items should assume that students know the quadratic formula.* *^Items should limit coefficients to real numbers.* *Items should not use complex roots.*
b) Interpret "=" as an equivalence between two values and use this interpretation to solve problems.		b) * Determine the role of hypotheses, logical implications, and conclusions in algebraic arguments about equality and inequality. *^For example, an item might require understanding that neither of the following statements can be reversed:* $y = x - 1$ *implies* $y^2 = (x-1)^2$ *or* $f(x) = 0$ *implies* $g(x)*f(x) = 0$.
c) Verify a conclusion using simple algebraic properties derived from work with numbers (e.g., commutativity, properties of 0 and 1). *^For example, an item might require understanding that if Sam is 3 years older than Ned, 20 years from now Sam will still be 3 years older than Ned.*	c) Make, validate, and justify conclusions and generalizations about linear relationships. *^Items should require inductive and deductive reasoning when recognizing, expressing, or using the connections among and between linear relationships.*	c) Use algebraic properties to develop a valid mathematical argument. *^Items should address properties of equality and properties of operations. For example, an item might require an explanation for why division by zero is undefined.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).

Exhibit 2.6. Algebra (continued)		
Alg – 4. Equations and inequalities (continued)		
Grade 4	**Grade 8**	**Grade 12**
	d) 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).	# d) Analyze situations, develop mathematical models, or solve problems using linear, quadratic, exponential, or *logarithmic equations or inequalities symbolically or graphically. *^Items should not involve complex roots.* *Items can include real number coefficients.*
	e) Interpret relationships between symbolic linear expressions and graphs of lines by identifying and computing slope and intercepts (e.g., in y = ax + b, know that *a* is the rate of change and *b* is the vertical intercept).	e) Solve (symbolically or graphically) a system of equations or inequalities and recognize the relationship between the analytical solution and graphical solution. *^Items should limit systems of equations to two linear equations or one linear equation and one quadratic equation.* *Items can assess compound inequalities.*
	f) Use and evaluate common formulas (e.g., relationship between a circle's circumference and diameter [C = πd], distance and time under constant speed). *^Items should utilize formulas that come from a familiar context or situation.*	# f) Solve problems involving special formulas such as: $A = P(I + r)^t$ or $A = Pe^{rt}$. *^Items should present special formulas and define all variables in presented special formulas.* *+For example, a mathematical literacy item might involve comparing amounts that would be paid back from loans of equal value but with different interest rates.*

Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Exhibit 2.6. Algebra (continued)		
Alg – 4. Equations and inequalities (continued)		
Grade 4	**Grade 8**	**Grade 12**
		# g) Solve an equation or formula involving several variables for one variable in terms of the others. *+Items should assume that students know the quadratic formula.*
		h) * Solve quadratic equations with complex roots. *^Items should assume that students know the quadratic formula.*

* Objectives that describe mathematics content beyond that typically taught in a standard 3-year course of study (the equivalent of 1 year of geometry and 2 years of algebra with statistics).
# Grade 12 objectives that provide opportunities for questions in mathematical literacy.

Revisions of the 2017 Content Objectives

Revisions to the 2017 NAEP mathematics content objectives resulted from consideration of a wide range of relevant sources. These included research on mathematical development and learning, each state's standards and frameworks for mathematics instruction and assessment in the United States, reviews of state standards in comparison to NAEP objectives (e.g., Johnston et al., 2018), research on the alignment between NAEP items and common standards (e.g., Daro, Hughes, & Stancavage, 2015), policy statements informing state standards (e.g., NCTM, 2000, 2014, 2018; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010), the *Guidelines for Assessment and Instruction in Statistics Education* (GAISE, Franklin et al., 2007, in press), *Guidelines for Assessment and Instruction in Mathematical Modeling Education* (GAIMME, Garfunkel & Montgomery, 2016), the content of leading international assessments (e.g., PISA, OECD, 2019 and TIMSS, NCES, 2019), the professional judgment and experience of Panel members, and feedback obtained from readers of draft versions of the Framework.

Though overlapping, these sources were not in complete agreement regarding the mathematics students need to know and be able to do. Using this range of sources resulted in a set of objectives that cannot and will not be representative of what every child in the U.S. is taught by a given grade, nor will they conform precisely to the stated achievement objectives of any single state or professional organization. At the same time, the resulting objectives are tightly linked to acknowledged aspirations for the mathematics U.S. students should have an opportunity to learn. The content delineated here focuses on mathematical ideas that students are likely to have encountered in school.

Revisions attended to both current state standards – where the nation is now – and where the nation is likely headed. Updates to the content objectives were also motivated by several other considerations, including precision and accuracy of the language used to describe an objective, developmental appropriateness of objectives at a particular grade level based on current research and state policy, and shifts in content emphases since the last framework update. In the case of a limited number of objectives that are not common in the majority of U.S. state standards, guidance came from the ways leading states and nations situate those topics in their respective content objectives.

Restructuring of "Mathematical Reasoning" as a Subtopic

Mathematical Reasoning subtopics appeared in the previous NAEP Mathematics Assessment Framework (2017) in Number Properties and Operations, Geometry, Data Analysis, Statistics, and Probability, and Algebra. With the introduction of the NAEP Mathematical Practices (see Chapter 3), most of the Mathematical Reasoning objectives will be measured by items aligned to a content objective that are classified with one of the NAEP Mathematical Practices. To preserve attention to content that was uniquely present in some of the Mathematical Reasoning objectives, some content from those objectives was incorporated into other subtopics' objectives (e.g., Number and Operations subtopic 3.e in grades 4 and 8 was "Interpret…" and is now "Interpret, explain, or justify…").

Changes at Grade 4

In the early grades, up through grade 4, there is a distinction between NAEP content area arrangement and the arrangement common in many states' assessment standards. Most state assessments use three to five areas in the early grades but these do not parallel the five areas used in NAEP. At the same time, it must be noted that analysis of state standards has indicated that some content in the previous objectives is now not regularly part of U.S. schooling until grade 5 or later (Daro, Hughes, & Stancavage, 2015; Hughes, Daro, Holtzman, & Middleton, 2013; Johnston et al., 2018). To address this, some objectives were removed at grade 4. In many cases, grade 8 objectives were similar and more appropriately timed to assess students on mathematics they would have had a chance to learn. Additionally, research comparing states' standards for curriculum and instruction with NAEP assessment objectives suggested that some content commonly taught by grade 4 was absent from NAEP (Johnston et al., 2018). Careful review of this analysis led to the modification or addition of objectives at grade 4. Research and development on the use of the equal sign as an equivalence between two values and its importance in the foundation for algebraic thinking (Carpenter, Franke, & Levi, 2003) has meant states include more attention to it. This greater attention led to the addition of one related objective in grade 4 Algebra. Increased work with certain concepts in early grades since the last NAEP mathematics framework update led to one addition and several modifications of grade 4 Number Properties and Operations objectives. Similarly, several grade 4 objectives in Data Analysis, Statistics, and Probability were modified to reflect current language use for noticing, using, and interpreting data.

Changes at Grade 8

Since the last NAEP framework update, there have been shifts in state standards in expectations about understanding and use of rates, recognition of pattern, and greater attention to data, statistics, and probability in grades 5, 6, 7, and 8 (i.e., after grade 4; Johnston et al., 2018). As a result, the grade 8 objectives in Data, Statistics, and Probability were revised to clarify expectations, and three grade 8 objectives were deleted because similar grade 4 objectives or grade 12 objectives were more appropriately timed to assess what students have an opportunity to learn.

Changes at Grade 12

At grade 12, as in the other grades, descriptions of objectives were edited to clarify measurement intent. Added in grade 12 were two objectives in Geometry and Measurement: one about periodicity of functions and one on applying geometric properties among similar figures in two and three dimensions. In some cases where an objective was identified as beyond what is commonly taught in grade 12 an asterisk (*) was added. Also, to support the possible reporting of Mathematical Literacy as a particular way in which students know and do mathematics at grade 12, a number sign (#) was added to indicate objectives relevant to the exploration of this reporting.

Changes in Item Distribution

As noted above, the last decade has seen a shift of data and related topics to grades 5, 6, 7, and 8. Hence, the proportion of items for Data Analysis, Statistics, and Probability went up for grade 8 (from 15% to 20%) and down for grade 4 (from 10% to 5%). Concurrently, greater attention to fractions in grade 4 across states led to an increase in the proportion of Number Properties and Operations items (from 40% to 45%). Measurement in contexts that are not geometric play a smaller role in grade 8 than geometry topics and the proportion of such items was reduced (from 15% to 10%). By grade 12, most new measurement ideas are in geometric contexts and, as in the previous framework, measurement and geometry continue to be treated together in the item distribution for grade 12. In fact, the distribution of items for each content area at grade 12 remains the same, reflecting the delineation of essential concepts in the literature on high school learning (NCTM, 2018).

CHAPTER 3

NAEP MATHEMATICAL PRACTICES

Interest in students' mathematical practices has been growing for over 40 years. Seminal work by authors such as Collins and Stevens (1983), Lave (1988), Saxe (1988), and Schoenfeld (1985) focused on the cognitive skills and strategies used by mathematics experts and adults "in the wild" (i.e., outside of school). This line of research led to a distillation of the specific behaviors engaged during mathematical reasoning and problem solving, illuminating what are now called "practices" of mathematics.

Mathematics education research has also experienced a "social turn" (Lerman, 2000), marked by a shift toward investigating mathematics learning as it is situated in social activity, including discourse practices (Adler, 1999; Bell & Pape, 2012; Black, 2004; Civil & Planas, 2004; Enyedy, 2003; Ernest, 1998; Moschkovich, 2007, 2008; NCTM, 1991; van Oers, 2001). Students use their mathematical knowledge and skill in the social settings of school and home, on the basketball court, or in games they play with friends. The 2025 NAEP Mathematics Framework captures this broader and more complete picture of what it means to know and do mathematics. For the first time, NAEP Mathematics includes mathematical practices as a fundamental component of the assessment (see Exhibit 3.1). This chapter offers a brief overview of the research literature on mathematical practices as a whole and describes these five key NAEP

Exhibit 3.1. Summary of NAEP Mathematical Practices

NAEP Mathematical Practice 1: Representing
Recognizing, using, creating, interpreting, or translating among representations appropriate for the grade level and the mathematics being assessed.

NAEP Mathematical Practice 2: Abstracting and Generalizing
Decontextualizing, identifying commonality across cases, items, problems, or representations, and extending one's reasoning to a broader domain appropriate for the grade level and the mathematics being assessed.

NAEP Mathematical Practice 3: Justifying and Proving
Creating, evaluating, showing, proving, or refuting mathematical arguments/suppositions in developmentally and mathematically appropriate ways.

NAEP Mathematical Practice 4: Mathematical Modeling
Making sense of a scenario, identifying a problem to be solved, mathematizing it, and applying the mathematization to reach a solution and checking the viability of the solution in developmentally and mathematically appropriate ways.

NAEP Mathematical Practice 5: Collaborative Mathematics
The social enterprise of doing mathematics with others through discussion and collaborative problem solving whereby ideas are offered, debated, connected, and built-upon toward solution and shared understanding. Collaborative mathematics involves joint thinking among individuals toward the construction of a problem solution in developmentally and mathematically appropriate ways.

Mathematical Practices in depth. As was the case with the content areas in Chapter 2, these five areas are not meant to be inclusive of all possible mathematical activity.

Selecting Mathematical Practices for NAEP

The five NAEP Mathematical Practices are a particular distillation – for the purposes of assessment – of more than 40 years of research and development. They reflect a review of current scholarship, national and international assessment frameworks, national standards, and state standards more broadly.

To understand what mathematical practices are, it may be helpful to consider what they are not. Although practices underlie and contribute to mathematical reasoning, they are not completely synonymous with it because many other skills contribute to mathematical reasoning, such as working memory (Geary, Hoard, Byrd-Craven & DeSoto, 2004) and computational fluency (Geary, Liu, Chen, Saults, & Hoard, 1999). Similarly, although mathematical practices may contribute to conceptual understanding, the two are not interchangeable. On some accounts, conceptual understanding is knowledge of the underlying structure and relations represented in mathematics that transcends application of familiar algorithms (Eisenhart et al., 1993; Hiebert & LeFevre, 1986). In contrast, practices are fluid and responsive to both familiar and unfamiliar problems. Indeed, it is just as likely that conceptual understanding improves students' mathematical practices as it is that practices themselves improve conceptual understanding.

An increasing emphasis on mathematical practices is evident in state and national standards (NCTM, 1991, 2000, 2014). It is now generally agreed that knowing and doing mathematics entail engaging in practices such as generalizing, conjecturing, justifying, mathematizing, solving problems, communicating, and sense-making (Barbosa, 2006; Goos, 2004; Goos, Galbraith & Renshaw, 2002; Hufferd-Ackles, Fuson, & Sherin, 2004; Hussain, Monaghan, & Threlfall, 2013; Lau, Singh, & Hwa, 2009; Truxaw & DeFranco, 2008). As students grapple with and discuss mathematical ideas and problems – individually and together – they engage in such mathematical practices, which serve to familiarize them with the norms of doing mathematics (Herbel-Eisenmann & Cirillo, 2009). The inclusion of NAEP Mathematical Practices is not separate from the mathematics content of Chapter 2. These practices are described separately to indicate the significant change to the NAEP Mathematics Framework in sufficient detail.

The term "mathematical practices" has been used by the field in a variety of ways, with state standards and NCTM standards offering two widely disseminated descriptions. Five specific practices have been selected for emphasis on the 2025 NAEP Mathematics Assessment; these are referred to throughout the Framework and the Assessment and Item Specifications as the NAEP Mathematical Practices. As further detailed in Chapter 4, the assessment is designed to measure content and practices together. However, not all items will include an assessed NAEP Mathematical Practice. In fact, not all NAEP content objectives need to be assessed alongside a NAEP Mathematical Practice. Some items will continue to assess content outside of the particular NAEP Mathematical Practices, such as items that focus on algorithms, procedural fluency, precision, tool use, or mathematical practices other than the five that are the focus for the NAEP Mathematics Assessment.

There are commonalities across the NAEP Mathematical Practices and the practices described in policy documents and common in state standards. For example, the NAEP Mathematical Practices and the NCTM Process Standards include communication and collaboration, while communication is a subtext in several of the mathematical practices common in state standards (e.g., in critiquing the reasoning of others). Representing in the doing, teaching, and learning of mathematics is a process standard in NCTM's *Curriculum and Evaluation Standards for Mathematics* (1989), *Principles and Standards for School Mathematics* (*PSSM,* 2000), and *Catalyzing Change* (2018) and is also a NAEP Mathematical Practice. The NCTM Process Standards include reasoning and proof, and states' standards for mathematical practice include constructing viable arguments; both are similar to the NAEP Mathematical Practice of Justifying and Proving. The NAEP Mathematical Practice of Abstracting and Generalizing is similar to a common state standard for mathematical practice about reasoning abstractly and quantitatively. Mathematical Modeling is in most states' standards for mathematical practice as well as a NAEP Mathematical Practice.

The Achievement Level Descriptions (ALDs) provide examples of what students performing at the *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* achievement levels should know and be able to do in terms of the mathematics content areas and NAEP Mathematical Practices identified in the Framework (see Appendix A1). The breadth and relative emphasis of mathematics content and NAEP Mathematical Practices, as presented in the Framework and Assessment and Item Specifications, must be represented on the assessment as a whole. Therefore, assessment developers need to create a pool of items for use on the NAEP Mathematics Assessment that reflects both the performance expectations in the Framework and Assessment and Item Specifications and the range of performances illustrated in the NAEP mathematics ALDs. Because a variety of knowledge and skills can be embodied at different levels of student performance, considering achievement levels while developing items is important.

Some illustrations in Chapter 3 include *ALD Notes for Item Developers.* These notes provide descriptions of how NAEP achievement level language relevant to the NAEP Mathematical Practices and content objectives is currently reflected in the given item and how the achievement level connection might be affected by revisions to the item.

Operationalizing the NAEP Mathematical Practices

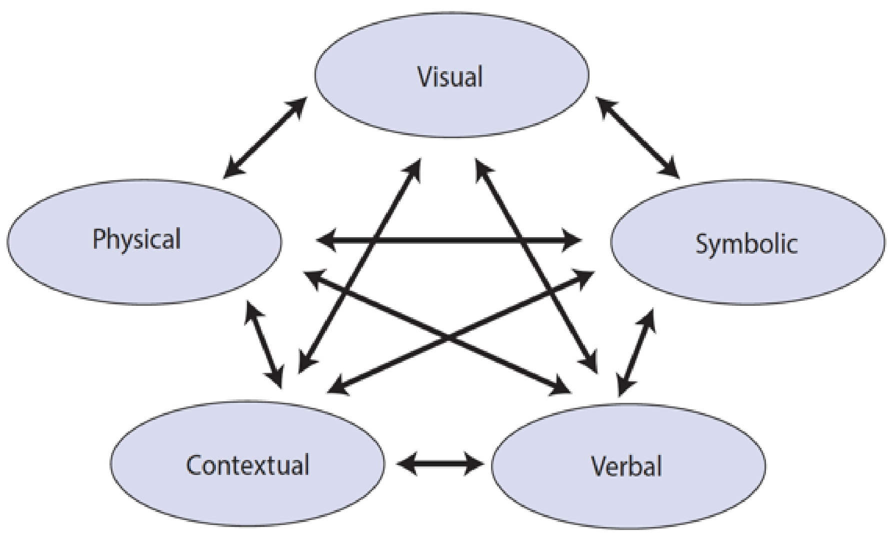
A description of each NAEP Mathematical Practice follows. Although each practice is treated as distinct, they are interrelated with one another and with content, as is demonstrated in the examples provided throughout. In designing NAEP items, it may be impossible to completely isolate a particular mathematical practice in an item. When items assess multiple aspects of mathematics, it should be possible to identify a primary content focus and a primary practice focus. The former has been done on NAEP Mathematics Assessments for many years, and the latter should be possible moving forward. Further, the practices fundamentally intersect with, and develop in relation to content. In this sense, the practices cut across grade levels, as well as across *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* achievement levels. This approach to mathematical practices is reflected in policy and state standards, where mathematical content standards are offered and described by grade levels, while practices cut across grade levels. Just as some mathematics content objectives are more likely to interact with others in items, some mathematical practices are more likely to be found in connection with certain mathematics objectives. At the end of this chapter, Exhibit 3.25 provides examples of where and how the five NAEP Mathematical Practices might be assessed within the NAEP mathematics content areas at each grade level. The tables are illustrative, not exhaustive of ways practices could be assessed within content areas.

NAEP Mathematical Practice 1: Representing

> *Representing: Recognizing, using, creating, interpreting, or translating among representations appropriate for the grade level and the mathematics being assessed.*

Representing mathematical ideas and using mathematical representations to make sense of and solve problems is central to mathematics. Students create representations themselves, or in collaboration with other students, and they reason from or translate between standard representations (e.g., graphs, tables, geometric drawings) (Lesh, Post, & Behr, 1987; NCTM, 2014). Tripathi (2008) argues that variety in representations "is like examining a concept through a variety of lenses, with each lens providing a different perspective that makes the picture (concept) richer and deeper" (p. 439). Exhibit 3.2, from *Principles to Actions* (NCTM, 2014, p. 25) illustrates some of the types of representation and the relationships among them.

Exhibit 3.2. Types and Connections Among Mathematical Representations

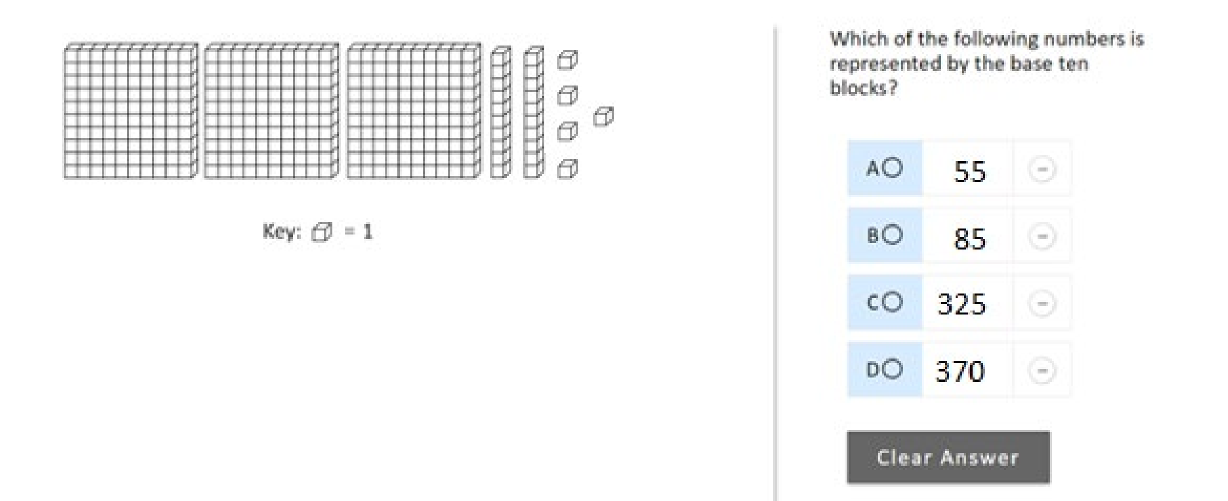


According to the National Research Council (NRC, 2009), students, especially young ones, benefit from using physical objects or acting out processes during problem solving. Base 10 blocks (or blocks/tiles representing other bases), fraction strips/bars, red–black integer tiles, and algebra tiles are all examples of physical representations of number and operation that are used to enhance students' understanding of concepts in elementary and middle grades. These visual and physical representations connect, eventually, to symbolic representations as well. Visual representations also play a particularly powerful role in helping students make sense of problems and understand mathematical concepts and procedures. For instance, arrays of squares in a grid can be used to represent area models for mathematical operations such as multiplication and division in early elementary grades, then later for multiplication of algebraic expressions. Additionally, students create, use, and reason about multiple representations for a given mathematical idea or relationship in contextually relevant ways.

The item in Illustration 3.1 is adapted from Exhibit 3.3 in the Framework. The item provides an image of base ten blocks and asks students to determine the number shown. In answering the question, students connect a visual representation of a number to its symbolic representation in base 10.

Illustration 3.1. Representing Example: Base-10 Blocks
based on **Exhibit 3.3**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Representing	Num – 1.b	SR – MC



Scoring Information	
Key	C. 325
ALD Notes for Item Developers	
Basic	The item assesses understanding of a visual representation of a familiar set of numbers – whole numbers.
Proficient	The item could be revised to present a representation of a whole number using base 10 blocks and ask for a different representation of the same number also using base 10 blocks.
Advanced	The item could be revised to require a description of why a base 10 representation of a number is incorrect along with a corrected representation.

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-4M9 #15 M347601.

The grade 8 item in Exhibit 3.4, from the 2003 NAEP Mathematics Assessment, demonstrates how students might provide a verbal representation from a graphical representation, or generate several alternative representations based on a problem situation. The item asks a student to take a graphical representation and work backward to a context that could fit that representation.

Alternatively, students could be asked to create their own graphical representation of a bicycle trip over time from a given verbal description of a trip. More realistic graphs of trips could be presented; for example, the item might offer a graph of a bicycle trip with more of a range and variety of speeds, including where the speed is zero at times mid-trip. Students could be given several different explanations that were provided by hypothetical students and asked to decide if those explanations correctly match the representation in the graph, or what an alternative explanation might be.

Exhibit 3.4. Grade 8 (and/or Grade 12) NAEP *Bicycle Trip* Item

The graph above represents Marisa's riding speed throughout her 80-minute bicycle trip. Use the information in the graph to describe what could have happened on the trip, including her speed throughout the trip.

During the first 20 minutes, Marisa

From 20 minutes to 60 minutes Marisa

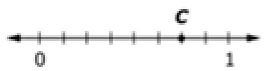
From 60 minutes to 80 minutes Marisa

The Smarter Balanced (SBAC) item in Illustration 3.2 is adapted from Exhibit 3.5 in the Framework. The item provides a point on a number line that represents a distance, along with additional written information. As they work to solve the problem, students are expected to engage with the measurement represented on the number line in conjunction with some additional information, recognize the representation of a fraction, and apply it within the given context.

Illustration 3.2. Representing Example: Number Line
based on **Exhibit 3.5**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Representing	Num – 3.f	SCR – FIB

Valeria and Diego walked home from school. The distance Valeria walked, in miles, is represented by point *C* on the number line.



Diego walked $\frac{2}{8}$ mile less than Valeria.

Enter the distance Diego walked as a fraction of a mile.

$\frac{\square}{\square}$

Scoring Information	
Key	4/8 (or equivalent)
ALD Notes for Item Developers	
Basic	The item could be revised to remove reference to Diego and ask students to identify the fraction of a mile that Valeria walked represented by point *C* on the number line.
Proficient	The item assesses use of a number line to solve a fraction computation problem involving comparing distances.
Advanced	The item could be revised to provide a problem situation whose solution can be determined by multiplying a whole number and a fraction less than 1. In the item stem could be a visual consisting of a set of wholes equal to the whole number factor. The item could require an explanation for how the wholes could be used to determine the solution.

This item was revised for the purpose of this document. The original version of this item appeared in SBAC as Item ID 3218.

The SBAC item in Illustration 3.3 is from Exhibit 3.6 in the Framework. The item asks students about two more ways of representing. In it, students select the written statement that could be represented by the given equation, connecting a context to a symbolic representation.

Illustration 3.3. Representing Example: Connecting Context to a Symbolic Representation ***based*** **on Exhibit 3.6**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Number Properties and Operations	Representing	Num – 3.e	SR – MC

Which situation can be represented by this equation?

$4 \div \frac{1}{8} = \square$

Ⓐ Jack has 4 pieces of fabric. Each piece is $\frac{1}{8}$ of a yard long. How many yards of fabric does Jack have?

Ⓑ Jack has 4 pieces of fabric. He gets $\frac{1}{8}$ more yards of fabric. How many yards of fabric does Jack have now?

Ⓒ Jack has 4 yards of fabric. He gives away $\frac{1}{8}$ of his pieces of fabric. How many pieces of fabric does Jack have left?

Ⓓ Jack has 4 yards of fabric. He cuts the fabric into pieces $\frac{1}{8}$ of a yard long. How many pieces of fabric does Jack have?

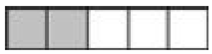
Scoring Information	
Key	D. Jack has 4 yards of fabric. He cuts the fabric into pieces 1/8 of a yard long. How many pieces of fabric does Jack have?
ALD Notes for Item Developers	
Basic	The item assesses the translation from one representation of a fraction operation (numeric) to another (verbal).
Proficient	The item could be revised to provide a number line from 0 to 4 partitioned into eighths and ask for an explanation for how the number line represents the quotient of 4 and 1/8.
Advanced	The item could be revised to provide the expression 4 ÷ (1/8) and a correct visual measurement representation of the quotient but an incorrect numerical representation of the quotient. The directive could be to explain the relationship between the visual and numeric representations of the quotient provided and determine whether each could be a correct representation.

This item was revised for the purpose of this document. The original version of this item appeared in SBAC as Item ID 3274.

Translating from one mathematical representation to another is a component of the practice of representing. For example, the item on the left in Illustration 3.4 asks students to write a fraction to describe the shaded part of a figure, moving from the visual to the numeric. In contrast, the item on the right in Illustration 3.4 (nonexample) asks students to choose the set of fractions ordered from least to greatest. Although a response to the nonexample item reveals something about what a student knows about fractions, it does not assess the practice of representing because the selection of the correctly ordered list does not meaningfully convey understanding of the representing of relative fraction size.

Illustration 3.4. Example and Nonexample of Representing

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4 8	Number Properties and Operations	Representing Other	Num – 1.e Num – 1.i	SCR – FIB SR – MC

Example	Nonexample
 What fraction of the figure is shaded? Answer: ______________	In which of the following are the three fractions arranged from least to greatest? A. $\frac{2}{7}, \frac{1}{2}, \frac{5}{9}$ B. $\frac{1}{2}, \frac{2}{7}, \frac{5}{9}$ C. $\frac{1}{2}, \frac{5}{9}, \frac{2}{7}$ D. $\frac{5}{9}, \frac{1}{2}, \frac{2}{7}$ E. $\frac{5}{9}, \frac{2}{7}, \frac{1}{2}$
Scoring Information	
Key 2/5	**Key** A. 2/7, 1/2, 5/9

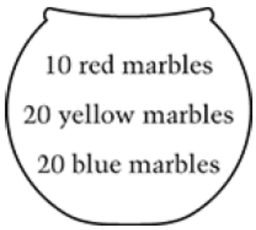
Original Grade 4 Item Information: NAEP Item ID 2007-4M7 #6 M139301, administered in 2007, aligned to grade 4 Number Properties and Operations objective 1.e, with a difficulty of easy.
Original Grade 8 Item Information: NAEP Item ID 2007-8M9 #12 M013631, administered in 2007, aligned to grade 8 Number Properties and Operations objective 1.j, with a difficulty of medium.

While the shaded figure in the example item in Illustration 3.4 is central to the item's assessment of representing, the inclusion of an image in an item does not automatically address the NAEP Mathematical Practice of Representing. Items may include images to convey information that could be provided another way, such as through written text. For an item to assess the NAEP Mathematical Practice of Representing, provided images would need to be representations with which students engage mathematically and that are critical to the solution process.

The item in Illustration 3.5 uses an image to convey information needed to solve a problem. The image might help students make sense of the item context, and the number of each color of marbles is essential for a complete item, but the picture of the bowl of marbles is not essential for the mathematical activity required to answer the question. That is, students do not engage mathematically with the image as they work to determine a response to the item. Therefore, the item does not assess the practice of representing.

Illustration 3.5. Representing Nonexample: Image Does Not Address Representing

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Other	Data – 4.b	SCR – composite



The bowl above contains the indicated number of marbles. The marbles are well mixed in this bowl. Juan will randomly pick a marble from the bowl. Juan believes that his chance of picking a blue marble is the same as his chance of picking a yellow marble. Is Juan correct?

Fill in the correct oval below.

⬭ Yes ⬭ No

Explain your answer.

Scoring Information

Key	Correct oval: Yes Explanation addresses that the chances of picking a blue marble or a yellow marble are equal. Example explanations: • This is true because the number of blue marbles in the bowl is equal to the number of yellow marbles. • The probability of selecting a blue marble is equivalent to the probability of selecting a yellow marble

This item was revised for the purpose of this document. The original version of this item appeared in the 2011 NAEP Mathematics Assessment with NAEP Item ID 2011-8M12 #6 M1532E1.

NAEP Mathematical Practice 2: Abstracting and Generalizing

> *Abstracting and Generalizing: Decontextualizing, identifying commonality across cases, items, problems, or representations, and extending one's reasoning to a broader domain appropriate for the grade level and the mathematics being assessed.*

Abstracting

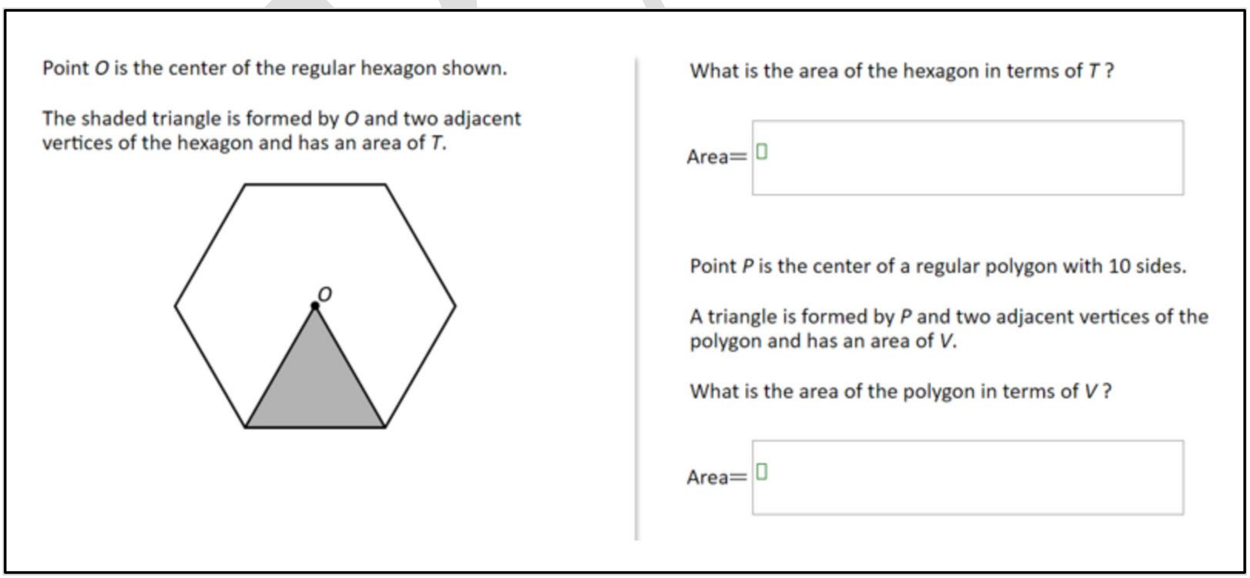
Students learning and doing mathematics also engage in the practice of abstracting and generalizing. An essential element of mathematical learning and problem solving is the ability to reason abstractly and to develop, test, and refine generalizations. In reasoning abstractly, students engage in the process of *decontextualizing*: Students abstract ideas in a given problem or context and express and manipulate them in a manner independent of their contextual references. Decontextualizing can foster an understanding of the relationships among problem contexts and written or symbolic forms, as well as an understanding of how mathematical expressions might be transformed to facilitate a solution strategy. Abstracting is also a critical activity for fostering generalizing; it enables a consideration of concepts and relationships decontextualized from specific examples or cases, which can support the formation of a more general rule or relationship.

Young students, for instance, can notice patterns of additive commutativity, such as 3 + 7 yielding the same sum as 7 + 3. In this instance, decontextualization would include finding a way to represent this relation independent of particular numbers, as a more general identity. Younger students might express this general identity verbally or with pictures, or with the use of a generic example. Older students might express this identity algebraically as $a + b = b + a$. Reasoning abstractly can also support recognizing similar mathematical structures across different problems or domains. For example, one could see the multiplication of two binomials $(2x + 7)(3x + 2)$ as a more general version of multiplying 27 by 32.

Consider the 2017 NAEP grade 8 Geometry item in Exhibit 3.7. This item requires students to express the area of the hexagon in terms of the area of the given shaded triangle. Students are then asked to extend their reasoning to a 10-sided figure. Thus, students are first challenged to reason structurally by mentally comparing the area of the triangle formed by the hexagon's center and two adjacent vertices with the area of the entire figure. Students are then further tasked with extending their reasoning from the specific case of the hexagon to another regular polygon.

Although a student could solve the problem in Exhibit 3.7 by drawing a 10-sided polygon and the specified triangle, and then counting the number of triangles that comprise the polygon, a student could also carry out this operation mentally rather than drawing it out. Also, the item could be revised to elicit decontextualizing beyond the hexagon, thinking about the relationship between the specified triangle and any regular polygon. In the later grades, students could be expected to express their reasoning algebraically and develop and prove a conjecture about the general relationship between the triangle and any n-sided regular polygon.

Exhibit 3.7. Grade 8 NAEP 2017 Geometry Item



Abstracting can occur across different domains. It can be addressed in reasoning about figures and their relationships in geometry, about number theory in number properties and operations, or about equivalence or functional relationships in algebra. How one decontextualizes or reasons with structure will differ across the domains, but these are processes students can employ in all five content areas included in the NAEP Mathematics Assessment.

Generalizing

Mathematics education researchers and policymakers have defined generalizing in a number of ways. Historically, generalization has been defined as an individual, cognitive construct (e.g., Carraher, Martinez, & Schliemann, 2008), where generalization is the act of identifying a property that holds for a larger set of mathematical objects or conditions than the number of individually verified cases. For instance, Harel and Tall (1991) described generalization as the process of "applying a given argument in a broader context" (p. 38), and Radford (2007) argued that generalization involves identifying a commonality based on particulars and then extending it to all terms.

More recently, researchers have begun to address generalizing as a construct that is both social and cognitive, that is, it can occur either individually or collectively. Therefore, for NAEP, generalizing is an individual or collective practice of (a) identifying commonality across cases, (b) extending reasoning beyond the domain in which it originated, and/or (c) deriving broader results from particular cases (Ellis, 2007). Its social dimensions make it relevant to the NAEP Collaborative Mathematics practice.

Several aspects of mathematical reasoning can foster generalizing. As mentioned above, abstracting and decontextualizing are important mental actions that support generalizing. Other actions that support generalizing include visualizing, focusing, reflecting, connecting, and expressing. *Visualizing* involves seeing patterns or structural relationships, as well as imagining a set of relationships beyond what is perceptually available. *Focusing* is attending to particular details, characteristics, properties, or relationships above others. This can include examining a particular case in a pattern or attending to figural or numerical cues. *Reflecting* involves actions such as thinking back on the operations one has carried out, observing one's method in solving problems, or examining the rules that govern a given pattern. *Connecting* is the identification of relationships among tasks, representations, or properties. Making connections between representations or identifying and operating on structural similarities can foster the development of generalizations. Finally, *expressing* involves depicting a generalization verbally or in writing. Describing generalizations in words can support the subsequent development of algebraically-represented generalizations.

Like abstracting, generalizing can occur across the content areas and grade bands. Existing NAEP Mathematics Assessment items contain a number of generalization tasks in which students are asked to determine a rule guiding the pattern of number terms in a sequence. In some items, potential rules are provided for students who are prompted only to attend to the action required to move from one term in the sequence to the next. In other items, students must determine a rule themselves, such as for the grade 12 item in Illustration 3.6 (from Exhibit 3.8 in the Framework). It is worth noting that for items such as the one shown in Illustration 3.6, there could be any number of non-equivalent rules to describe the pattern, so it may be more appropriate to ask students to provide "a" rule rather than "the" rule.

Illustration 3.6. **Abstracting and Generalizing Example:** **Write a Rule to Describe a Pattern** *based on* **Exhibit 3.8**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Algebra	Abstracting and Generalizing	Alg – 1.b	ECR – composite

Sequence I: 3, 5, 9, 17, 33, . . .

Sequence I, shown above, is an increasing sequence. Each term in the sequence is greater than the previous term.

a. Make a list of numbers that consists of the positive differences between each pair of adjacent terms in Sequence I. Label the list Sequence II.

b. If this same pattern of differences continues for the terms in Sequence I, what are the next two terms after 33 in Sequence I ?

6th term ______________________

7th term ______________________

c. Write an algebraic expression (rule) that can be used to determine the n^{th} term of Sequence II, which is the difference between the $(n+1)^{\text{st}}$ term and the n^{th} term of Sequence I.

Scoring Information	
Key	a. 2, 4, 8, 16 b. 6^{th} term = 65 7^{th} term = 129 c. Difference between *n*th and (*n*+1)st term is 2^n
ALD Notes for Item Developers	
Basic	The item could be revised to remove Part C and assess the creation and extension of patterns.
Proficient	The item could be revised to remove Part C and assess the creation and extension of patterns along with an explanation for the choices made.
Advanced	The item assesses use of structures and patterns to determine a complex rule.

This item was revised for the purpose of this document. The original version of this item appeared in the 2005 NAEP Mathematics Assessment with NAEP Item ID 2005-12M3 #17 M095401.

Notice that for part c of this grade 12 item, students are expected to write a formal algebraic rule for the n^{th} term of Sequence II. In other items, students may be tasked with determining a recursive rather than an explicit rule to find the n^{th} term in a sequence.

Determining a rule for a pattern is a common focus of grade 4 generalization items, such as the adapted grade 4 TIMSS (2011) item shown in Illustration 3.7. The original TIMSS item was multiple choice, with response options providing choices for Steve’s rule. Here it has been modified to a fill-in-the-blank item requiring students to determine correct values for the rule and avoiding the opportunity for students to check provided response options to determine which could be the rule.

Illustration 3.7. Abstracting and Generalizing Example: Complete a Rule to Describe a Pattern

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Algebra	Abstracting and Generalizing	Alg – 1.a	SCR – FIB

Steve used the same rule to get the number in the □ from the number in the △.

Complete this rule so that it could be Steve's Rule.

multiply by ☐, then add ☐

Scoring Information

Key	multiply by 2, then add 2

This item was revised for the purpose of this document. The original version of this item appeared in the 2011 TIMSS as Item ID M031251.

Students can also be challenged to engage in the processes of generalizing in items that do not rely on pattern sequences, as in Exhibit 3.9. This item could support a number of possible generalizing processes, as well as the opportunity for abstracting. For instance, one could consider that for each coin (nickel, dime, quarter), there are two possible outcomes, H or T. Thus, a student could either systematically list outcomes to determine that there are 8 total outcomes or could begin to think structurally to reason that for three coins and two outcomes per coin, there must be $2^3 = 8$ total outcomes. Alternatively, through systematic listing a student could determine that there are 1 + 3 + 3 + 1 outcomes, corresponding to 1 outcome with exactly zero Ts, 3 outcomes with exactly one T, 3 outcomes with exactly two Ts, and 1 outcome with exactly three Ts. Extending to the 4-coin case, for instance, students might determine that the number of outcomes is 1 + 4 + 6 + 4 + 1, corresponding to 1 outcome with exactly zero Ts, 4 outcomes with exactly 1 T, 6 outcomes with exactly 2 Ts, 4 outcomes with exactly three Ts, and 1 outcome with exactly four Ts (and symmetrically but opposite for the number of Hs).

Exhibit 3.9. Grade 8 and/or Grade 12 task (adapted from 2013 Grade 8 NAEP item)

Three students each have a coin, one has a nickel, one has a dime, and the third student has a quarter. They flip their coins at the same time. Each coin can land either heads up (H) or tails up (T). List all the different possible outcomes for how the coins could land in the chart below. The list has been started for you.

Nickel	Dime	Quarter
H	H	H
H	H	T

What if a 4th student joins the group with a half-dollar coin? How many different ways could the 4 coins land? What if a 5th student joined with a penny–how many different ways could the 5 coins land?

One aspect of generalizing is identifying commonality across cases. Students might notice that the outcomes for the 3-coin and 4-coin cases can be structured according to the rows in Pascal's triangle. Or, students might reason that, like the 3-coin case, each of the positions in the 4-coin case has two possible outcomes, H or T, and thus the total number of possible outcomes must be $2^4 = 16$, and more generally for *n* coins, 2^n. An item like the one in Exhibit 3.9 affords a number of rich generalizing opportunities, regardless of whether students are expected to recognize that 2^n is the sum of the coefficients of the binomial expression $(a + b)^n$ (e.g., $2^4 = 1 + 4 + 6 + 4 + 1$).

An item assessing generalizing may call on structural reasoning, requiring students to break mathematical components of an items apart to identify the building blocks needed to answer the question (Cuoco, Goldenberg, & Mark, 1996; Küchemann & Hoyles, 2009). As such, a distinction needs to be made between items that ask students to reason structurally and items that prompt students to identify known quantities or properties. Consider the grade 8 item in Illustration 3.8. Since one example of each set is needed to determine the correct response, students do not need to consider the structure of each set. Therefore, this item does not ask students to reason structurally about described sets of geometric objects (see Chapter 4, p. 187 for additional information about the item type represented in Illustration 3.8).

Illustration 3.8. Abstracting and Generalizing Nonexample: Recognizing Properties of Geometric Objects

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Geometry	Other	Geom – 3.g	SR – grid

Three questions about lines and angles in the plane are given in the table.

Indicate whether the answer to each question is 0, 1, 2, 4, or 8.

Make one selection for each question to show your answer.

Questions	0	1	2	4	8
At how many points do two different parallel lines meet?	○	○	○	○	○
How many right angles are formed by a pair of perpendicular lines?	○	○	○	○	○
How many angles that measure less than 180 degrees are formed when two parallel lines are cut by a transversal?	○	○	○	○	○

Clear Answer

Scoring Information

Key

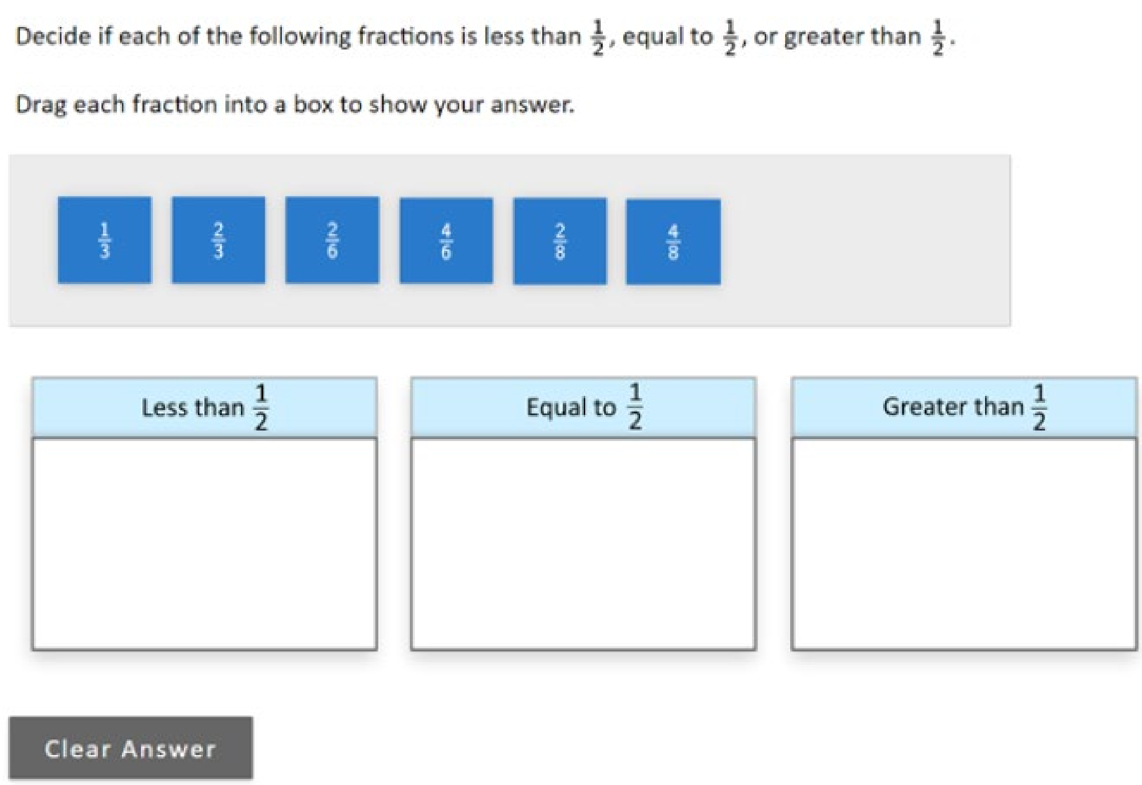
Questions	0	1	2	4	8
At how many points do two different parallel lines meet?	◉	○	○	○	○
How many right angles are formed by a pair of perpendicular lines?	○	○	○	◉	○
How many angles that measure less than 180 degrees are formed when two parallel lines are cut by a transversal?	○	○	○	○	◉

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #10 M3821MS.

Since the practice of abstracting and generalizing involves reasoning about mathematical structures and systems, items that focus on concrete examples likely do not assess this practice. Consider the grade 4 item in Illustration 3.9. In this item, students are asked to compare given fractions to the benchmark number 1/2; they are not asked to determine the structure of a fraction that is less than, equal to, or greater than 1/2. The thought process behind a student's matching results are unknown. Instead, the evidence provided from a response to this item indicates whether or not the fractions were compared correctly (see Chapter 4, p. 183 for additional information about the item type represented in Illustration 3.9).

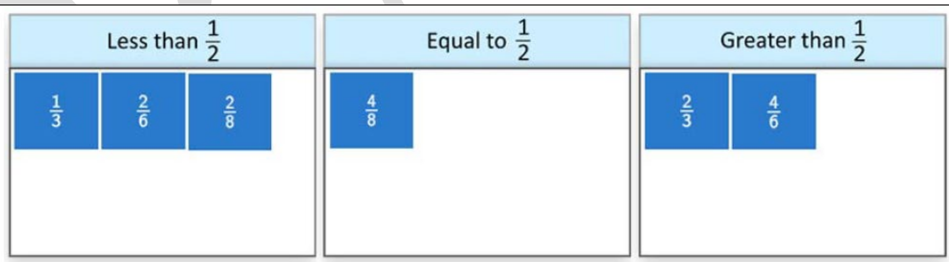
Illustration 3.9. Abstracting and Generalizing Nonexample: Using Benchmarks

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Other	Num – 1.i	SR – matching



Scoring Information

Key



This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-4M1 #6 M3714MS.

Abstracting and generalizing support students' problem-solving activity. The types of structural elements students identify and abstract will influence the generalizations they make, and students' processes of generalization can, in turn, affect other aspects of problem solving.

NAEP Mathematical Practice 3: Justifying and Proving

> *Justifying and Proving: Creating, evaluating, showing, proving, or refuting mathematical arguments/suppositions in developmentally and mathematically appropriate ways.*

Justifying and proving are essential in all content areas and grade levels. Traditionally, proof was viewed as a form of mathematical argumentation pertaining first to high-school geometry and not visited again until pre-calculus courses with proofs of trigonometric identities and proofs by mathematical induction. However, this changed in the last quarter of the 20th century. The *Principles and Standards for School Mathematics* emphasized the importance of justifying and proving at all levels of mathematics, noting that "reasoning and proof should be a consistent part of students' mathematical experience in prekindergarten through grade 12" (NCTM, 2000, p. 56). Similarly, state standards highlight the activities students engage in as they learn to create valid mathematical arguments: making and investigating conjectures, developing particular forms of argument (e.g., deductive, inductive), and using a variety of proof methods (e.g., direct, counterexample). These are all considered components of the practice of justifying and proving.

Mathematical justification includes creating arguments, explaining why conjectures must be true or demonstrating that they are false, exploring special cases or searching for counterexamples, understanding the role of definitions and counterexamples, and evaluating arguments (Ellis, Bieda, & Knuth, 2012). A valid justification should show why a statement or conjecture is true or not true generally (i.e., for all cases) and, especially by grades 8 and 12, should do so by providing a logical sequence of statements, each building on already established statements, ideas, or relationships.

Consider the item in Illustration 3.10, which asks students to choose a counterexample to a statement. Knowing that a single example is all that is needed to contradict a supposition or conjecture is a fundamental building block to establishing a mathematical argument, as examples cannot show something is true generally, but can show that something is false. This item asks for identification of the counterexample presented in a set of response options, and the required action reflects emerging development of the practice of justifying and proving.

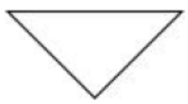
Illustration 3.10. Justifying and Proving Example: Choose the Counterexample

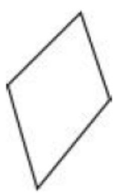
Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Geometry	Justifying and Proving	Geom – 1.e	SR – MC

Alan says that if a figure has four sides, it must be a rectangle. Gina does not agree. Which of the following figures shows that Gina is correct?

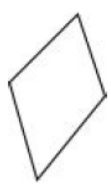
A.


B.


C.


D.


Scoring Information

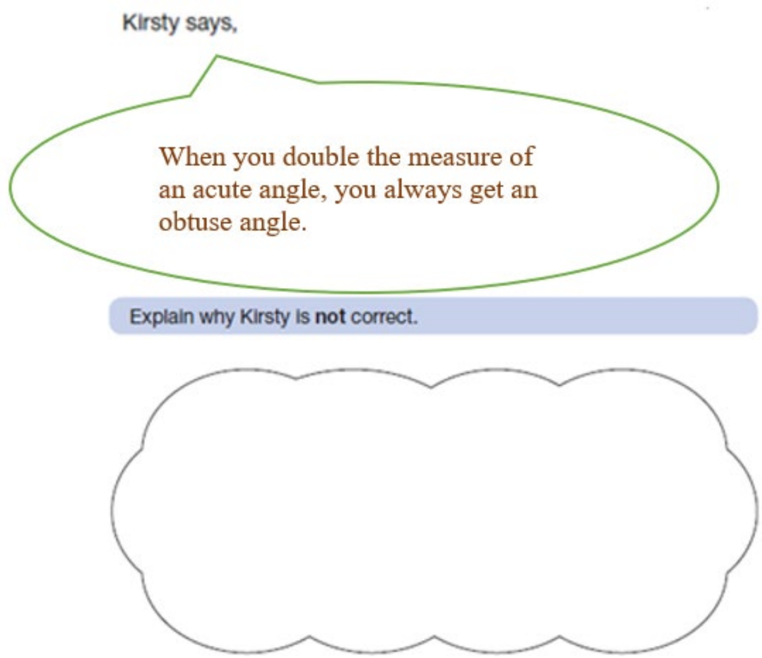
Key	D. 

This item was revised for the purpose of this document. The original version of this item appeared in the 2003 NAEP Mathematics Assessment with NAEP Item ID 2003-4M6 #7 M046401.

The adapted England *Key Stage* released item in Illustration 3.11 provides a mathematical statement and asks for an explanation for why the statement is incorrect. As with the previous illustration, this item, at a minimum, requires presentation of a counterexample. Alternatively, a student might respond in a way that generalizes the range of angle sizes for which the given statement is false. Although a response detailing when the statement is false lends itself to the practice of abstracting and generalizing, this type of response is not needed to receive credit. Instead, a student could provide a single example of when doubling an angle measure results in an angle that is not obtuse. Therefore, the item does assess the practice of justifying and proving but does not assess the practice of abstracting and generalizing.

Illustration 3.11. Justifying and Proving Example/ Abstracting & Generalizing Nonexample: Generate a Counterexample

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Geometry	Justifying and Proving	Geom – 1.b	SCR



Scoring Information

Key		
	An explanation that includes a correct counter example, e.g. • When you double 10° it is not obtuse • 2 × 27° = 54° • Double 45° is a right angle not obtuse **OR** An explanation that demonstrates where the statement in the question is not correct, e.g. • If the acute angle is less than 45° then doubling it will be less than 90°, so it won't be obtuse (more than 90°).	**Do not** accept vague or incomplete explanations, e.g. • Sometimes it will be acute • Some acute angles are half an obtuse angle, but not all • When you double an acute angle, you get a right angle **Do not** accept explanations which include incorrect mathematics or incorrect information that is relevant to the explanation, e.g. • 20°C × 2 = 40°C • 20% x 2 = 40%

This item was revised for the purpose of this document. The original version of this item appeared in the 2003 England *Key Stage 2* as Item 13.

In geometry, definitions are often used to justify mathematical statements. The item in Illustration 3.12, adapted from a 2009 NAEP item, provides a mathematical statement that students are required to support or refute. Students engage in a fundamental aspect of justifying by using the definition of a function to form the requested explanation.

Illustration 3.12. Justifying and Proving Example: Supporting with a Definition

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Algebra	Justifying and Proving	Alg – 1.g	SCR

This table shows all of the ordered pairs (x, y) that define a relation between the variables x and y.

x	-2	-1	0	1	2	3
y	3	0	-1	0	3	8

Explain whether or not y is a function of x.

Scoring Information

Key	Sample Explanation: y is a function of x because for each x-value (domain) there is only one y-value (range) that is associated with it.

This item was revised for the purpose of this document. The original version of this item appeared in the 2009 NAEP Mathematics Assessment with NAEP Item ID 2009-12M2 #7 M1906E1.

As mentioned above, the item in Illustration 3.12 is an adaptation of a NAEP released item. A key revision to the original item was to combine a yes or no response and a request for an explanation into a single directive.

A justification is *not* based on authority, perception, popular consensus, or examples alone. As students engage in justifying, they may be tempted to rely on external sources to verify their ideas, such as their teacher or a textbook (Harel & Sowder, 1998). Students may also want to use examples to support their claims, concluding that a conjecture must be true because it holds for several different cases. Examples can and do play an important role in justifying and proving, particularly in terms of helping students make sense of statements, gain a sense of conviction, or revealing an underlying structure that could lead to a proof. But they do not suffice as a mathematical justification or proof except for proofs by exhaustion or counterexample.

Consider the grade 4 item in Exhibit 3.10. Some students may use specific examples in their arguments, which would show limited ability to justify a claim. A complete response to this item requires students to indicate why the claim is true for all numbers.

Exhibit 3.10 Grade 4 Number Properties and Operations Proof Item

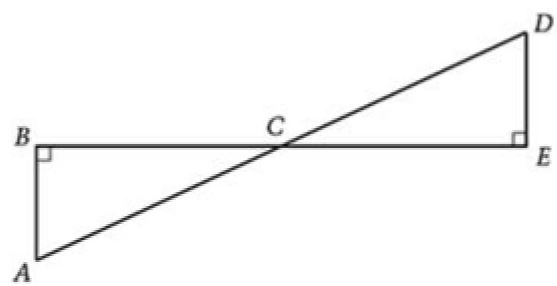
Elise claims that if you multiply any whole number by 6, you will always get an even number for the answer. Provide an argument for why Elise is correct.

A grade 4 proof for the claim in Exhibit 3.10 could involve demonstrating with either pictures or symbols that the answer can always be separated into two equal parts, because 2 is a factor of 6, or that the 2 answer can always be divided by 2 or cut in half because 2 already divides 6. An argument such as 6 x NUMBER = 3 x NUMBER + 3 x NUMBER might also be provided by fourth graders, demonstrating symbolically that the result can be split into two equal parts. Arguing from examples alone is not a justification, but in providing examples students may discover the key piece to demonstrate that 2 will always be a factor of the product.

A formal proof is a specific type of argument "consisting of logically rigorous deductions of conclusions from hypotheses" (NCTM, 2000, p. 55). Here "argument" is being used to include both an assertion and a defense of the assertion. The argument constitutes a proof when the defense of the assertion is clear, complete, and convincing, containing logically connected mathematical statements that are based on valid definitions and theorems. In grade 12, students are expected to develop formal mathematical proofs. A proof uses definitions and theorems that are available without further justification and a proof is valid only if the assumptions upon which it relies have already been shown to be true.

Often, the phrase "mathematical proof" conjures an image of the traditional two-column proof that is typical in high-school geometry classrooms. This form of proof can be helpful for supporting students' efforts to develop a clear chain of statements, each relying on the prior, and for making sure that each statement is justified, as illustrated in Exhibit 3.11.

Exhibit 3.11. Grade 12 NAEP Geometry Proof Item



Given: C is the midpoint of $\overline{BE}$.
$\angle B$ and $\angle E$ are right angles.

Prove that $\overline{AC} \cong \overline{DC}$ and give a reason for each statement in your proof.

This item lends itself well to a two-column proof, particularly because it stipulates that a reason must be provided for each statement in the proof. One proof is as follows:

Statement	Reason
C is the midpoint of $\overline{BE}$	Given
$\angle B$ and $\angle E$ are right angles	Given
$\overline{BC} \cong \overline{EC}$	Definition of midpoint
$\angle B \cong \angle E$	Right angles are congruent
$\angle ACB \cong \angle DCE$	Vertical angles are congruent
$\triangle ACB \cong \triangle DCE$	Angle-Side-Angle (or Leg-Angle)
$\overline{AC} \cong \overline{DC}$	Corresponding parts of congruent triangles are congruent

Although this proof follows a typical form of school mathematics proof, there is nothing about the prompt that stipulates that the proof must occur in a two-column format. A narrative form of the proof in answer to the item in Exhibit 3.11 could also be appropriate, as seen below:

> The measures of $\angle BCA$ and $\angle ECD$ are equal because vertical angles have the same measure. We also know that the measures of $\angle B$ and $\angle E$ are the same because they are both right angles. Since C is the midpoint of $\overline{BE}$, $\overline{BC} \cong \overline{EC}$. So, by the angle-side-angle rule, triangle *ACB* is congruent to triangle *DCE*. Therefore, $\overline{AC} \cong \overline{DC}$ because corresponding parts of congruent triangles are congruent.

Understanding that a single counterexample undermines a general claim is an important part of justification. Within an item, the presence of two contradictory statements provides impetus for the exploration of examples that refute one of the statements. This can be seen in the grade 12 algebra item in Exhibit 3.12. Here, one could identify a value for x that is, for instance, less than 5 but not also greater than –3 (e.g., $x = -10$). That single counterexample is sufficient to show that Dave's claim cannot be correct because $x = -10$ does not satisfy the statement $-3 < x < 5$.

Exhibit 3.12. Grade 12 NAEP Algebra Counterexample Item

Question A: If x is a real number, what are all values of x for which $x > -3$ and $x < 5$?

Question B: If x is a real number, what are all values of x for which $x > -3$ or $x < 5$?

Barbara said that the answers to the two questions above are different.

Dave said that the answers to the two questions above are the same.

Which student is correct?

◯ Barbara ◯ Dave

Explain why this student is correct. You may use words, symbols, or graphs in your explanation.

The questions at the start of the item in Exhibit 3.12 could be altered to give a grade 8 item:

Question A: If x is a number, what are all values of x for which $x \geq -3$.
Question B: If x is a number, what are all values of x for which $x > -3$.

The rest of the item would remain the same.

Similarly, only one counterexample is needed to refute Pat's claim in the grade 8 number properties and operations item in Illustration 3.13 (from Exhibit 3.13 in the Framework). Multiplying 6 by any real number less than 1 will yield a result less than 6, confirming Tracy's claim and refuting Pat's claim (see p. 137 for a discussion on this item related to collaborative mathematics).

Illustration 3.13. Justifying and Proving Example: Confirming/ Refuting Claims *based on* **Exhibit 3.13**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Number Properties and Operations	Justifying and Proving	Num – 3.d	SCR

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 Information

Key	Examples of correct reasons: • If you multiply by a number *smaller* than 1, the result is less than 6. • $6 \times 0 = 0$ • $6 \times \frac{1}{2} = 3$ • $6 \times -1 = -6$

ALD Notes for Item Developers

Basic	The item could be revised to require the selection of the description for why the product of 6 and 1/2 is less than 6, focusing on the meaning of multiplication, not the value of the product.
Proficient	The item assesses understanding and use of a counterexample to refute a claim.
Advanced	The item could be revised to ask for the conditions, if any, and a justification for the conditions, under which the product of 6 and an irrational number is a rational number. Note that this revision also assesses Abstracting and Generalizing.

This item was revised for the purpose of this document. The original version of this item appeared in the 1992 NAEP Mathematics Assessment with NAEP Item ID 1992-12M14 #2 M054801.

Learning to search for counterexamples and explaining why they are justifications is only one aspect of refutation. Attempting to prove that a conjecture is false can also lead to the development of new insights or ideas, as well as to the formation of different conjectures that can then be explored, refuted, or proved.

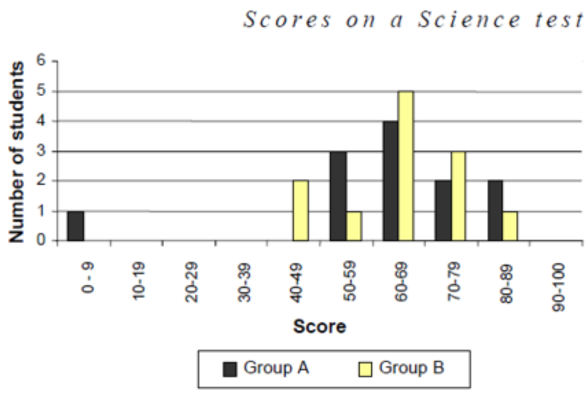
The item in Illustration 3.14 is a released PISA (2006) Data Analysis, Statistics and Probability item that assesses justifying and proving. This item provides a data set and asks for a mathematical argument that counters a claim, thereby assessing justifying and proving (see the explanation on p. 128 for why this item does not also assess mathematical modeling).

Illustration 3.14. Justifying and Proving Example: Providing a Mathematical Argument

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Data Analysis, Statistics, and Probability	Justifying and Proving	Data – 1.c	SCR

The diagram below shows the results on a Science test for two groups, labelled as Group A and Group B.

The mean score for Group A is 62.0 and the mean for Group B is 64.5. Students pass this test when their score is 50 or above.



Looking at the diagram, the teacher claims that Group B did better than Group A in this test.

The students in Group A don't agree with their teacher. They try to convince the teacher that Group B may not necessarily have done better.

Give one mathematical argument, using the graph, that the students in Group A could use.

Scoring Information

Key	One valid argument is given. Valid arguments could relate to the number of students passing, the disproportionate influence of the outlier, or the number of students with scores in the highest level. • More students in Group A than in Group B passed the test. • If you ignore the weakest Group A student, the students in Group A do better than those in Group B. • More Group A students than Group B students scored 80 or over.

This item was revised for the purpose of this document. The original version of this item appeared in the 2006 PISA as Item ID M513Q01 – 0 1 9.

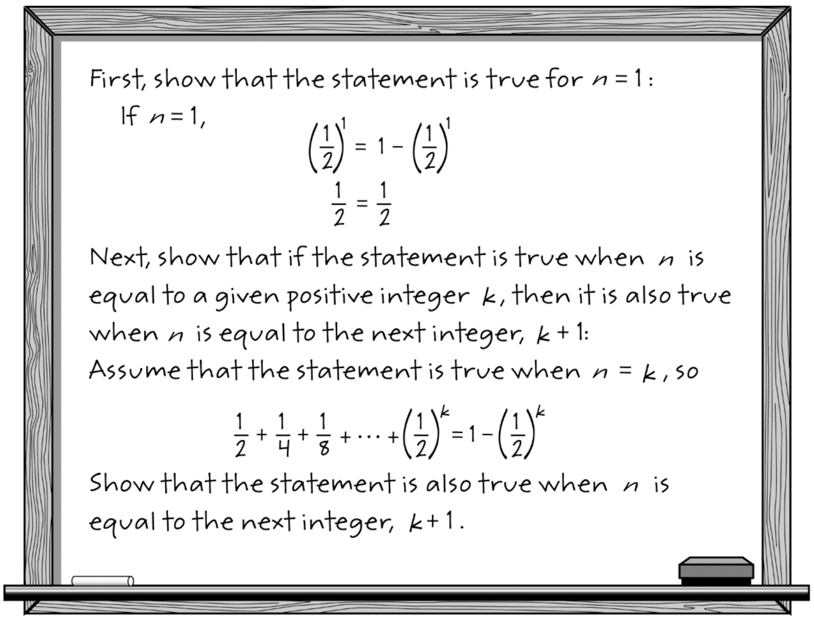
In addition to the various formats one can use to develop or present proofs, there are also many other ways of mathematically proving, disproving, or justifying a mathematical answer, including proving by exhaustion (i.e., verifying every possible case) and employing mathematical induction. Often, it may be easier to use a particular mode of argumentation based on the nature of the claim. Some NAEP items require a specific mode of proof, such as the grade 12 number properties and operations item in Exhibit 3.14.

Exhibit 3.14. Grade 12 NAEP Number Properties Mathematical Induction Item

A student was asked to use mathematical induction to prove the following statement.

$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\left(\frac{1}{2}\right)^n = 1-\left(\frac{1}{2}\right)^n$ for all positive integers n

The beginning of the student's proof is shown below.



Complete the student's proof by showing that if the statement is true when $n = k$, then it is also true when $n = k + 1$, where k is any positive integer.

Here, a student must use the tools of mathematical induction to complete the provided argument:

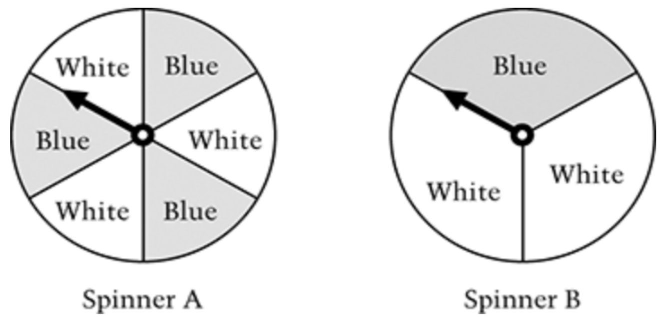
For $n = k + 1$, $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\left(\frac{1}{2}\right)^{k+1}$ can be expressed as $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\left(\frac{1}{2}\right)^{k}+\left(\frac{1}{2}\right)^{k+1}$. We know from the above statement that $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\left(\frac{1}{2}\right)^{k}$ is equal to $1-\left(\frac{1}{2}\right)^{k}$, so substituting that yields $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\left(\frac{1}{2}\right)^{k+1} = 1-\left(\frac{1}{2}\right)^{k}+\left(\frac{1}{2}\right)^{k+1}$. Simplifying the expression on the right gives us $\frac{2^{k+1}-1}{2^{k+1}}$, or $1-\left(\frac{1}{2}\right)^{k+1}$.

Knowing a variety of approaches to generating a proof and knowing which one to select for a particular circumstance is an important aspect of justifying and proving.

Another element of justifying and proving is evaluating the validity of a purported proof. This involves not only deciding whether a proof is valid in terms of its conclusion, but also whether a given proof relies on correct assumptions, makes use of merited conclusions and logic, and explains the entire statement or conclusion. These skills can be fostered by challenging students to judge the appropriateness of a given argument (e.g., a formal or informal proof; Knuth, Choppin, & Bieda, 2009). Some NAEP items could be adjusted or expanded to include evaluating the justifications or proofs of others. For instance, the grade 8 NAEP item in Exhibit

3.15 addresses the question of maximizing the probability of landing on blue. Although a correct response is specific to the provided spinners, students engage in a fundamental aspect of justifying by using the definition of probability to form the requested explanation.

Exhibit 3.15. Grade 8 NAEP Probability Spinners Item



Lori has a choice of two spinners. She wants the one that gives her a greater probability of landing on blue.

Which spinner should she choose?

◯ Spinner A ◯ Spinner B

Explain why the spinner you chose gives Lori the greater probability of landing on blue.

Original item was grade 4.

Asking students to explain *why* the spinner they chose gives Lori the greater probability of landing on blue would foreground justifying. Students could also be given a version of this task in which other students' explanations for choosing Spinner A are provided, and then be asked which of the explanations is the most convincing to them and *why* it convinces them. Versions of the examples below might be offered as text, or by avatars, or through video.

1. Andreas says Spinner A has a greater chance for landing on blue because it has three blue sections and Spinner B only has one blue section.
2. Basil says that Spinner A will have a greater probability of landing on blue because the area of two of the blue sections on Spinner A is equal to the area of the one blue section on Spinner B.
3. Calista says that Spinner A has a greater chance of landing on blue because she tried it out. Calista spun each spinner 10 times. For Spinner A, the arrow fell on blue 6 times. For Spinner B, it only fell on blue 2 times.
4. Dora says that Spinner A will have a greater probability because it is one-half blue, but Spinner B is only one-third blue and one-half is more than one-third.

Engaging in justifying and proving is a way for students to explore why a particular assertion must be true. Granted, some proofs might only serve to verify the truth of a statement without helping students understand why; researchers refer to these as "proofs that prove" rather than "proofs that explain" (Hanna, 1990). Certainly not all proofs are explanatory, but in many cases, justifying or evaluating a given argument can help students understand why a conjecture is true. While investigating the reasons a conjecture might be true, students attend to particular features and consider relationships, examine multiple factors that are relevant to the problem statement, return to the meanings of terms and operations, or notice similarity or difference across cases. By exploring these factors, students gain new insight into the conjecture or deepen their understanding of fundamental mathematical ideas.

The grade 8 algebra item in Exhibit 3.16 foregrounds generalizing but could be revised into a justification task. In the item as given, the pattern that the number of diagonals d is equal to the number of sides $n - 3$ is readily apparent from the provided cases. However, adding a prompt asking *why* the equation $d = n - 3$ is a reasonable conjecture for any convex polygon would foreground justifying and proving. A valid justification might involve drawing a few cases, reasoning that from any given vertex one cannot draw a diagonal to itself and one cannot draw a diagonal to the two adjacent vertices (because this makes up two of the sides of the polygon), which means that three of the vertices cannot have diagonals drawn to them while the remaining vertices can.

Exhibit 3.16. Grade 8 NAEP Algebra Generalization Item

From any vertex of a 4-sided polygon, 1 diagonal can be drawn.

From any vertex of a 5-sided polygon, 2 diagonals can be drawn.

From any vertex of a 6-sided polygon, 3 diagonals can be drawn.

From any vertex of a 7-sided polygon, 4 diagonals can be drawn.

How many diagonals can be drawn from any vertex of a 20-sided polygon?

Answer:_________________

The item in Exhibit 3.16 also could be revised into a task to justify why the total number of diagonals that can be drawn for any given convex polygon is $n(n - 3) / 2$. Justifying could take the form of first describing why the number of diagonals that can be drawn from a vertex is $n - 3$ (as above) and then reasoning that since there are n vertices, one could draw $n(n - 3)$ diagonals. However, this would mean that each diagonal would be drawn twice, to and from each vertex. Therefore, in order to avoid double counting the diagonals, one must divide by 2, yielding the expression $n(n - 3) / 2$. To further illustrate the difference between a proof that proves and one that explains, note that the expression for the total number of diagonals can also be proved by induction. Such a proof by induction would verify the statement without revealing why it is true.

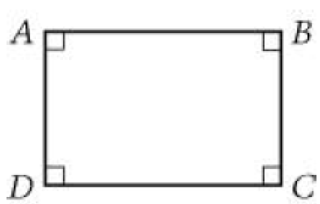
As seen in the examples in this section, the practice of justifying and proving goes beyond explaining how one knows something. That is, for the purposes of the NAEP Mathematics Assessment, explanations and justifications are not always equivalent.

The grade 8 item in Illustration 3.15 asks students to explain why a drawn shape is or is not a parallelogram. A correct response is based on comparing Sara's figure to known characteristics of a parallelogram. While the question asked may serve as a starting place for the student to consider properties of every parallelogram, the required response is specific to only the drawn shape. As stated previously, descriptions from or about examples do not suffice as a mathematical justification. Therefore, this item does not assess justifying and proving. While an item requiring a justification could start from a single example, the response would need to extend to the set of all parallelograms, beyond the constraints of the provided example.

Illustration 3.15. Justifying and Proving Nonexample: Does Not Require a Complete Application of the Definition of a Parallelogram

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Geometry	Other	Geom – 1.c	SCR

Sara was asked to draw a parallelogram. She drew the figure below.



Is Sara's figure a parallelogram? Why or why not?

Scoring Information

Key	Yes, it is a parallelogram. Some possible correct explanations (any response may contain more than one of these): • Opposite sides are parallel. • Opposites sides are equal in length. • It's a rectangle and all rectangles are parallelograms. • All angles are equal (which would imply it's a rectangle). • Opposite angles are equal in measure. • Adjacent angles are supplementary.

This item was revised for the purpose of this document. The original version of this item appeared in the 2007 NAEP Mathematics Assessment with NAEP Item ID 2007-8M11 #13 M106901.

Justifying and proving can help students develop a new and deeper understanding of the mathematics content at hand. Making sense of others' justifications or proofs – and determining their validity – can help students generate new ideas, conjectures, and generalizations, or can support their efforts to develop a new theory to be tested. That is, justifying and proving is an important mode of communication. Proofs can reveal the tools, strategies, modes of thinking, and resources used by those who created them.

NAEP Mathematical Practice 4: Mathematical Modeling

> *Mathematical Modeling: Making sense of a scenario, identifying a problem to be solved, mathematizing it, and applying the mathematization to reach a solution and checking the viability of the solution in developmentally and mathematically appropriate ways.*

Mathematical modeling involves student choice, including the assumptions made in the posing of answerable questions in an open-ended situation. The practice of modeling requires students to make sense of a scenario, identify a problem to be solved, mathematize it, and apply the mathematization to reach a solution and check the viability of the solution. Mathematical modeling also requires discussions and decisions about what is valuable (Burroughs & Carlson, 2019).

At an introductory level, modeling involves steps such as selecting and applying mathematical processes or expressing mathematical concepts and processes (such as mathematical operations) using visual, physical, or symbolic representations. At a more advanced level, a series of processes may be needed to mathematize a messy real-world situation prior to selecting and applying the mathematics. Follow-up work can involve analyzing and evaluating the results obtained from doing the mathematics. A full cycle in the mathematical modeling process includes: (a) identifying the problem; (b) making assumptions that often simplify the problem and then identifying variables; (c) mathematizing the situation; (d) analyzing and assessing solutions; and (e) translating the solution(s) back into the real world and examining their feasibility, and if not feasible, changing the simplifying assumptions and iterating the process. Finally, if there seems to be a feasible real-world solution, there are two additional steps: (f) implementing the model; and (g) reporting out results (Garfunkel & Montgomery, 2016, pp. 12–13).

It is important to distinguish between the process of mathematical modeling and the noun "model," which is an object and a term sometimes used as a synonym for a mathematical representation. For example, when a line or other function is fitted to a bi-variate scatter plot, the function is referred to as a model for the data, meaning a representation of the data. However, the practice of mathematical modeling involves far more than just using a representation. As described above, mathematical modeling is a multistep process, which may involve aspects of representing, particularly building or interpreting a representation. However, the NAEP Mathematical Practice of Mathematical Modeling is distinct from that of Representing in that the use of representations in modeling is necessarily in service of the overarching purpose of identifying and finding solutions for problems in real-world situations. For the purposes of assessment and item development, tasks that assess mathematical modeling may call upon the use of representations but representing is not the primary focus of such a task. Rather, items assessing the NAEP Mathematical Practice of Mathematical Modeling may focus on multiple steps of the cycle of mathematical modeling driven by that overarching purpose. For example, given an open-ended situation, students could generate questions they would need to explore or identify some assumptions as they begin the modeling process. In such scenarios, students would engage in the first two steps of the modeling process.

Scenario-based tasks are particularly useful in assessing student achievement in the practice of mathematical modeling. Consider the Lunch Problem scenario in Exhibit 3.17 (based on Garfunkel & Montgomery, 2016, pp. 32–35).

Exhibit 3.17. Grade 4 Example: Adaptation of GAIMME Lunch Problem Scenario

> *[Task is introduced through video: A school food service director states during the morning announcements that the school is planning a "Garden Bar" as an option for school lunch <video/image of a garden bar with a variety of fruits and vegetables>] The director says, "The cafeteria staff and I would like your input, so we know that the fruits and vegetables included will be eaten. To assist us in our decision-making process, we are establishing a task force to help us gather your suggestions and will take your suggestions into account when making our decision."]*
>
> You volunteer for the task force.
>
> At the first meeting, the team works to determine what they need to know and how to go about gathering that information. Some of the questions your team identifies are:
>
> "How many students are in the school? Do students like some of these choices more than others? Do some of these choices cost more than others? If so, which ones might we have some left over, which might we run out of? Should the school's cost of these items be considered?"

From the scenario launch, several questions might be asked. Students who address these questions would be engaging in steps 1 and 2 of the modeling cycle (identifying the problem and making assumptions).

Other tasks built from a similar scenario, about a pizza party for a grade 8 class, could be posed in different ways depending on the aspect(s) of the modeling process being assessed. For example, grade 8 students could be given the open prompt: "How many and what types of pizzas should be ordered for an 8th grade party?" Some possible questions for students to address as they attempt to model this situation are: "How many students do we expect to feed? How can we find out what types of pizza they like? Should we survey some of the students? How do we decide who to survey? What size pizzas should we order? What is the cost of each size of pizza?" Here students would need to devise a survey (identify the problem), narrow down to choices of pizza and sizes of pizza (make assumptions; identify variables), and as they begin to investigate costs of sizes and types of pizza, they would need to create estimates for the cost of the party (mathematize the situation; analyze and assess solutions).

At grade 12, a similar scenario-based open-ended task might include items based on a scenario such as: "What is the best type of computer for the school district to order for students to use in computer labs?" Some possible issues students may need to address as they attempt to model this situation are: "How many computers are needed in a school lab, and how do we know? Is there a break on cost if a large number of computers are purchased at the same time? Which types of classes will need access to the computers? What types of software will be needed for the classes?

Do any of the companies offer deals for software along with the computer purchase? How much money can be spent per student?" There are many decisions to be made about what to include and what to assume to address this task. The problem also evokes initial mathematization processes when students ask questions like: "How much money per student?" or "Are there deals for software inclusion or a price break on a large order?"

Exhibit 3.18 is an example where some initial information is provided and students could work to develop a mathematical model (possibly in teams). While the task as posed does not involve the complete modeling cycle from (a) to (g), it does assess parts of (b), (c), and (d) (identify variables; mathematize the situation; analyze and assess the solution). Exhibit 3.18 demonstrates how an item that has scaffolded the early work of (a) and (b) in the modeling process will assess only a limited part of mathematical modeling practice. Similarly, absent in Exhibit 3.18 are the creative and interpretive aspects (e), (f), and (g). At the same time, students do have choice in how to calculate an effective tax rate and leeway in interpreting and explaining the feasibility of their approach – in how they defend their position on the highest effective tax rate. An extension of the item to include (e) (translating for applicability) might also involve building a general symbolic model for the effective tax rate.

Exhibit 3.18. Grade 12 Example: Modeling Income Tax Scenario

In one state, individuals with an income of $10,000 or less per year pay no income tax; those with income greater than $10,000 per year pay a 6% tax on all income over $10,000. The effective tax rate is defined as the percent of total income that is paid in tax.

(a) What would a resident who made $40,000 pay in tax? A resident who made $50,000?
(b) What would the effective tax rate be for the resident with the $40,000 income? For the resident with the $50,000 income?
(c) Find a way to calculate the effective tax rate in general, for any resident in the state.
(d) Is there a highest effective tax rate a resident could pay? Defend your position on the highest possible tax rate.

Access to digital modeling tools, such as graphing tools and spreadsheet tools, would be important in the assessment of students' modeling practice on tasks like Exhibit 3.18, especially for parts c and d. The effective tax rate (ETR) can be expressed as the ratio of tax T to income I, or T/I (identify variables). Students first need to compute the tax on income I, with the given 6% rate after the first \$10,000 of income, arriving at $T = 0.06(I - \$10{,}000)$ (mathematize the situation). A symbolic model for the effective tax could be $\text{ETR} = T/I = 0.06(I - 10{,}000)/I$ (mathematize the situation). To answer questions about the highest possible tax rate, students could create a graphical model of ETR as a function of income, I. The mathematization process for this task starts with decisions about using ratios and percent and then could evolve to developing an algebraic expression to model ETR or even a graph of ETR as a function of income (analyzing and assessing the solution).

Modeling processes also often arise in data analysis and statistics. The task in Exhibit 3.19 is an example taken from the on-line bank of tasks available from Levels of Conceptual Understanding in Statistics (LOCUS, 2019).

Exhibit 3.19. Grade 8 LOCUS Data Modeling Task

> The student council members at a large middle school have been asked to recommend an activity to be added to physical education classes next year. They decide to survey 100 students and ask them to choose their favorite among the following activities: kickball, tennis, yoga, or dance.
> (a) What question should be asked on the survey? Write the question as it would appear on the survey.
> (b) Describe the process you would use to select a sample of 100 students to answer your question.
> (c) Create a table or graph summarizing possible responses from the survey. The table or graph should be reasonable for this situation.
> (d) What activity should the student council recommend be added to physical education classes next year? Justify your choice based on your answer to part (c).

As posed, this task covers the entire modeling cycle and closely follows the four-stage statistical investigation process as outlined by Franklin and colleagues (2007): (a) identifying a statistical question for investigation; (b) gathering appropriate data; (c) analyzing the data; and (d) communicating the results. The task assesses several content objectives in the data analysis, statistics, and probability area, including posing a statistical question, addressing issues of bias in surveys, and creating tables and graphical representations of data. Though the task as written addresses a full modeling cycle, some parts could be supplied to students and then students could be asked to engage in a narrower aspect of the modeling process.

The mathematical literacy-focused modeling task in Illustration 3.16 was adapted from a water crisis task developed for use with teachers (Aguirre et al., 2019). In this task, students are asked to think as a member of a community working to solve a problem. They are not asked to work through the entire modeling process, which would take more time than a scenario-based task would allow. Instead, the content is scaffolded to provide access to aspects of the modeling process as a path to a possible solution to the question posed. Through the task, students need to determine variables of interest, analyze the model presented for the community, and translate this model to the science club members' town.

Illustration 3.16. Mathematical Modeling Example: Scenario-Based Task with a Mathematical Literacy Context

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Data Analysis, Statistics, and Probability	Mathematical Modeling	Data – 3.c Data – 3.d	SR – MS, MC SCR

Members of the science club saw a news clip about a water crisis.

Science club members began the process of answering the question: "How many bottles of water would be needed to supply drinking water to each person in our town should our water supply become harmful to drink?" They create a mathematical model to help them answer this question.

1. Which variables **must** the science club members consider when creating a model to answer the question? Select **all** that apply.

   a. the cost of a bottle of water
   b. the amount of water a person drinks *
   c. the weight of the water to be supplied
   d. the time needed to distribute the water
   e. the number of people in the community *
   f. the length of time the bottled water is needed *

2. The science club members begin exploring the amount of water needed. Which question would **not** help the members determine the number of bottles of water that would be needed?

   a. What is the amount of water in each bottle?
   b. Does every person need the same amount of water?
   c. Does bottled water taste better than water from a faucet? *
   d. What amount of water does a person need to drink each day?

3. The science club members want to learn what they can from the water crisis. They read that from May 2018 through the end of August 2019, one company donated over 6.5 million bottles of water to the community.

   - Write **two** additional questions to which the members need answers so they can determine how many bottles of water the community actually needed.
   - Explain how knowing answers to these two questions will help the science club members in the process of answering their question.

4. There were about 96,000 residents in the community in 2018. To meet the drinking water needs of that community required more than 25 million 0.5-liter bottles.
   The science club members' town has a population of about 4,000 people.

   Estimate the number of 0.5-liter bottles needed to meet the town's drinking water needs. Justify your response.

A real-world situation such as a water crisis provides a wealth of material from which a modeling task can be built that calls on students to determine and apply relevant information to solve a problem.

Not all representations of mathematical thought address a component of the modeling process. The grade 8 item in Illustration 3.17 (from which Exhibit 3.9 was adapted) asks students to list all of the possible outcomes of flipping three coins. Absent from this released item are key aspects of mathematical modeling discussed previously, including student choice and discussions and decisions about what is valuable. To address the practice of mathematical modeling in a coin-flipping situation, a more open-ended task could be developed in which student thinking is in service of the overarching purpose of identifying and determining a solution for a problem in a real-world situation. For example, an item could state, “Someone puts an unfair coin in a stack with 5 fair coins. All of the coins look identical. Create a process that uses only coin flips and mathematics for determining which of the coins is the unfair coin.”

Illustration 3.17. Modeling Nonexample: List Possibilities without Connection to the Modeling Process

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Other	Data – 4.e	SCR

A nickel, a dime, and a quarter are flipped at the same time. Each coin can land either heads up (H) or tails up (T). List all the different possible outcomes for this event in the chart below. The list has been started for you.

Nickel	Dime	Quarter
H	H	H
H	H	T

Scoring Information	
Key	HHH – given HHT – given HTT HTH THT TTH TTT THH

This item was revised for the purpose of this document. The original version of this item appeared in the 2013 NAEP Mathematics Assessment with NAEP Item ID 2013-8M3 #2 M1499E1.

As a further example, reconsider the item in Illustration 3.14 (p. 118) discussed in the context of justifying and proving. In the item stem, students are given several constraints, including the graph from which their response should be built, minimizing the need for mathematization of the problem situation. The closed nature of the task does not allow for choice in determining the problem or for the identification of assumptions or variables. What the item does do is allow students to answer a question for a provided graph. Therefore, the item in Illustration 3.14 more closely resembles statistical reasoning and not the NAEP Mathematical Practice of Mathematical Modeling.

As mentioned, the terms “represent” and “model” are often used interchangeably. For the purposes of NAEP, creating or using a mathematical representation may indicate the practice of representing, but may or may not be invoked by the practice of mathematical modeling. For example, “Use an algebraic model to estimate height” was the description in the NAEP Questions Tool of the item shown in Illustration 3.18. In this item, students use a given representation to solve a problem, assessing the practice of representing. However, the practice of mathematical modeling is not assessed in part because students are not asked to situate the equation within the modeling cycle.

Illustration 3.18. Modeling Nonexample/ Representing Example: Evaluating a Formula to Answer a Question

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Algebra	Representing	Alg – 4.f	SR – MC

Archaeologists measure the lengths of certain bones to estimate a dinosaur’s height. When the length t of the tibia, or leg bone, is known, a dinosaur’s height h can be estimated by the following formula, where t and h are in centimeters.

$$h = 73 + 2.5t$$

If the length of the tibia of a certain dinosaur is 400 centimeters, what is its estimated height in centimeters?

A. 402.5

B. 473

C. 475.5

D. 1,000

E. 1,073

Scoring Information	
Key	E. 1,073

This item was revised for the purpose of this document. The original version of this item appeared in the 2013 NAEP Mathematics Assessment with NAEP Item ID 2013-8M3 #14 M151101.

Note that items developed using a definition of modeling other than that described in the Framework may not assess the NAEP Mathematical Practice of Mathematical Modeling. For example, the definition of “modeling” used in PARCC item development has more in common with the NAEP Mathematical Practice of Representing than the NAEP Mathematical Practice of Mathematical Modeling. Attending to the requirements for representing and modeling will be useful in item development and allow for distinguishing NAEP Mathematical Practices assessment intent.

Illustration 3.19 shows the first part of a released PARCC item assessing PARCC’s definition of modeling but not the NAEP Mathematical Practice of Mathematical Modeling. In terms of the NAEP Mathematical Practices, the request in the item for a “model” calls for students to engage in representing, not modeling. The structure of the item provides too many constraints within which students must work, minimizing the need for mathematization of the problem situation or other substantive engagement in the cycle of mathematical modeling defined for the 2025 NAEP

Mathematics Assessment. The information provided in the item stem is sufficient to develop a verbal description or algebraic representation that can be used to determine a solution to the posed problem. Absent are allowances for student choice and discussions and decisions about what is valuable. Additionally, in contrast to the item in Exhibit 3.18, the item in Illustration 3.19 does not require analysis and assessment of the solution. Therefore, while this item may be an example of “modeling” for PARCC, it is **not** an example of “mathematical modeling” for NAEP. Attention to this nuance is included here because item writers often develop material for several different assessments and item development context (i.e., which assessment) matters.

Illustration 3.19. Modeling Nonexample/ Representing and Mathematical Literacy Example

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Algebra	Representing	Alg – 4.d	ECR

A family compares the costs of renting a truck from two different companies for its 2-day move to another state. The costs are shown in the table.

Truck Rental Costs

Item	Company X	Company Y
base rental charge	$29.95 per day	$19.95 per day
mileage charge	59 cents per mile	79 cents per mile
drop-off charge	$150	included
insurance	$18 per day	$26 per day

Part A

Create a model that can be used to determine the rental cost of each truck for the 2-day move. Describe the process you used to determine your model.

Use your model to determine the number of miles when the rental costs of the two trucks will be equal.

Enter your answers in the space provided.

- Math symbols
- Relations
- Geometry
- Groups
- Trigonometry
- Statistics
- Greek

Scoring Information

Key	Student response includes the following 4 elements. • Valid definition of variables • Valid model of the rental costs for Company X • Valid model of the rental costs for Company Y • Correct number of miles when the rental costs of the two trucks will be equal

This item was revised for the purpose of this document. The original version of this item appeared in the 2018 PARCC as Item ID VH145748.

Although modeling tasks – especially separate aspects of the modeling process – could be posed to individual students, in the workplace mathematical modeling is often done in teams. The importance of preparing students to solve problems is regularly identified as a 21st century skill. The U.S. Department of Labor (2010) has noted:

> The ability to work as part of a team is one of the most important skills in today's job market. Employers are looking for workers who can contribute their own ideas, but also want people who can work with others to create and develop projects and plans (p. 57).

In school mathematics, students already often work together in groups on mathematical tasks and a mathematical modeling situation provides an inviting context for the use of collaborative tasks. The practice of mathematical modeling is also a natural place to use scenario-based tasks. Many of the sample tasks provided in this section could best be done by groups or pairs students. When a task is worthy of group effort, the assessment could focus on group responses, solutions, and problem-solving activity. Such an assessment approach is central to the final practice of the NAEP Mathematics Framework, collaborative mathematics.

NAEP Mathematical Practice 5: Collaborative Mathematics

> *Collaborative Mathematics: The social enterprise of doing mathematics with others through discussion and collaborative problem solving whereby ideas are offered, debated, connected, and built-upon toward solution and shared understanding. Collaborative mathematics involves joint thinking among individuals toward the construction of a problem solution in developmentally and mathematically appropriate ways.*

Collaborative mathematics refers to the talk and actions students engage in with one another as they participate in a necessary collaboration, where the mathematics is too complex or messy for an individual to meet its demands alone (Fiore et al., 2017). As a practice, collaborative mathematics exists alongside other mathematical practices. That is, as students work together towards a shared goal, they may also engage in representing, abstracting and generalizing, justifying and proving, and mathematical modeling. Assessing collaborative mathematics requires developing items that foreground and require the doing of mathematics collaboratively, engaging processes that are fundamentally about *joint thinking* (Teasley & Roschelle, 1993). Collectively, these processes include sharing ideas with others; attending to and making sense of the mathematical contributions of others; evaluating the merit of others' ideas through agreement or disagreement; and productively responding to others' ideas through building on or extending ideas and connecting or generalizing across ideas.

Collaborative mathematics processes are largely understood as discursive in nature and occurring through social interaction during mathematical activity. The NCTM's policy documents reflect a long-standing focus on discourse and communication. Beginning with the Mathematics as Communication standard (NCTM, 1989) and attention to discourse (NCTM, 1991), mathematics educators have argued that when students write and talk about their thinking, they not only clarify their own ideas, but they also offer valuable information for assessment.

Given the discursive nature of collaborative mathematics, NAEP Mathematics Assessment items that measure collaborative processes should likewise be discursive in nature, offering students examples of social interaction or imagined utterances around mathematics to which they are tasked to respond in key ways. These include being asked to make sense of others' thinking, express and defend agreement or disagreement, and extend an idea. Tasks might also be genuinely collaborative in nature, asking assessed students to work together in a team during the assessment, such as on a mathematical modeling task.

The discursive nature of collaborative mathematics also means that it is highly contextualized activity, tied to cultural ways of working together both in and out of the classroom. As stated in the opening of the chapter, while state standards have long included mathematical practices, and collaboration among students has long been emphasized, instruction that engages students in mathematical practices generally, and through collaborative activity in particular, may not yet be pervasive. Without careful attention to opportunities to learn, the assessment may privilege particular out-of-school cultural repertoires for collaboration, particularly around critique.

The assessment of collaborative activity is not new. The international *Programme for International Student Assessment*, for example, assesses collaborative problem solving, defined as:

> the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills, and efforts to reach that solution. (OECD, 2017, p. 6)

As illustrated in the components from a PISA scenario-based collaborative problem-solving task (Exhibits 3.20 and 3.21), the task structure involves a dialogue between a team of avatars and the assessed student. The problem task is on the right of the screen, while the running dialogue is on the left (Exhibit 3.20). The assessed student is to choose a discursive response to productively move the collaboration forward. In the example offered in the subsequent screenshots in Exhibit 3.21, one can see that the components of the task emerge as interactional contributions are offered by each avatar ("Brad" and "Rachel") and the assessed student ("you") through item response choices.

Exhibit 3.20. Sample PISA Collaborative Problem-Solving Item

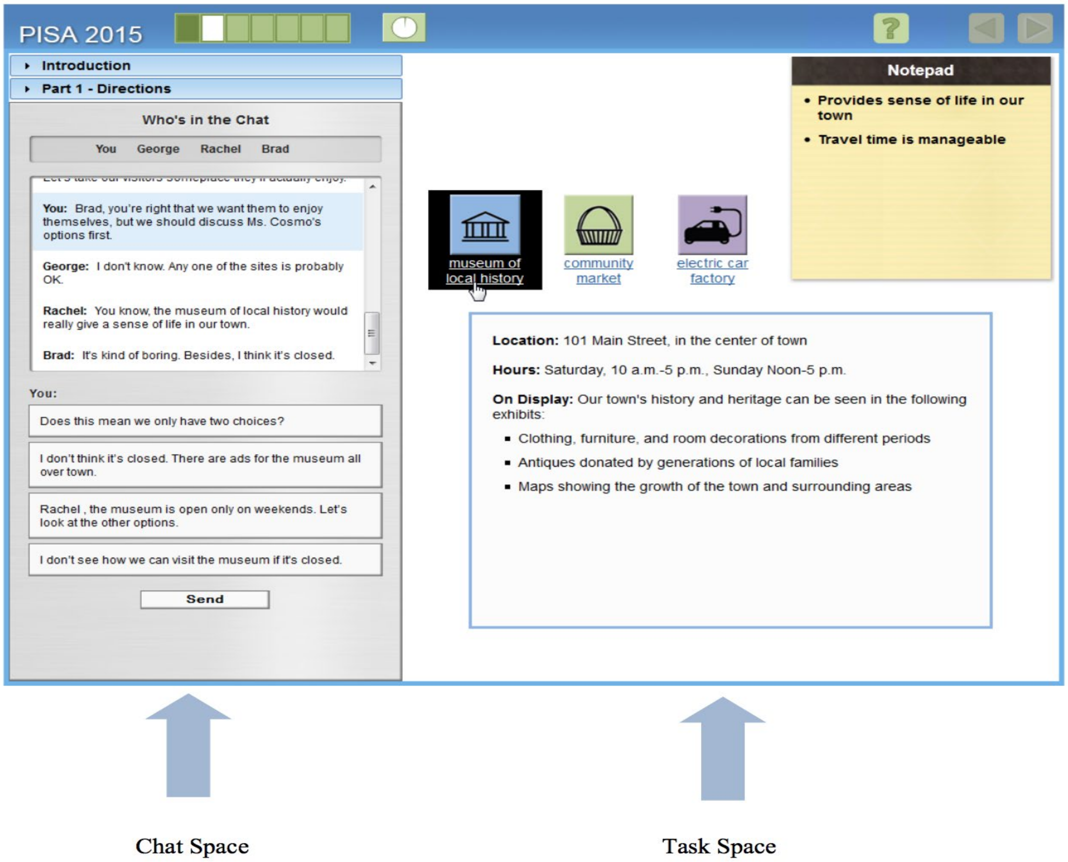
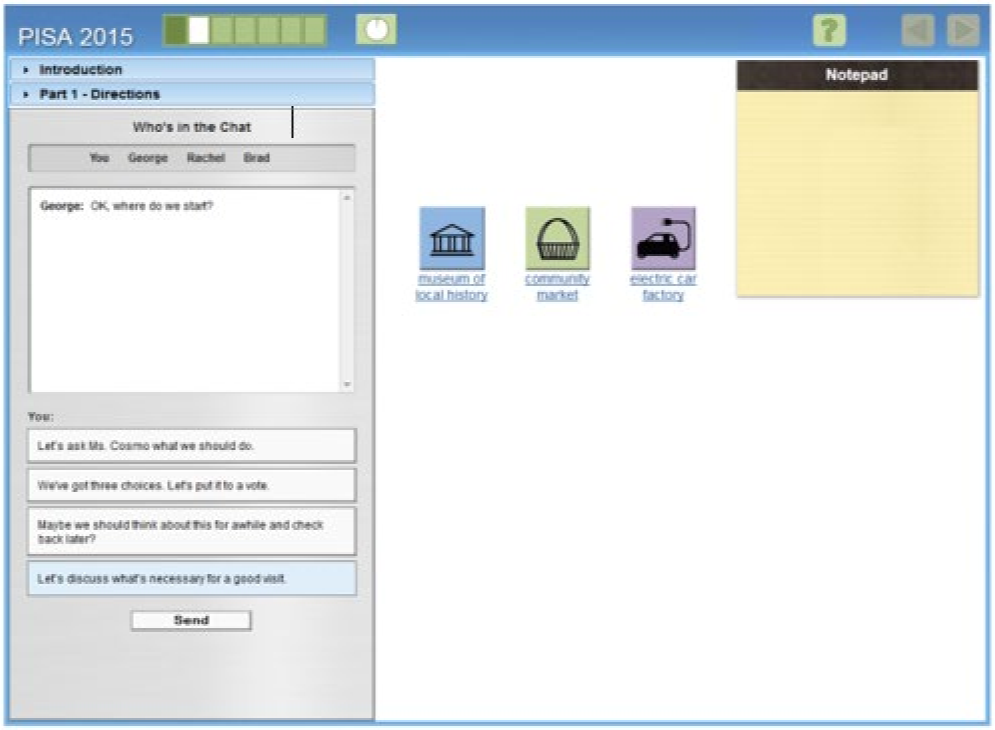
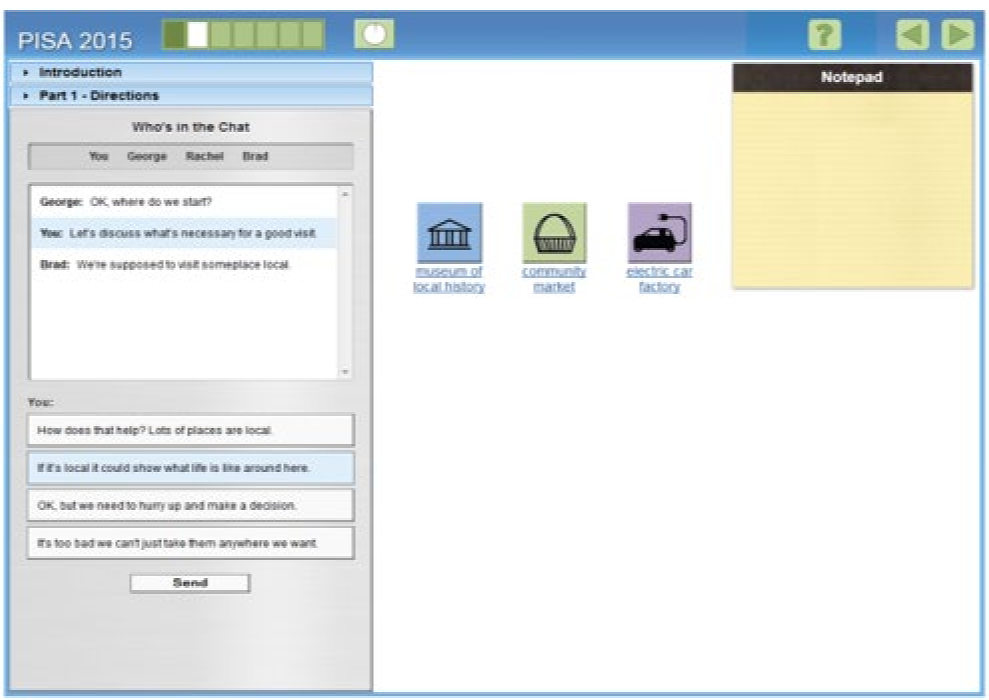


Exhibit 3.21. Example PISA Collaborative Problem-Solving Interaction



Brad mentions that the group is supposed to visit someplace local.



While PISA collaborative problem-solving items are helpful in highlighting discursive assessment, PISA items are not specifically focused on mathematics. Rather, PISA assesses three generic collaborative problem-solving competencies: establishing and maintaining a shared understanding; taking appropriate action to solve the problem; and establishing and maintaining team organization. Additionally, PISA's collaborative problem-solving items are intended to assess problem solving competencies such as exploring and understanding; representing and formulating; planning and executing; and monitoring and reflecting.

Some of these competencies may apply to collaborative mathematics, but the aim for NAEP is to assess the collaborative processes involved in *mathematics* in particular. The following sections describe three measurable skills involved in collaborative mathematics:

- attending to and making sense of the mathematical contributions of others
- evaluating the mathematical merit of the contributions of others
- responding productively to others' mathematical ideas

Collaborative mathematics begins with the sharing of ideas in the form of a conjecture or other contribution that is meant to be communicated to others. A first joint act is made up of both this sharing and how others attend to the conjecture and make sense of it (Forman, Larreamendy-Joerns, Stein, & Brown, 1998). To do so, students must establish a shared understanding about what the problem is and how the problem is being interpreted (Lerman, 1996).

While classroom studies document the importance of making sense of peers' ideas during collaborative mathematics activity, most research on the discursive processes in making sense of student thinking has looked at teacher talk moves rather than student talk moves (Chapin, O'Connor, O'Connor, & Anderson, 2009). These moves are nevertheless relevant in framing how students make sense of one another's mathematical thinking. For example, people *elicit* and *probe* ideas. Individuals then express and check personal understanding of another's thinking by repeating or *revoicing* the idea (Enyedy, et al., 2008). During a collaborative mathematics assessment task, students can elicit, probe, and revoice peers' ideas to demonstrate and check for understanding.

Revoicing is a particularly powerful discursive opportunity to assess whether a student has understood the mathematical contribution of others. Revoicing is defined as "when one person re-utters another's contribution through the use of repetition, expansion, or rephrasing" (Enyedy et al., 2008, p. 135). From an assessment perspective, students can be asked to revoice (or put into their own words) the expressed mathematical ideas of another student/avatar, or to justify its mathematical appropriateness.

The item in Illustration 3.20, adapted from Exhibit 3.12 in the Framework (discussed previously in the justifying and proving section of this chapter, p. 116) assesses collaborative mathematics. In both the original and the adapted item, students are asked to make sense of the mathematical contributions of others as they evaluate the correctness of the given statements. The adapted item includes an opportunity for students to consider original and revoiced statements. Note also that the names in the adapted item are different than those used in the original item. This change is to increase the diversity of the names contained in exhibits and illustrations throughout this document.

Illustration 3.20. Collaborative Mathematics Example: Revoicing
based on **Exhibit 3.12**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Algebra	Collaborative Mathematics Justifying and Proving	Alg – 4.c	SCR – composite

Two questions are shown.

Question A: If x is a real number, what are all values of x for which $x > -3$ <u>and</u> $x < 5$?

Question B: If x is a real number, what are all values of x for which $x > -3$ <u>or</u> $x < 5$?

Kala made a statement about the answers to the two questions being different.
Samir made a statement about the answers to the two questions being the same.

Two statements are shown. Label each with the name of the person who said it (Kala or Samir). Then explain whether Kala or Samir made a correct statement. You may use words, symbols, or graphs in your explanation.

_______________'s Statement	_______________'s Statement
The compound inequality $-3 < x < 5$ and the compound inequality $x > -3$ or $x < 5$ have different solution sets.	The compound inequality $-3 < x < 5$ and the compound inequality $x > -3$ or $x < 5$ have the same solution set.

Scoring Information

Key

<u>Kala</u>'s Statement	<u>Samir</u>'s Statement
The compound inequality $-3 < x < 5$ and the compound inequality $x > -3$ or $x < 5$ have different solution sets.	The compound inequality $-3 < x < 5$ and the compound inequality $x > -3$ or $x < 5$ have the same solution set.

Sample Response:
Kala's statement is correct because the solution set to $-3 < x < 5$ is all values between -3 and 5, and the solution set to $x > -3$ or $x < 5$ is the set of all real numbers.

This item was revised for the purpose of this document. The original version of this item appeared in the 2013 NAEP Mathematics Assessment with NAEP Item ID 2013-12M99 #1 M1934E1.

Once students attend to and make sense of the thinking of others, they must evaluate the mathematical reasonableness of their peer's mathematical contribution. Generally, students express their evaluation of the mathematical reasonableness of an idea through agreement or disagreement, including some explanation or justification. Agreeing or disagreeing emerges out of shared understanding (Nathan, Eilam, & Kim, 2007). This skill is critical to the development of productive mathematical argumentation. Experimental and classroom studies have found that students' ideas can be evaluated and become influential due to issues of status or authority rather than mathematics sense-making (Cohen & Lotan, 1997; Engle, Langer-Osuna, & McKinney de Royston, 2014).

Exhibit 3.22 shows a grade 4 SBAC (2018) item suited to assess the collaborative skill of evaluating the mathematical merit of the contributions of others. In the item, the assessed student is offered a strategy for solving a problem by an imagined student, Connor. The assessed student is asked to evaluate Connor's stated strategy and decide whether or not he is correct and why. Digitally based administration of this and similar items could provide the assessed student the opportunity to read or hear (through voiceover) Connor's own utterances, make sense of Connor's thinking, and then choose an evaluation with explanation. Note that hearing Connor's words does not make the item collaborative. Collaborative mathematics is tied to the nature of the item, which illustrates a very basic instance of looking into another person's strategy, requiring students to attend to and make sense of Connor's mathematical contribution and evaluate the mathematical merit of Connor's assertion.

Exhibit 3.22. Grade 4 SBAC Number Properties Collaborative Mathematics Item

3607

Conner solves 8× 16. He says, "I can find the product if I multiply 8× 15 and then add 8."

Select the statement that **best** explains if Connor's strategy is correct.

Ⓐ Connor is correct, because he can change the 16 to use an easier number to multiply, like 15.

Ⓑ Connor is incorrect, because 8× 16 is the same as 4 groups of 8, plus 4 groups of 8.

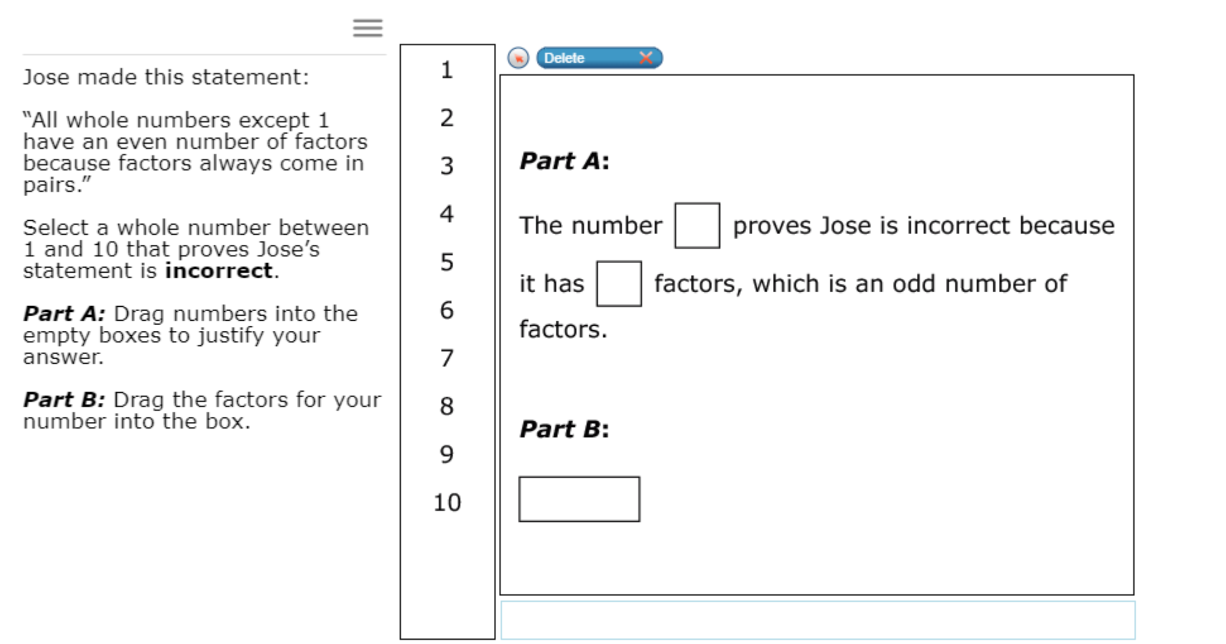
Ⓒ Connor is correct, because 8× 16 is the same as 15 groups of 8, plus 1 group of 8.

Ⓓ Connor is incorrect, because he should add 16 instead of 8.

Illustration 3.21, from Exhibit 3.23, shows another grade 4 SBAC item. Like the previous item, the assessed student is offered a glimpse into the thinking of an imagined peer, Jose. Here, Jose offers a conjecture about number. The assessed student is asked to critique Jose's conjecture by offering a counterexample that proves Jose's statement false. A digitally based assessment means the assessed student could have the opportunity to read or hear (through voiceover) Jose's own utterance, make sense of Jose's thinking, and then choose a number that proves Jose's statement is false. The item addresses the practice of collaborative mathematics through the required response to the imagined utterance by Jose. The item also addresses the practice of justifying and proving through the required completion of a counterexample to refute Jose's statement.

Illustration 3.21. **Multi-Practice Example:** **Collaborative Mathematics with Justifying and Proving**
based on **Exhibit 3.23**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Collaborative Mathematics Justifying and Proving	Num – 5.b	SR – matching



Scoring Information	
Key	4; 3; 1, 2, 4 OR 9; 3; 1, 3, 9
ALD Notes for Item Developers	
Basic	The item could be revised to require the selection of a provided description for why 4 is a factor of itself, focusing on the meaning of factor as applicable to any whole number.
Proficient	The item assesses consideration of a mathematical statement made by another in concert with understanding of factors to complete an argument that refutes the given statement.
Advanced	The item could be revised to ask for a written response to Jose that provides justification for whether his statement is true or false.

This item was revised for the purpose of this document. The original version of this item appeared in the SBAC as Item ID 3322.

Consider again the item in Illustration 3.13 (p. 117), a grade 12 NAEP Mathematics Assessment item also suited to assess collaborative mathematics. In the item, the assessed student is given an exchange by two imagined students, Tracy and Pat. As stated previously, collaborative mathematics begins with the sharing of ideas in the form of a conjecture or other contribution that is meant to be communicated to others. In this item, Tracy offers a conjecture about which Pat expresses and explains disagreement. The assessed student is asked to evaluate these utterances and decide which is correct and to explain their evaluation. In this way, students are asked to attend to and make sense of the mathematical contributions of Tracy and Pat, evaluate the mathematical merit of Tracy's and Pat's contributions, and respond productively to Tracy's and Pat's mathematical ideas. The assessed student has the opportunity to read or hear (through voiceover) Tracy and Pat's own utterances. The conversational formats in Illustration 3.13 and 3.21 are preferable to items that might offer paraphrased positions that the assessed student is tasked to evaluate.

A third mathematics-specific collective process involves responding productively to others' mathematical ideas. In particular, students learn to build on, extend, and connect across mathematical ideas. These discursive acts depend and build on the acts of making sense of and evaluating others' mathematical thinking. Once a shared mathematical idea is understood, students can further contribute to the mathematical discussion by acting upon those shared ideas. Connecting across students' mathematical ideas is a core discursive component of productive collaborative mathematics (Stein, Engle, Smith, & Hughes, 2008). By connecting ideas, students are able to notice and explain how two seemingly different strategies hold the same mathematical ideas. Students also build on or extend an idea through new examples, next steps, or logical deductions.

The grade 4 item in Illustration 3.22 has potential, but as written does not assess the practice of collaborative mathematics. In the item, students are asked to describe a next step when a first step is given in the process of subtracting. Although the item does ask students to determine next steps in determining the difference for which Mark provided a first step, students do not need to determine what Mark's first step was or evaluate the meaningfulness of Mark's work in the context of the problem. To more strongly address collaborative mathematics, the item could be revised to ask students to critique multiple statements from virtual peers, each describing a way to determine the difference. Taken in this direction, students would need to attend to and make sense of the thinking of others, evaluate the mathematical reasonableness of the presented mathematical contributions, and express their evaluation of the mathematical reasonableness of the presented ideas. Note that the items in Illustrations 3.10 (p. 111) and 3.11 (p. 112) have the potential to assess collaborative mathematics as well but would also need to be revised to address at least one of the three measurable skills for collaborative mathematics.

Illustration 3.22. Potential to Be a Collaborative Mathematics Example

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Representing	Num – 3.a	SCR – composite

$$\begin{array}{r} 143 \\ -\ 48 \\ \hline \end{array}$$

Mark needs to solve the problem shown.

He will solve the problem in two steps.

First, Mark subtracts 43 from 143.

What does Mark need to do next to complete the problem?

What is the answer to **143 − 48** ?

Scoring Information	
Key	a) Subtract 5 from 100. OR Subtract 5 from the result found in step 1. b) 95
Correct	Both parts correct
Partial	Part a) correct only OR Part b) correct only
Incorrect	Incorrect response

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-4M1 #8 M3744E0.

While the item in Illustration 3.22 does not assess collaborative mathematics, it does assess the practice of representing. In the item, students are presented with a symbolic representation of subtraction with which they engage mathematically as they consider the verbal representation of Mark's first step and ways of representing a next step in the solution process. That is, as they complete the item, students use and reason about presented representations within a mathematical context.

The practice of collaborative mathematics can also be assessed through a scenario-based task. Illustration 3.23 presents such a task, built from a classroom-based situation that also involves mathematical modeling. Much of the text in the illustration describes intended collaborative components. The mathematical content of this grade 4 task focuses on fair sharing, a common elementary mathematics activity. Similar classroom activities can provide a solid foundation from which a collaborative mathematics task can be built.

Illustration 3.23. Collaborative Mathematics and Mathematical Modeling Example: Scenario-Based Task Situated in a Classroom Setting

Grade Level	Content Area	Assessed Practice(s)	Objective IDs	Item Formats
4	Number Properties and Operations	Collaborative Mathematics Mathematical Modeling	Num – 4.e Num – 4.f	SR – MC SCR

Start with video clip or an avatar that sets up the task: Teacher shows students a bin containing same-sized cubes. The cubes are different colors. Students pose questions about the cubes in the bin: How many cubes are there? How many cubes of each color are there? What will the cubes be used for? If we share the cubes, should everyone get the same number? What if there are cubes leftover? The students decide to fair share the cubes so that everyone gets the same number of cubes.

1. The teacher or teacher avatar states that each student gets the same number of cubes and then asks for a question that can be answered. Four students each present a question. One question cannot be answered. The assessed student is asked to select from presented response options the question that **cannot** be answered. Examples of response options are:
   a. How will the cubes be used?
   b. How many cubes are leftover?
   c. How many cubes does each student get?
   d. How many cubes of each color are there?

2. An avatar sets up the process for distributing the cubes to each group of students in the room. Students or student avatars offer methods of sharing that they think are likely to lead to each student getting the same number of cubes. Videos are shown of the actions. The assessed student is asked to select a method and describe why the method is likely to lead to each student getting the same number of cubes. Two potential methods are listed.
   a. Each student reaches into the bin and grabs a handful of cubes.
   b. Each group of four students gets a scoop of cubes from the bin.

3. A video shows each group of four students getting a scoop of cubes from the bin. A group discusses how to share their cubes so that each student in the group gets the same amount. The assessed student is asked to describe a method to fair share the cubes within a group.

Video shows the students in each group with the same number of cubes, plus some leftover. Alternatively, an avatar representing each group could name the number of cubes each student in his/her group has and the number of cubes leftover. Not all groups have the same number of cubes. The teacher or teacher avatar creates a table displaying the number of cubes a student in each group has, plus the number of cubes leftover.

This table shows the number of cubes a student in each group has and the number of cubes each group has leftover.

Group	Number of Cubes Each Student Has	Number of Cubes Leftover
P	15	2
Q	12	1
R	18	1

S	14	0
T	13	0

4. The assessed student is asked to explain a way to share the cubes so that each student in the class gets the same number of cubes. As part of the response, students are asked to tell how many cubes each student gets and how many cubes are leftover.

Developing collaborative mathematics tasks can be challenging and time-consuming, but the challenges may be lessened through collaborative item development. Additionally, existing resources may provide inspiration for task development. For example, the task in Illustration 3.23 was adapted from a sharing task used in an elementary classroom (Wickstrom & Aytes, 2018).

Balance of Mathematical Practices

The target percentage range of items for each NAEP Mathematical Practice is given in Exhibit 3.24. Most NAEP Mathematics Assessment items will feature one of the five NAEP Mathematical Practices (55 to 85 percent). The range of 55 to 85 percent allows flexibility in assessment and item development across grades 4, 8, and 12 while also ensuring that the majority of the assessment is designed to capture information on student knowledge while engaging in NAEP Mathematical Practices. All NAEP Mathematical Practices will be represented in all grades and at least at the minimal levels. The relative emphasis on justifying and proving is based on its centrality across a range of mathematical activity; for example, the Smarter Balanced Assessment targets justifying across multiple content categories, including modeling and data analysis, and communicating reasoning at every grade level.

Exhibit 3.24. Percentage Distribution of Items by NAEP Mathematical Practice

NAEP Mathematical Practice Area	Percentage of Items
Representing	10–15
Abstracting and Generalizing	10–15
Justifying and Proving	15–25
Mathematical Modeling	10–15
Collaborative Mathematics	10–15
Other	15–45

The remaining balance of items (15 to 45 percent) fall into the “Other” category and will assess knowledge of content without designing the item to also assess a particular NAEP Mathematical Practice. Examples might include items that emphasize mathematical facts or procedural fluency or items that target practices that are not included in the five identified for the NAEP Mathematics Assessment. As noted earlier in this chapter, this could also include items that focus on algorithms, precision, or tool use.

Challenges

Together, the past several decades of research on mathematics thinking and learning and the consensus judgment of experts in mathematics education provide strong warrants for incorporating mathematical practices into the NAEP Mathematics Assessment. Despite widespread consensus on their importance, there are many challenges to assessing the NAEP Mathematical Practices. One is the interrelated nature of mathematical practices. Second, there is not consensus on how to define, let alone assess, mathematical practices. Finally, given the state of research and item development, it will be challenging to have sufficient numbers of items that assess student achievement with each NAEP Mathematical Practice, presenting challenges to reporting results on the Practices.

Although these challenges are formidable, they are not insurmountable. Existing state assessment programs include mathematical practices in their assessments. PISA has also been assessing mathematical practices for some time. Challenges can be addressed as the mathematical practices are incorporated into the 2025 NAEP Mathematics Assessment and refined over successive administrations. In addition, a special study to examine ways to report on mathematical practices to the general public is described in Appendix C. Despite these challenges, NAEP is clearly advancing mathematical practices as a core component of student achievement in mathematics, with the opportunity to become a leader in designing valid ways to assess the practices and report the results.

Exhibit 3.25A. Practices and Content Illustrations — Grade 4

In each cell, practice descriptors are included for a particular content area. The entries in this table are intended to be illustrative, not comprehensive. Included with some of the descriptors is italicized text providing the location of an item that is reflective of the descriptor.

Representing Grade 4				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Represent numbers or operations using visual models (e.g., base 10, number lines, fraction strips). *Illustration 3.1* *Illustration 4.13b* Recognize, translate between, interpret, and compare written, numerical, and visual representations of large numbers (e.g., thousands).	Select appropriate units related to representing or measuring an attribute of an object. Create visual representation of measurements or relationships between measurements.	Draw or sketch figures from a written description. Represent or describe figures from different views. Use a geometric model of a situation to draw conclusions.	Create a visual graphical, or tabular representation of a given data set. Compare and contrast different visual and graphical representations of a univariate distribution.	Recognize, describe, or extend numerical and geometric patterns using tables, graphs, words, or symbols. Translate between different representations of numerical expressions using symbols, tables, diagrams or written descriptions.

Exhibit 3.25A. Practices and Content Illustrations — Grade 4 (continued)

Abstracting and Generalizing Grade 4				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Identify patterns in numbers or figures and generalize patterns in written or pictorial forms. Describe or extend a pattern or relationship to a larger set of numbers. Find structural relationships among sets of numbers. Generalize understanding of place value.	Make generalizations about areas of squares or rectangles Extend quantified attributes to a larger set.	Generalize geometric properties by making connections across different figures and families of figures (e.g., triangles, quadrilaterals, polygons, polyhedra). Extend a geometric relationship from one or more figures to a family of figures.	Interpret graphical or tabular representations of data in terms of generalized phenomena (e.g., middle or median, range, mode, or shape). Make general conclusions about graphs of single sets of data (e.g., pictographs, bar graphs, dot plots).	Generalize a pattern appearing in a sequence or table, using words or symbols. *Illustration 3.7* Given a description, extend a pattern or sequence.

Exhibit 3.25A. Practices and Content Illustrations — Grade 4 (continued)

Justifying and Proving Grade 4				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Defend or counter claims about why a numerical relationship or pattern is valid or will always hold. *Illustration 3.21* Evaluate the appropriateness of an argument provided about properties or operations.	Defend or counter a claim about physical attributes, comparisons, or measurement properties. Choose a counterexample that disproves a claim about properties such as area, length, or volume.	Validate geometric conjectures (e.g., distinguish which objects in a collection satisfy a given geometric property and defend choices).	Evaluate the characteristics of a good survey and justify a survey's validity. Defend or counter conjectures offered based on a data set.	Make and justify conclusions and generalizations about numerical relationships. Given a pattern or sequence, construct, explain, or justify a rule to generate the terms of the pattern or sequence.

Exhibit 3.25A. Practices and Content Illustrations — Grade 4 (continued)

Mathematical Modeling Grade 4				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Use physical or virtual materials to build a model of a number pattern or to predict or estimate results of a continued pattern. Select and defend an appropriate method of estimation as a model for an estimation problem. Select appropriate properties or operations that can be used to build a model of a situation or solve a problem. *Illustration 3.23*	Identify the attribute(s) appropriate to measure in a given situation. Mathematize a contextual measurement situation to lead to a solution.	Use existing geometric models to solve mathematical or real-world problems.	Identify a statistical question to investigate in a given, open-ended or data-rich situation.	Identify a mathematical problem from a given situation that could be modeled numerically. Identify the variables needed to create an algebraic model of a situation.

Exhibit 3.25A. Practices and Content Illustrations — Grade 4 (continued)

Collaborative Mathematics Grade 4				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Add to a numerical model provided by others to complete a mathematical task. Evaluate others' interpretations of numbers from real-life contexts. Analyze the effect of another's estimation method on the accuracy of results.	Evaluate the validity of a measurement claim posed by others. Analyze others' solutions and suggest a critique of their solutions in a situation involving measurement. Attend to and make sense of the mathematical contributions of others in a situation involving measurement (e.g., revoice the work of others to clarify meaning of choice of measurement units).	Express and justify agreement or disagreement with a claim made by others in a geometric problem situation. Build on the work of others to geometrically model a situation.	Recognize and critique misleading arguments from data (e.g., from media or other people).	Verify the conclusions of others using algebraic/numerical properties.

Exhibit 3.25B. Practices and Content Illustrations — Grade 8

In each cell, practice descriptors are included for a particular content area. The entries in this table are intended to be illustrative, not comprehensive. Included with some of the descriptors is italicized text providing the location of an item that is reflective of the descriptor.

Representing Grade 8				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Represent word problems through visual models. Recognize, apply, create, or translate across multiple representations of fractions (e.g., visual models of equivalent fractions) and rational numbers (decimals, fractions, percents). *Illustration 4.9*	Select or use appropriate measurement instruments to determine the attributes of an object. Create visual representation of measurements or relationships between measurements.	Represent or describe figures from different views. Visualize and solve problems using geometry (e.g., using 2-D representations of 3-D objects). Use a geometric model of a situation to draw conclusions. Represent problem situations with geometric models to solve mathematical or real-world problems.	For a given set of data, create a visual graphical, or tabular representation. *Illustration 4.12* Compare and contrast different visual and graphical representations of univariate and bivariate data. *Illustration 4.18a* Justify the use of a particular representation of data over another. Interpret visual representations to compare data sets, to draw inferences, or to make conclusions across two or more distinct data sets. Create and use scatterplots to represent the relationship between two variables and to estimate the strength of the relationship (strong, weak, none).	Use or create a graphical representation of a situation to draw conclusions. Translate between different representations of expressions using symbols, graphs, tables, diagrams or written descriptions. *Illustration 3.3*

Exhibit 3.25B. Practices and Content Illustrations — Grade 8 (continued)

Abstracting and Generalizing Grade 8				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Determine an expression for a recursive pattern. Generalize, describe, or compare numerical properties and operations across different domains. Extend a pattern or relationship to a larger set of numbers. Find and generate structural relationships among sets of numbers. Generalize findings about rational and irrational numbers.	Extend quantified attributes to a larger set. Make connections between representations of different measurement systems.	Describe the general effects of dilations, translations, and rotations for two-dimensional figures. Identify common elements across different figures and families of figures (e.g., triangles, quadrilaterals, polygons, polyhedra). Extend a geometric relationship from one or more figures to a family of figures.	Interpret graphical or tabular representations of data in terms of generalized phenomena (e.g., shape, center, spread, clusters). Generalize trends in data to suggest interpretations or infer conclusions.	Generalize a pattern appearing in a sequence, table, or graph using words or symbols. Develop general rules for translating functions and graphs. Create connections across representations.

Exhibit 3.25B. Practices and Content Illustrations — Grade 8 (continued)

Justifying and Proving Grade 8				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Defend a claim about why a numerical relationship or pattern is valid or will always hold. Find a counterexample to refute a claim about number properties or operations. Evaluate the appropriateness of a provided argument about properties or operations.	Defend a claim about physical attributes, comparisons, or measurement properties. Evaluate the validity of a provided argument making use of measurement. Find a counterexample to disprove a claim about properties such as area, length, or volume.	Verify properties of rotations, reflections, or translations. Create, test and validate geometric conjectures (e.g., distinguish which objects in a collection satisfy a given geometric definition and defend choices). Defend claims about similarity of two-dimensional figures. Analyze a provided argument about geometric attributes or relationships.	Evaluate the characteristics of a good survey or of a well-designed experiment and defend the validity of surveys or experiments. Offer counter arguments in relation to conjectures about bivariate data.	Develop a valid mathematical argument based on properties of slope and intercept for linear functions. Justify functional relationships across different representational forms, such as tables, equations, verbal descriptions, or graphs.

Exhibit 3.25B. Practices and Content Illustrations — Grade 8 (continued)

Mathematical Modeling Grade 8				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Build a model of a situation for an estimation problem. Communicate and defend a decision about a physical or virtual model involving number and/or operation to an audience for feedback.	Mathematize a contextual measurement situation to lead to a solution. Evaluate the reasonableness of a model unit for an attribute in a real context.	Visually model the effects of successive (or composite) transformations of figures in the plane. Construct geometric models using physical or virtual materials to solve mathematical or real-world problems.	Identify a statistical question to investigate in a given, open-ended or data-rich situation. Create or use a statistical model to answer a statistical question or make a prediction about a data set. Create or use a statistical model to assess the validity of a statistical claim.	Identify the variables needed to create an algebraic model of a situation. Write algebraic relationships, expressions, equations or inequalities to model real-world situations.

Exhibit 3.25B. Practices and Content Illustrations — Grade 8 (continued)

Collaborative Mathematics Grade 8				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Build on a numerical model provided by others to complete a mathematical task. Analyze the effect of another's estimation method on the accuracy of results. Reflect on the work of others to extend a numerical pattern.	Evaluate the validity of a measurement claim posed by others. Engage in joint thinking to reach consensus about a measurement situation. Analyze others' solutions and suggest a critique of their solutions in a situation involving measurement.	Express and justify agreement or disagreement with a claim made by others in a geometric problem situation. Build on the work of others to geometrically model a situation. Evaluate the merit of others' geometric ideas. Connect across geometric ideas contributed by others in a problem-solving situation.	Choose a worthwhile statistical question from a set offered by others about a problem situation or context involving data. Recognize and critique misleading arguments from data (e.g., from media or other people). Revoice the work of others in addressing a statistical or probabilistic situation. Analyze the models constructed by others to evaluate a new data set.	Verify the conclusions of others using algebraic properties.

Exhibit 3.25C. Practices and Content Illustrations — Grade 12

In each cell, practice descriptors are included for a particular content area. The entries in this table are intended to be illustrative, not comprehensive. Included with some of the descriptors is italicized text providing the location of an item that is reflective of the descriptor.

Representing Grade 12				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Create and justify solutions to word problems through numeric representations and operations. Represent, interpret, or compare expressions or problem situations involving absolute values.		Represent or describe figures from different views. Visualize and solve problems using geometry (e.g., using 2-D representations of 3-D objects). Represent problem situations with geometric models to solve mathematical or real-world problems.	For a given set of data, create a visual graphical, or tabular representation of the data. Compare and contrast different visual and graphical representations of univariate and bivariate data. Interpret visual representations to compare data sets, to draw inferences, or to make conclusions across two or more distinct data sets. Create and use scatterplots to represent the relationship between two variables and to estimate the strength of the relationship (strong, weak, none).	Use or create a graphical representation of a situation to draw conclusions. *Illustration 4.18b* Translate between different representations of expressions using symbols, graphs, tables, diagrams or written descriptions. Express linear and exponential sequences in recursive or explicit forms given a table.

Exhibit 3.25C. Practices and Content Illustrations — Grade 12 (continued)

Abstracting and Generalizing Grade 12				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Determine a generalized expression for a recursive pattern. Extend properties of numbers from one system to another (for instance, extend the properties of exponents to rational exponents). Generalize, describe, or compare numerical properties and operations across different domains or number systems. Extend a pattern or relationship to a larger set of numbers. Find and generate structural relationships among sets of numbers.	Generalize the effect of proportions and scaling for area and volume. Extend trigonometric formulas to determine triangle unknowns.	Generalize relationships such as congruence, similarity, orientation between figure and their images under transformation. Extend a geometric relationship from one or more figures to a family of figures. Develop generalizations about transformations that preserve the area or volume of figures.	Interpret graphical or tabular representations of data in terms of generalized phenomena (e.g., shape, center, spread, clusters). Organize and display data in order to recognize and make inferences from patterns in the data. Notice patterns of outcomes in a probability situation. Generalize trends in data to suggest interpretations or infer conclusions. Develop generalizations about how linear transformations of one-variable data affect mean, median, mode, range, interquartile range, and standard deviation.	Extend and generalize numerical patterns, including arithmetic and geometric progressions. *Illustration 3.6* Compare and generalize properties of linear, quadratic, rational, and exponential functions. Identify commonalities within and across function families. Develop general rules for translating functions and graphs. Create connections across representations.

Exhibit 3.25C. Practices and Content Illustrations — Grade 12 (continued)

Justifying and Proving Grade 12				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Find a counterexample to refute a claim about number properties or operations. Prove numerical relationships through developing deductive arguments, engaging in proof by exhaustion, or employing mathematical induction. Evaluate the validity of a provided argument about properties or operations. *Illustration 3.13* Analyze or interpret a proof by mathematical induction about the properties of numbers. *Exhibit 3.14* Justify relationships between properties of number systems, including natural numbers, integers, rational numbers, real numbers, and complex numbers.	Justify or prove a claim about physical attributes, comparisons, or measurement properties. Explain why a given attribute can be appropriately measured by the chosen quantity and unit. Evaluate the validity of a provided argument making use of measurement. Find a counterexample to disprove a claim about properties such as area, length, or volume. Prove conjectures about trigonometric identities.	Justify relationships of congruence and similarity; apply these relationships using scaling and proportional reasoning. *Exhibit 3.11* Create, test and validate geometric conjectures (e.g., distinguish which objects in a collection satisfy a given definition and defend choices). Analyze a provided argument about geometric attributes or relationships. Use given definitions and theorems to prove geometric conjectures. Develop justifications and proofs that rely on a variety of representational modes (e.g., two-column, paragraph). Discuss the implications that a definition of a type of figure has on the figure properties.	Critique the validity of surveys or experiments. Justify or prove conjectures about probability. Create and explore counting arguments in order to develop and justify conjectures.	Create, validate and justify conclusions and generalizations about functional relationships. *Illustration 3.12* Verify a conclusion using algebraic properties. Prove algebraic relationships through developing deductive arguments, finding counterexamples, engaging in proof by exhaustion, and employing mathematical induction. *Exhibit 3.12*

Exhibit 3.25C. Practices and Content Illustrations — Grade 12 (continued)

Mathematical Modeling Grade 12				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Select appropriate properties or operations that can be used to build a model of a situation or solve a problem. Create a physical or virtual model involving number and/or operation.	Select or use a model unit for an attribute to be measured and defend the use of that unit. Mathematize a contextual measurement situation to lead to a solution. Create a model to convert between two measurement systems. Construct scale drawings to be used as measurement models of objects in problem situations.	Create a geometric model of a physical object. Discuss differences in solutions caused by having used a simplified model. Use existing geometric models to solve mathematical or real-world problems. Visually model the effects of successive (or composite) transformations of figures in the plane. Construct geometric models using physical or virtual materials to solve mathematical or real-world problems. Predict the results of combining, subdividing, and transforming geometric figures.	Identify a statistical question to investigate in a given, open-ended or data-rich situation. Use a statistical model to answer a statistical question or make a prediction about a data set. Create a probability model to calculate or estimate the probability of an event. Compare and contrast theoretical probabilities with results from experimental probabilities in a simulation.	Identify a mathematical problem from a given situation that could be modeled algebraically. Identify the variables needed to create an algebraic model of a situation. Write algebraic relationships, expressions, equations or inequalities to model real world situations. Revise an existing algebraic model based on introducing new variables or parameters. Build or apply a mathematical model of a financial situation. (e.g., a monthly family budget, or a car loan).

Exhibit 3.25C. Practices and Content Illustrations — Grade 12 (continued)

Collaborative Mathematics Grade 12				
Number Properties and Operations	**Measurement**	**Geometry**	**Data Analysis, Statistics, and Probability**	**Algebra**
Build on a numerical model provided by others to complete a mathematical task. # Analyze the effect of another's estimation method on the accuracy of results. Reflect on the work of others to extend a numerical pattern. Evaluate the mathematical reasonableness of a peer's mathematical contribution.	Evaluate the validity of a measurement claim posed by others.	Express and justify agreement or disagreement with a claim made by others in a geometric problem situation. Attend to the contributions of others in collaboratively generating a geometric proof. Build on the work of others to geometrically model a situation. Generalize across geometric ideas contributed by others in a problem-solving situation.	Revoice/restate the work of others in addressing a statistical or probabilistic situation. Analyze the models constructed by others to evaluate a new data set.	Verify the conclusions of others using algebraic properties.

CHAPTER 4

OVERVIEW OF THE ASSESSMENT DESIGN

This chapter provides an overview of the major components of the mathematics assessment design, which includes the types of assessment tasks and item formats and how they can be used to expand the ways in which students are asked to demonstrate what they know and can do in mathematics. In addition, this chapter describes how the assessment is distributed across the five mathematics content areas described in Chapter 2 and the five NAEP Mathematical Practices in Chapter 3. The 2025 Framework intentionally emphasizes increased access for students – including English language learners and students with disabilities – to demonstrate their mathematics understanding. Scholarship has demonstrated that students of various ethnic, racial, economic, and cultural backgrounds have salient differences that matter to the format and design of assessment items for inclusiveness (Solano-Flores, 2011). In particular, the 2025 NAEP Mathematics Assessment will continue to use concepts of universal design for assessment to increase inclusiveness and assessment validity (Thompson, Johnstone, & Thurlow, 2002).

Previous NAEP Mathematics Assessments included only discrete items, which stand alone or comprise a composite item. Discrete items consist of selected response and constructed response item types. In order for students to demonstrate what they know and can do with respect to the range of mathematics content knowledge and NAEP Mathematical Practices in the Framework, the 2025 NAEP Mathematics Assessment includes a new item assessment format: scenario-based tasks. Scenario-based tasks have both context and extended storylines to provide opportunities to demonstrate facility with the integrated nature of mathematics content knowledge and NAEP Mathematical Practices.

Two fundamental aims motivate the expansion. There is a need to ground the NAEP assessment in relevant tasks and familiar contexts to provide a better measure of student content knowledge and mathematical practices (Eklöf, 2010). Second, by expanding item types and thoughtfully using technology, the NAEP Mathematics Assessment continues to provide greater access to all students, diversifies the ways in which student achievement can be recognized and measured, and more robustly assesses both what students know and what they can do. For example, graphics can be presented in color with greater clarity and with a tool to zoom in and out (Sireci & Zenisky, 2006).

Technology provides opportunities for assessment, but with each opportunity come myriad constraints and repercussions that must be considered. For example, introducing a new format for items on the NAEP Mathematics Assessment that is interactive or discussion-based requires that great care be taken to ensure that the design is accessible to students, that students have ample time to understand how to engage with the item, and that students have had opportunities to experience the task type. Familiarity with digital technology in general, and with specific digital tools in particular, can influence student performance (Dunham & Hennessey, 2008). Another potential threat to assessment validity is the accessibility of tools and the affordances for students with and without certain disabilities. Due to differential access to, use of, and outcomes stemming from student experiences with technologies in and out of school (Warschauer & Matuchniak, 2010), development work should address known and potential implementation challenges and identify ways to mitigate issues of access in doing the assessment that could occur in under-resourced communities (Warschauer, 2016). A goal of the NAEP Mathematics Assessment is not to disadvantage students by virtue of the assessment's technology.

Item Development

Chapter 2 describes, for each grade level, the content objectives in each of five areas of mathematics: Number Properties and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra. Chapter 3 describes the five NAEP Mathematical Practices that are the targets for assessing mathematical activity across all grade levels: Representing, Abstracting and Generalizing, Justifying and Proving, Mathematical Modeling, and Collaborative Mathematics. Combined, the guidelines in those two chapters along with this chapter focus on translating the intent of the Framework into development of items used on the assessment.

The guidelines offered here highlight only some of the critical considerations in item development, concentrating on topics specific to the NAEP Mathematics Assessment. Item writers should refer to general directions for developing items provided by the National Assessment Governing Board and their designees in addition to the information in this document.

Item Characteristics

The specific components of an item are determined by the item format. However, two components are constant across all item formats: 1) the item stem, and 2) the response. The item stem, also known as the stimulus, is the introduction to the item and the question asked of or directive given to the students. The item stem should provide all of the necessary information for students to respond, clearly laying out for the student what is being asked and the expected response method. The response method is determined by the item format.

Illustration 4.1 provides a multiple-choice item with the main item components labeled. Note that the rationales – the support for the inclusion of the response options as correct or plausible – are utilized during item development and item review but are **not** part of a student-facing NAEP item.

Illustration 4.1. Components of a Multiple-Choice Item

The expanded form of a number is shown.

$$8\times10{,}000 + 4\times100 + 6\times10$$

What is the number in standard form?

Item Stem

A. 846
B. 8,460
C. 80,460
D. 84,600

Response Options

Rationales for response options:

Rationale A: Uses the digits 8, 4, and 6 in order of appearance, but not with associated place value.

Rationale B: Uses the digits 8, 4, and 6 in order of appearance, but writes 8 as the digit in the thousands place.

Rationale C: Correct. 80,000 + 400 + 60 = 80,460.

Rationale D: Uses the digits 8, 4, and 6 in order of appearance, but writes 4 as the digit in the thousands place and 6 as the digit in the hundreds place.

General Principles of Item Writing

NAEP items will be developed in accordance with recommended practice and the Governing Board Item Development and Review Policy (2002). See the Board's policy for principles about item writing that apply to all NAEP assessments.

Types of Tasks, Items, and Supporting Tools

The 2025 NAEP Mathematics Assessment will include existing and new discrete items as well as scenario-based tasks.

Scenario-Based Tasks

The goal of scenario-based tasks is to provide evidence of students' ways of knowing and doing mathematics. Current and future NAEP Mathematics Assessments can take advantage of evolving digital technologies to create the next generation of scenario-based tasks, as well as yet-to-be-imagined items and tasks. Other NAEP frameworks have set a foundation for scenario-based tasks. For example, since 2009 the NAEP Science Framework has called for the use of interactive computer tasks, and the NAEP Technology and Engineering Literacy (TEL) Framework has done so since its start in 2014. Examples of scenario-based tasks from TEL can be found at www.nationsreportcard.gov/tel_2014/#tasks/overview.

The defining features of the scenario-based task for the 2025 NAEP Mathematics Assessment are an authentic context, in which students can imagine themselves, with a motivating question or goal along with item design that supports exploration. The motivating goal for a scenario-based task might be to solve a particular problem or to complete a certain mission within the scenario. The goal provides the driving rationale for the tasks that the student will perform. It offers a storyline that helps build needed background, defines the task's relevance and coherence, and motivates the student to engage with the scenario-based task.

Within one scenario-based task, a student may complete multiple items that vary in format with both constructed- and selected-response item types (details of item types are provided in the next section). Within a scenario-based task, each item is in some way related to, or builds on, the next item as part of the cohesive experience. Such tasks may be well suited to addressing the intersecting nature of the mathematics content and the NAEP Mathematical Practices illustrated in Exhibits 3.25A, 3.25B, and 3.25C in Chapter 3. Scenario-based tasks may also be especially well-suited to measuring the highly iterative or interactional nature of the NAEP Mathematical Practices described in Chapter 3.

An advantage of digital delivery of the assessment is that scenario-based tasks can use multimedia (e.g., images, video, and animation, in addition to future technologies) to present the settings for the assessment items. As a result, non-mathematical linguistic demand might be reduced while maintaining mathematical rigor. Multimedia can also better scaffold the background understanding that examinees may need to complete a given item. For example, video segments or animations that a student observes, along with text, numbers, and graphics, can convey information necessary for the task to be accomplished. In developing such scenario-based tasks, related design decisions should serve a particular purpose and not be extraneous or be presented simply for visual interest. While in many cases relevant multimedia content can have a positive impact on student engagement and performance, it is also possible that it may introduce competition of attention between visual and auditory channels (Folk et al., 2015). When multimedia content is included in a scenario-based task, developers need to ensure that the multimedia is used productively and minimizes such competition.

Within a scenario-based task, students are given opportunities to select tools from a toolkit and use them to solve problems. For example, students might be asked to select a graphing or spreadsheet tool or to use a simulation. Various digital and physical tools may be made available, depending on the scenario. These might take the form of chat/texting, or presentation tools for communication tasks, if deemed relevant to the mathematical understanding being assessed.

When designing tools for a scenario-based task, it is necessary to determine which elements of a tool are needed for the activities in the scenario and which features are used by students. For example, only those functions of a spreadsheet tool that are directly relevant to a given item might be provided. It is not necessary to provide all of the other features of the spreadsheet tool. In fact, including every feature could be distracting to students and could produce measurement error. Additionally, students are not expected to know how to use all tools in a scenario-based task prior to starting the task. In these cases, instructions and practice using the tool are embedded in the task before the tool is needed or used to complete the task.

An important consideration for assessment developers when designing scenario-based tasks is to ask what is gained through the selection of a scenario as assessment context. A robust scenario will allow examinees to interact with task components in multiple ways, explore alternative outcomes and explanations, find multiple solution paths, and demonstrate their thinking. Students could also evaluate the outcomes of the choices they make and convey their understanding of mathematical concepts in diverse ways. For example, one scenario-based task may engage students in a range of mathematical practices and foreground one content area.

Interactive scenario-based tasks can elicit rich data, providing evidence of NAEP Mathematical Practices that are difficult to measure with more conventional items and tasks. For example, measuring collaboration has long been a challenge in assessment. Novel methodological approaches have explored discipline-specific student collaborative activity through the use of performance outcomes and process data from scenario- and simulation-based collaborative assessment (Andrews et al., 2017). These approaches can be used to better assess the NAEP Mathematical Practice of Collaborative Mathematics.

As illustrated in the PISA example in Chapter 3 (see Exhibit 3.20, p. 134) validated scenario-based tasks that assess collaborative problem solving already exist. In that example, the task was structured as a dialogue with a collaborative team made up of avatars and assessed students in a way that is nearly impossible to do using only discrete item sets. In contrast, Exhibit 4.1 (based on a grade 8 Stacking Chairs task from the Silicon Valley Mathematics Initiative, 2016) illustrates a set of discrete items that are scenario-based presented in a non-digital environment. Notably lacking from this example are supporting multimedia and tools.

Exhibit 4.1. Grade 8 Scenario Example

You, Lee, and Pat are the team organizing the spring concert at your school. The school has a large room with a stage but the team will need to arrange for renting chairs from a local company. The chairs must be put in a storage room before the concert. The chairs can be stacked. The team stacked some chairs and measured the heights of the stacks. Below are the notes the team made.

The height of stacked chairs

5 chairs are 51 inches high

3 chairs are 45 inches high

8 chairs are 60 inches high

1. How tall are two chairs stacked together? ________ inches

Lee suggests the chairs be stacked in groups of 10.

2. How tall is a stack of 10 chairs? ________ inches
   Show how you figured it out.

The team decides that groups of 10 chairs will take up too much floor space. The team wants an equation to know how tall a stack will be if you know the number of chairs.

3. Write an equation to find the height, *y*, if the number of chairs in a stack is *x*.

4. Explain how Pat can use the equation you wrote to determine the height of 28 chairs.

The storage room is 15 feet tall. Three feet of space above the stack of chairs is needed (to take chairs off the stack).

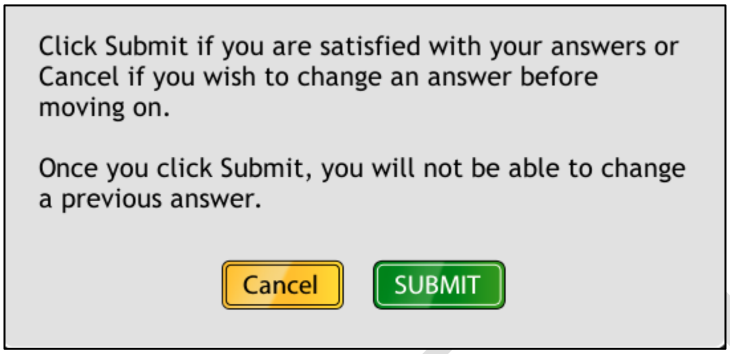
5. How many chairs can be in a stack and still fit in the storage room? ______ chairs
   Show how you figured it out.

6. There will be 200 chairs for the audience. What else would the team need to know in order to determine whether or not all 200 chairs will fit in the storage room? Why is the information needed?

Note that the response to item 4 in Exhibit 4.1 is dependent on the response to item 3. On a digitally based assessment, the task could be presented in a way that removes the dependency. In administering the item, students could be asked to review and submit their answer to item 3 before accessing item 4.

3. The number of chairs in a stack is represented by *x*. Write an equation to determine the total height, *y*, in inches, of the stack of chairs. When you are finished, click submit.

The image below shows text displayed to students during administration of TEL tasks that can be adapted for use on the mathematics assessment.



To allow for completion of item 4 without reference to a response provided for item 3, item 4 could be revised to give an equation that represents the height of a stack of *x* chairs and ask students to use the equation to determine the height of a stack of 28 chairs.

4. Lee writes the equation shown to determine the total height, *y*, in inches, of a stack of *x* chairs.

$$y = 36 + 3x$$

Explain how to use the equation that Lee wrote to determine the height, in inches, of a stack of 28 chairs. As part of your response, determine the height, in inches, of a stack of 28 chairs.

For additional examples that avoid dependencies between related item parts, see the TEL Andromeda Task: https://www.nationsreportcard.gov/tel/tasks/andromeda/.

An alternate version of the stacking chairs task is provided in Illustration 4.2. The item context puts the students in the task as part of a team determining whether chairs can be stacked in a storage room. Item content in Parts B, C, and D present content differentiation for grade 8 and grade 12. Included with these versions of the task are development notes and scoring information. Additional information on scoring is provided later in this chapter.

Illustration 4.2. Alternate Stacking Chairs Task

You, Chi, and Alma are the team organizing seating for the spring concert at your school. The audience will sit in chairs that must be put in a storage room after the concert. The team needs to determine whether all 200 chairs can be stored in the room.

The chairs are identical and can be stacked. The team stacked some chairs and measured the heights of the stacks. Below are the heights the team measured.

The height of stacked chairs

5 chairs are 51 inches high

3 chairs are 45 inches high

8 chairs are 60 inches high

Item - Part A

What is the height, in inches, of a stack of 2 chairs?
[Correct response: 42 (inches)]

Item - Part B

The team wants a way to determine the height of a stack of chairs when the number of chairs in the stack is known.

(Grade 8)
Write an equation that can be used to determine *h*, the total height, in inches, of a stack of *n* chairs.

On Screen: Click Submit if you are done with your answer or cancel if you wish to change your answer before moving on. Once you click Submit you cannot change the answer.

[Scoring Information:
Student response should be equivalent to h = 3n + 36.]

(Grade 12)
The team will stack chairs on a cart. The cart adds 18 inches to the total height of a stack.

Write an equation that can be used to determine *h*, the total height, in inches, from the ground to the top of *n* chairs stacked on a cart. Explain how you determined your equation.

On Screen: Click Submit if you are done with your answer or cancel if you wish to change your answer before moving on. Once you click Submit you cannot change the answer.

[Student response should mathematically support an equation equivalent to $h = 3n + 54$.
Sample student response: Since the height of a stack of 5 chairs is 51 inches and the height of a stack of 3 chairs is 45 inches, each additional chair increases the total height of the stack by 3 inches. Since $45 – 6 = 39$, the height of 1 chair is 39 inches. So, the height, h, of a stack of n chairs would be $39 + 3(n - 1)$ or $3n + 36$. Since the height of the cart adds 18 inches to the total height, the height of a stack of n chairs on a cart would be $h = 3n + 54$.]

Item - Part C

(Grade 8)

Chi writes the equation shown to determine *h,* the total height, in inches, of *n* chairs stacked on a cart.

$$h = 36 + 3n$$

Explain how to use the equation that Chi wrote to determine the height, in inches, of a stack of 28 chairs. As part of your response, determine the height, in inches, of a stack of 28 chairs.

On Screen: Click Submit if you are done with your answer or cancel if you wish to change your answer before moving on. Once you click Submit you cannot change the answer.

[Scoring information:
Student response should mathematically support a height of 120 inches.
Sample student response: Since n represents the number of chairs, substitute 28 for n. Multiply 28 by 3. Then add 36. The height of a stack of 28 chairs is 120 inches.]

(Grade 12)

Alma writes the equation shown to determine *h,* the total height, in inches, of a stack of *n* chairs.

$$h = 54 + 3n$$

After the chairs are stacked on the carts, they will be stored in a room that is 12 feet high. A space of 3 feet is needed above the top of each stack of chairs so that chairs can be taken off the cart.

The team has determined that no more than 10 carts can be put into the storage room. Using Alma's equation, determine whether or not all 200 chairs can be stacked on carts and stored in the room. Show your work or explain how you determined your answer.

[Scoring information:
Student response should mathematically support that all 200 chairs cannot be stored in the room. Correct response may or may not include reference to shorter doorway and vertical fit through doorway rather than vertical fit in the room itself.
Sample student response: Since 10 carts will fit in the storage room, each cart must have 20 chairs stacked on it: $200 \div 10 = 20$. Using the equation $h = 3n + 54$, the height of each stack will be 114 inches: $3 \times 20 + 54 = 60 + 54 = 114$.
Since 3 feet are needed at the top of each stack, the total height needed for each stack is $114 + 36 = 150$ inches. The height of the room is 12 feet, which is 144 inches. Since 150 inches are needed for each stack, all 200 chairs cannot be stored in the room.]

Item - Part D

(Grade 8)

The team will put 200 chairs in the storage room. What other information does the team need to know to determine whether all 200 chairs will fit in the storage room? Why is the information needed?

[Scoring information:
Student response should include information about floor space in the storage room, other dimensions of the doorway (e.g., width of opening) and other dimensions of the chair stack related to width and length of the stack (in addition to the height information). Justification might include a need for the stack to fit through the doorway of the storage room and establishing lower bounds on fit for both the doorway and the room.]

Item Development Information

Development - Part A
Objective Alignment: Algebra, 1.a
NAEP Math Practice Alignment: None

This item serves as a lead-in to the task. Although the relationship given by the heights of the chairs is linear, students may not use a linear relationship to determine the height of one chair. However, students will need to use the difference of 3 inches between the height of a stack of n chairs and the height of a stack of $n + 1$ chairs, focusing on the application of a determined pattern to answer the question asked.

Development - Part B
(Grades 8 and 12)
Objective Alignment: Algebra, 3.b
NAEP Math Practice Alignment: Abstracting & Generalizing; Representing

Although students could use the height from Part A to determine the equation, they do not need to. Instead, a student could use the difference of 3 inches between the height of a stack of n chairs and the height of a stack of $n + 1$ chairs to determine that 15 inches of a stack of 5 chairs are the seats. Since the total height is 51 inches, 36 inches are constant.

Development - Part C
(Grade 8)
Objective Alignment: Algebra, 4.a
NAEP Math Practice Alignment: none

Part C presents a correct equation and asks students to determine the height of a stack containing a specified number of chairs. This item relates to the open-ended item in Part D. Students need to provide an explanation for how the equation could be used to determine the height, which is less complex than the reasoning required to respond to Part D.

(Grade 12)
Objective Alignment: Algebra, 4.c
NAEP Math Practice Alignment: Justifying & Proving

Part C extends the thinking done in Part B by requiring students to use an equation to determine whether additional constraints can be met when placing the carts in a storage room. Students might approach this item by starting with the height of the room or the height of a stack of chairs.

Development - Part D
Objective Alignment: Algebra, 4.c
NAEP Math Practice Alignment: Justifying & Proving

Part D provides an opportunity for students to consider constraints and limitations to putting the chairs in the storage room. The open-ended nature of this question increases complexity while also allowing for the consideration of multiple measurements that impact the storage of the chairs in the room. The mathematical argument necessary to support why information is needed to determine whether the chairs can be stacked in the room addresses justifying and proving, as the provided information would be needed to determine whether any number of stacks of chairs could fit in the room. The constraints already given in the task, including how the chairs are stacked, precludes the item from addressing mathematical modeling.

Due to their capacity to replicate authentic situations (i.e., experiences that students may encounter in their lives), scenario-based tasks have the potential to provide a level of accessibility and support for student engagement with the assessment that other types of assessment tasks do not. Additionally, scenario-based tasks provide opportunities to simultaneously assess multiple practices or content areas. However, a block of scenario-based tasks may provide less measurement information than a block of discrete items in the same amount of assessment time; scenario-based tasks typically require a longer duration to reach optimal reliability (Jodoin, 2003).

Scenario-based tasks will take students about 10–20 minutes to complete. Longer scenario-based tasks may include a greater number of embedded assessment requirements and items to which a student is asked to respond. The discussion of the balance of item types later in this chapter provides a general range to allow item developers greater flexibility to fulfill assessment design blocks.

Leveraging Existing NAEP Items to Create Scenario-Based Tasks

All of the general principles for item writing discussed in this document apply to the development of scenario-based tasks. However, the development of a well-written scenario-based task is not easy. The authors of the 2019 Trends in International Mathematics and Science Study (TIMSS) Framework stated that TIMSS problem solving and inquiry tasks (PSIs), which have characteristics similar to NAEP scenario-based tasks, are challenging and time consuming to build (Mullis & Martin, 2016). Therefore, to aid in task development for NAEP mathematics, some illustrated suggestions are offered for the development process using existing NAEP TEL specifications and existing NAEP items as starting points.

The 2014 TEL *Assessment and Item Specifications* suggests the use of a scenario shell to help think through the components of a task, including the problem to be solved and the practices and objectives being assessed. The example from the TEL *Assessment and Item Specifications* is shown in Illustration 4.3a. An adaptation for NAEP Mathematics is shown in Illustration 4.3b.

Illustration 4.3a. NAEP TEL Sample Scenario Shell

Grade	4, 8, or 12
Major Assessment Areas	Technology and Society Design and Systems Information and Communication Technology
Context	*What is the context of the scenario?*
Problem	*What are the big ideas for the students?*
Available Resources and Information	*What is given to the student to solve the problem?*
Tools Used	*What domain-specific tools (virtual and actual) will the students use?*
Practices	*Which of the NAEP practices will be addressed?*
Assessment Targets	*Which of the NAEP targets will be addressed?*

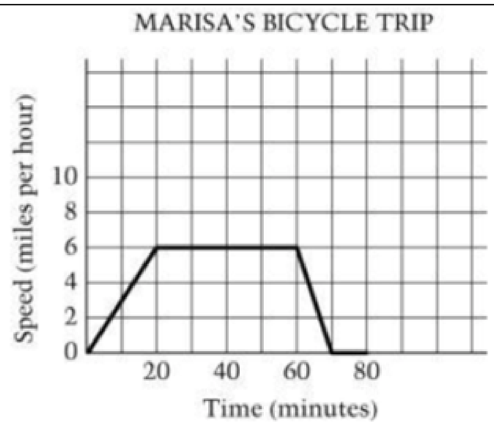
Illustration 4.3b. NAEP Mathematics Sample Scenario Shell

Grade	4, 8, or 12
Major Content Area	Number Properties and Operations Measurement Geometry Data Analysis, Statistics, and Probability Algebra
Context	*What is the context of the scenario?*
Problem	*What are the big ideas for the students?*
Available Resources and Information	*What is given to the student to solve the problem?*
Tools Used	*What domain-specific tools (virtual and actual) will the students use?*
NAEP Mathematical Practices	*Which, if any, of the NAEP mathematical practices will be measured?*
NAEP Mathematics Objectives	*Which of the NAEP content objectives will be measured?*

Bicycle Trip Example: Grade 8 Scenario-Based Task

Consider the *Bicycle Trip* item introduced in Chapter 3 and included again in Illustration 4.4 for reference.

Illustration 4.4. Marisa's Bicycle Trip



The graph above represents Marisa's riding speed throughout her 80-minute bicycle trip. Use the information in the graph to describe what could have happened on the trip, including her speed throughout the trip.

During the first 20 minutes, Marisa *[textbox]*

From 20 minutes to 60 minutes Marisa *[textbox]*

From 60 minutes to 80 minutes Marisa *[textbox]*

Using the NAEP Mathematics Sample Scenario Shell and the original *Bicycle Trip* item, an outline of a scenario-based task was developed (see Illustration 4.5). During this process, the context of the original item was revisited to consider topics of interest for eighth graders.

With the multimedia capabilities of online administration of scenario-based tasks, consideration was given to the unique opportunities for content presentation as a way to connect a version of the graphical representation from the original item to a different type of representation in the new task. The choice to use video clips as a mode of representation provides a level of engagement not offered by the original task.

Illustration 4.5. Grade 8 Scenario Shell Adaptation of *Marisa's Bicycle Trip*

Grade	8
Major Content Area	Algebra
Context	Ordering video clips of a bicycle trip
Problem	Given a graph and a set of video clips, order the clips to show Marisa's bicycle trip.
Available Resources and Information	video clips graphical representation
Tools Used	video player
NAEP Mathematical Practices	Representing
NAEP Mathematics Objectives	Algebra 2.a – Translate between different representations of linear expressions using symbols, graphs, tables, diagrams, or written descriptions.

After the scenario shell was completed, an initial draft of a portion of the task was developed. This draft is shown in Illustration 4.6. As revisions to the original graphic are likely needed and technology features will be applied, some italicized notes are included within the item to illustrate thinking about these item components. Additional item parts could be added to consider questions that can be answered about Marisa's trip or to apply the same relational thinking to a different context. As the multimedia features of the mathematics assessment are configured, this task could be developed further and continue to be refined.

To build from items in the existing item pool, a scenario-based task based on an unreleased NAEP item could include the original item as a part of the task. For example, using the *Bicycle Trip* item, the original item might be used as Part A, to have students talk about the rate at which Marisa rides. The new content of the task, the ordering of the video clips, could be included as Part B.

Illustration 4.6. Draft Grade 8 Scenario-Based Task Adaptation of *Marisa's Bicycle Trip*

You are producing a video to tell the story of Marisa's bicycle trip. You will order four video clips. To assist you, an editor has created a graph showing the relationship between the number of minutes Marisa rode her bicycle and her speed.

[*Art note: Adjust the time shown on the axis depending on whether videos are meant to be real time (e.g., a 4-minute trip) or representations of segments of a longer trip.*]

Watch each video clip. Then, put the four clips in order so that they represent the graph of Marisa's bicycle trip.

[*Technology implementation note: Create a tab for each of the four clips. Label each tab "Video Clip <letter>", with <letter> replaced with A, B, C, and D. Create a fifth tab for ordering the clips to show the trip. Label the fifth tab "Order the clips". Consider the potential to merge the ordered clips all together to show the trip in its entirety.*
Video Clip description and scoring order:
Clip A: shows Marisa riding at constant speed (order: second)
Clip B: shows Marisa stopped (order: fourth)
Clip C: shows Marisa riding at a decreasing speed (order: third)
Clip D: shows Marisa riding at an increasing speed (order: first)
Note that video clips should not give the actual speed at which Marisa is riding.]

Tab development:
The graph of Marisa's Bicycle Trip should be shown on each tab.

Text for use with Video Clips A, B, and C:
Watch the video clip. Then select the tab for the next video.
[*Include play button for the video.*]

Text for use with Video Clip D:
Watch the video clip. Then select the tab to order the video clips.
[*Include play button for the video.*]

Text for use on Order the Clips Tab:
Order the video clips so that they represent the graph of Marisa's bicycle trip. Explain the ordering of the video clips.

Drag each clip into a box.

[*Present the clips in a row: A, B, C, D. In a row beneath the clips, create four drop boxes, labeled "First", "Second", "Third", and "Fourth". Under this item part, include a response box for the explanation.*]

The draft grade 8 task requires students to synthesize multiple pieces of information to arrive at a solution. The setting of the task presents the content in a way that could not be done in a traditional item, and the motivating goal of producing a video provides an authentic context. Taken along with potential additional item parts, these features define this task as scenario based.

The focus on Algebra as a content area and Representing as a NAEP Mathematical Practice were inherent to the scenario-based *Bicycle Trip* task. However, this is not always the case. A scenario-based task may contain items aligned to different content areas and/or NAEP Mathematical Practices, with an identified overarching content area and practice defined by the task problem (i.e., the driving storyline for the task, such as in the *Stacking Chairs* task in Exhibit 4.1).

Bicycle Trip Example: Grade 4 Scenario-Based Task

The NAEP Mathematical Practice of Representing spans all grade levels. Therefore, a set of items inspired by the original NAEP *Bicycle Trip* can be developed utilizing the idea of connecting representations at each grade level, 4, 8, and 12. To this end, consider an adaptation of the *Bicycle Trip* item for grade 4. For this work, content that is **not** appropriate to assess at grade 4 was as important to consider as content that can be assessed at grade 4. For example, although grades 8 and 12 objectives address representations that show change over time, objectives at grade 4 do not.

Since the objective for the grade 8 tasks was Algebra 2.a, Algebra 2.a was initially considered as the objective for the grade 4 task. The guiding question, "How can the representation from the original item be adapted to meet the needs of a grade 4 task?" served as a starting point for the completion of the scenario shell shown in Illustration 4.7.

Illustration 4.7. Grade 4 Scenario Shell Adaptation of *Marisa's Bicycle Trip*

Grade	4
Major Content Area	Number Properties and Operations
Context	Ordering video clips of a bicycle trip
Problem	Given a representation and a set of video clips, order the distances indicated in the video clips from least to greatest.
Available Resources and Information	video clips graphical representation
Tools Used	video player
NAEP Mathematical Practices	Representing
NAEP Mathematics Objectives	Number Properties and Operations 1.i – Order or compare whole numbers, decimals, or fractions using common denominators or benchmarks.

Note that the objective changed from Algebra 2.a to Number Properties and Operations 1.i as the scenario shell was developed. This change stemmed from a desire to focus on a provided representation instead of translation between representational forms. To adapt the representation for grade 4, a diagram can be presented that shows four locations represented by images. The path Marisa rides connects the images, and each piece of the path is labeled. The video clips can show Marisa riding from one location to the next, indicating the distance between each location, with each distance measured in the same unit. Students can be asked to order the labels for the pieces of the path by distance, from least to greatest.

Bicycle Trip Example: Grade 12 Scenario-Based Task

To reimagine the task for grade 12, content not addressed at grade 8 but addressed at grade 12 was considered first. The comparable grade 12 objective, Algebra 2.a, expands the types of equations used but does not differentiate the types of interpretations students are to make. Therefore, a decision was made to increase the complexity of the video clips by including information about speed in each clip along with a set of clips that cannot be represented by any piece of the graph. The scenario shell for the grade 12 task is shown in Illustration 4.8.

Illustration 4.8. Grade 12 Scenario Shell Adaptation of *Marisa's Bicycle Trip*

Grade	12
Major Content Area	Algebra
Context	Ordering video clips of a bicycle trip
Problem	Given a graph and a set of video clips, order the clips to show Marisa's bicycle trip.
Available Resources and Information	video clips graphical representation
Tools Used	video player
NAEP Mathematical Practices	Representing
NAEP Mathematics Objectives	Algebra 2.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.

For the grade 12 task, six video clips can be presented. Two of the clips would show either a speed or an elapsed time that cannot be matched to a piece of the graph. However, each clip would be formatted similarly to provide sufficient context for students to determine speed and/or elapsed time. The item directions would ask students to watch each of the six clips, and then select and order **four** of the clips to show what is most likely Marisa's bicycle trip.

Identification and revision of a concept as a foundation for a scenario-based task is likely to happen in parallel with the identification and revision of mathematics objective targets. Concepts serving as candidates for a scenario will likely involve **at least** two actions, such as attending to relationships, visualizing, coordinating, comparing, contrasting, synthesizing, validating, predicting, or persuading via mathematical argument. For example, the original *Bicycle Trip* item involves imagining movement and coordinating between two representations (graphical and verbal). Additionally, the *Stacking Chairs* adaptation involves attending to relationships, coordinating representations (verbal, symbolic), and predicting (to identify what additional information is needed).

The item type(s) used within a scenario-based task should be based on the structure of the task and the measurement intent. The item types for a composite scenario-based task will be aligned to the item format that best supports the requested evidence. Therefore, the requirement for developing scoring guides for scenario-based tasks should follow the same principles as outlined in the discussions about item types (starting on p. 178).

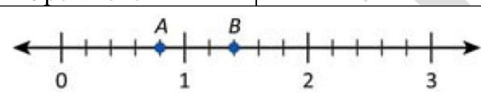
What a Scenario-Based Task Is Not

The inclusion of multiple parts is not sufficient to label an item set as a scenario-based task. One of the criteria for a task to be scenario based is that the scenario from which the task is built serves as a driving force through the completion of the task.

The item in Illustration 4.9 contains three parts. A correct response to each part requires use of the number line presented at the beginning of the item. While there is a connection between the item parts, there is no underlying storyline driving the mathematical activity required by the item as a whole.

Illustration 4.9. Nonexample of a Scenario-Based Task

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Number Properties and Operations	Representing	Num – 1.b	SCR – composite



Points *A* and *B* are plotted on the number line.

What number corresponds to point *A* ?

What number corresponds to point *B* ?

What number corresponds to the point that is located halfway between point *A* and point *B* ?

Scoring Information	
Key	a) 4/5 or 0.8 (or equivalent) b) 1 2/5 or 1.4 (or equivalent) c) 1 1/10 or 1.1 (or equivalent)

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-8M9 #12 M3566CL.

While development of scenario-based tasks is a complex and time-consuming process, focusing on the larger aspects of the task prior to development of the items that will comprise the task provides structure within which item writers can work. Additionally, the considerations listed below can be used to aid the item writer in task development.

- Use of an online environment to create authentic, relevant, and compelling ways of presenting and assessing content and practices.
- Contexts that are interesting to and appropriate for students at the grade level.
- Content and NAEP Mathematical Practices that make sense within the proposed context.
- Content that is out-of-bounds at a particular grade level as a check to ensure the task aligns to on-grade-level objectives.
- Progression of content through grade levels.
- Patience and persistence in iterating the development process and seeking feedback as the task becomes fully formed.

Item development is a complex endeavor involving many components, from conceptualization of content, to item format selection, to scoring considerations. To aid in providing structure within which these complexities can be thought through, design patterns have been conceptualized for use as item development tools (Mislevy & Haertel, 2006). Stemming from recent work in evidence-centered designs for assessment, design patterns leverage commonalties in item design so that differing components can be modified (e.g., providing structure for a set of collaborative mathematics tasks that assess different content objectives). As all scenario-based tasks have some common components and some components that vary, consideration should be given to the potential of design patterns to substantially support their development.

Item Types

Since 1992, the NAEP Mathematics Assessment has used two item types: multiple choice and constructed response. In 2017, the term multiple-choice was revised to "selected response" to account for the wider range of item formats available (e.g., matching) with digitally based assessments. Selected-response items require a student to select one or more response options from a given, limited set of choices. Constructed-response items require students provide a text-based or numerical response.

Innovative item types made possible by digital test administration are often referred to as technology-enhanced items (TEIs). TEIs have the potential to assess what students know and are able to do in a more authentic way than static selected-response items (Sireci & Zenisky, 2006). While item performance indicates that TEIs tend to be more difficult than multiple-choice items assessing the same content, both item formats appear to be well-correlated with student overall performance on an assessment (Crabtree, 2016). Therefore, TEIs are often viewed as a middle ground between traditional multiple-choice items that are frequently viewed as artificial but have high reliability, and traditional constructed-response items that allow for more authentic assessment of what students know and can do but are costly in terms of money and time spent during development, administration, and scoring and are likely to have lower reliability (Sireci & Zenisky, 2016).

Research on the development and performance of TEIs is ongoing, but what is known has guided the development of the recommendations in this chapter. As additional item-format-specific research is disseminated, assessment developers will be able to refine development and administration guidelines.

Some selected-response items, such as matching or multiple-selection items, have scoring guides to permit partial credit. Every constructed-response item has a scoring guide that defines the criteria used to evaluate students' responses. Some short constructed-response items can be scored according to guides that permit partial credit, while others are scored as either correct or incorrect. All constructed-response scoring guides are refined from work with a sample of actual student responses gathered during item pilot testing. Students are provided information on elements required for a complete response in some of the individual discrete item stems and/or in overviews of composite items. This provides all students with greater access to the task and defines the parameters for their responses, honoring their time and energy as they engage in the work.

In 2025, the NAEP Mathematics Assessment retains selected- and constructed-response item types. The evolving capabilities of digital technology and the addition of NAEP Mathematical Practices means the 2025 Framework includes the expansion of the two item types to allow for additional object-based and discourse/collaboration-based responses within discrete items and scenario-based tasks. Selected-response items for use on the NAEP Mathematics Assessment include a variety of formats. The listed formats reflect a subset of those with the potential to be developed. Any combination of these item formats in a single item constitutes a composite item.

Selected Response

- Single-selection multiple choice: Students respond by selecting a single choice from a set of given choices.
- Multiple-selection multiple choice: Students respond by selecting two or more choices that meet the condition stated in the item stem.
- Matching: Students respond by inserting (i.e., dragging and dropping) one or more source elements (e.g., a graphic) into target fields (e.g., a table).
- Zones: Students respond by selecting one or more regions on a graphic stimulus.
- Grid: Students evaluate mathematical statements or expressions with respect to certain properties. The answer is entered by selecting cells in a table in which rows typically correspond to the statements and columns to the properties checked.
- In-line choice: Students respond by selecting one option from one or more drop-down menus that may appear in various sections of an item.
- Conversational responses (new): Students respond by selecting from two or more choices of conversational responses as part of a discourse-based or collaborative item.

A new selected-response item type included for the 2025 NAEP Mathematics Assessment involves the use of discourse and collaboration responses. These types of items map most directly to the collaborative mathematics and modeling practices outlined in Chapter 3. Current examples ask a student to interact via a text-based scenario with avatars and choose (e.g., through multiple-choice, limited option selections) from given conversational responses to move the collaborative problem forward. Such a selected-response choice then provides some information about the level of collaborative mathematics the student exhibits.

Although conversational responses retain the structure of other selected-response item formats, they have the potential to be scored polytomously, meaning that some incorrect answer choices may be determined to be "more correct" than other incorrect answer choices. Therefore, response options in these items may have differing numbers of score points.

The table in Illustration 4.10 lists and describes selected-response item formats, indicates other names by which an item format might be known, and provides the location of exhibits and illustrations within the *Assessment and Item Specifications* of examples and nonexamples. At the beginning of the table are guidelines to assist with the development of selected-response items.

Illustration 4.10 Selected Response Item Information

Selected Response (SR) Development Guidelines

- The item stem (introduction to the item) includes only the information needed for students to respond.
- Response options are succinctly worded and avoid repetition of phrases in each choice.
- Response options are parallel in mathematical approach and general phrasing.
- Response options do not cue correct response or use exclusionary language (e.g., always, never).
- Incorrect Response options are plausible (e.g., through mathematical conceptions, common errors).
- Incorrect Response options connect to the mathematical construct being assessed.
- Rationales are provided for all response options.

NAEP Item Formats	**Similar Item Formats/ Abbreviations**	**Student Interaction**	**Location of Example Item(s)**
single-selection multiple choice	multiple choice (MC)	Student selects one of four given response options at grade 4. At grades 8 and 12, student selects one of five response options.	Illustration 3.1 Illustration 3.3
multiple-selection multiple choice	multiple select (MS)	Student selects two or more of the given response options.	Illustration 4.11
matching	drag and drop gap match	Student inserts one or more source elements (e.g., graphics) into target fields (e.g., cells of a table).	Illustration 3.21 Illustration 4.12
zones	hot spot (HS)	Students respond by selecting one or more regions on a graphic stimulus.	Illustration 4.13a Illustration 4.13b
grid	matching table	Students evaluate mathematical statements or expressions with respect to certain criteria. The response is entered by selecting cells in a table in which rows typically correspond to the statements and columns to the properties checked.	Illustration 3.8 Illustration 4.14
in-line choice	inline dropdown (IC)	Students respond by selecting one option from one or more drop-down menus that appear in various sections of an item.	Illustration 4.15

Discussion of Selected-Response Item Examples

Single-selection multiple choice

Multiple-choice items are an efficient way to assess knowledge and skills, and they can be developed to measure various levels of rigor. 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. Historically in NAEP, the stem is followed by either four or five response options, only one of which is correct. The item in Illustration 3.1 in Chapter 3 (p. 97) illustrates a straightforward stem with a direct question. The distractors are plausible, but only one response option is correct.

Multiple-selection multiple choice

As with single-selection multiple choice items, the stem of a well-designed multiple-selection multiple choice item 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. To avoid confusion for students, it is common in assessment development that the stem in multiple-selection items is followed by **more than** four response options with **more than** one correct response option (i.e., when single-selection items have four options with exactly one option correct). Directions for this item format should indicate either the number of correct responses or that students should select all of the correct responses. Due to the selection of multiple responses, some items allow for partial credit. For these items, scoring guides are developed to indicate how the partial credit is allocated.

Correctly responding to items using this format is more challenging than traditional multiple-choice items as students must determine not only the relationship between a response and the item stem but also the relationships among the response options (Baghaei & Dourakhshan, 2016). The item in Illustration 4.11 asks students to select all of the response options that represent a unit of measure for the length of time a person will drive. Using a multiple-selection multiple-choice item format allows for the assessment of student recognition of more than one appropriate unit, changing the measurement intent from that of an item asking students to select exactly one unit of measure.

Illustration 4.11. Selected-Response Example: Multiple-Selection Multiple-Choice Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Measurement	Other	Meas – 2.a	SR – MS

Ms. Taylor will drive from Maine to Florida.

Which of the following units of measurement could be used to measure the amount of time it will take her to complete the drive?

Select all the correct answers.

- A ☐ Days
- B ☐ Gallons
- C ☐ Hours
- D ☐ Miles
- E ☐ Pounds
- F ☐ Yards

Clear Answer

Scoring Information	
Key	A, C
Correct	Two correct selections and no incorrect selections
Partial	One correct selection and no incorrect selections
Incorrect	Two or fewer correct selections and one or more incorrect selections

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-4M1 #7 M3706MS.

Matching

Matching items take many forms, but each involves the dragging and dropping of one or more objects. For example, a matching item may require the dragging of text, numbers, or figures into indicated spaces; the ordering of presented text, numbers, or figures; or the matching of a subset of objects from one set of information to objects in another set.

Matching items can quickly become quite complicated based on the number of dragging and dropping actions required. Item writers should consider the number of actions in light of the measurement intent of the item – that is, how much information students need to provide to demonstrate evidence of understanding of the assessed objective. Additionally, when possible, the development of more objects to drag than locations in which to drop them tends to allow students to make an error in one placement without impacting the other placements.

The item in Illustration 4.12 asks students to drag each color into the correct piece of the circle graph. As each color is required to be represented in the circle graph, a one-to-one relationship between the colors and the pieces of the graph is the necessary structure.

Illustration 4.12. Selected-Response Example: Matching Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics and Probability	Representing	Data – 1.b	SR – matching

Ms. Glen asked students in her class to vote for their favorite color from a list of five colors.

She recorded the results in the table.

FAVORITE COLORS

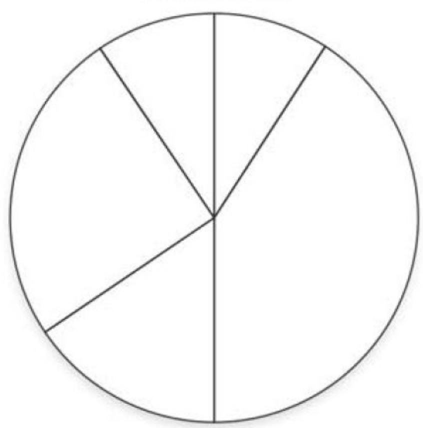
Color	Number of Votes
Blue	3
Green	13
Purple	5
Red	8
Yellow	3

Ms. Glen correctly created a circle graph (pie chart) to show the results of the vote but did not label the sectors.

Label each sector of the circle graph to show the results of the vote.

Drag colors to the circle graph to show your answer.

FAVORITE COLORS

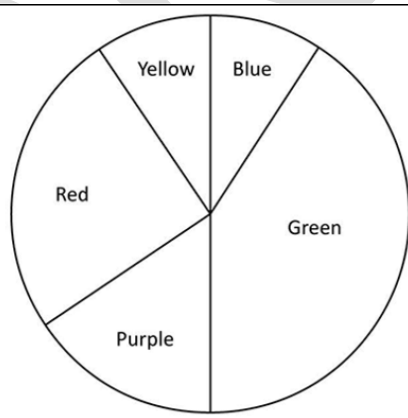


Blue
Green
Purple
Red
Yellow

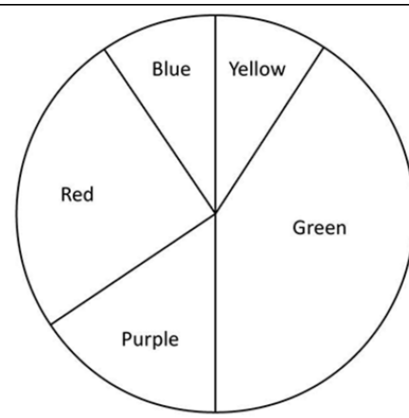
Clear Answer

Scoring Information

Key



OR



Correct	Five colors correctly placed
Incorrect	Incorrect response

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #2 M3806MS.

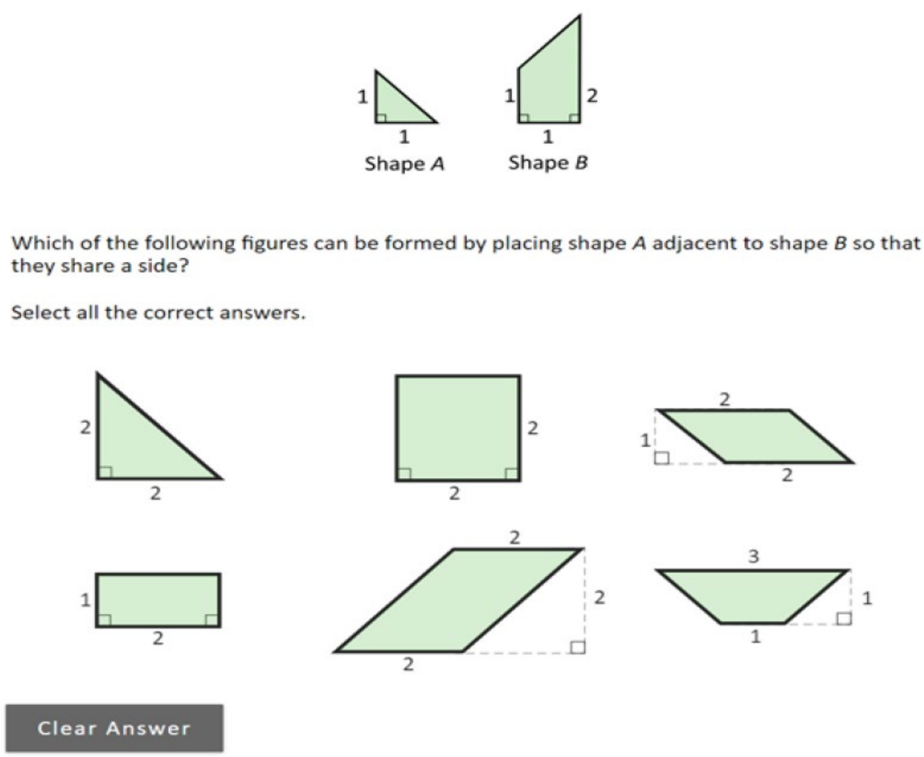
Zones

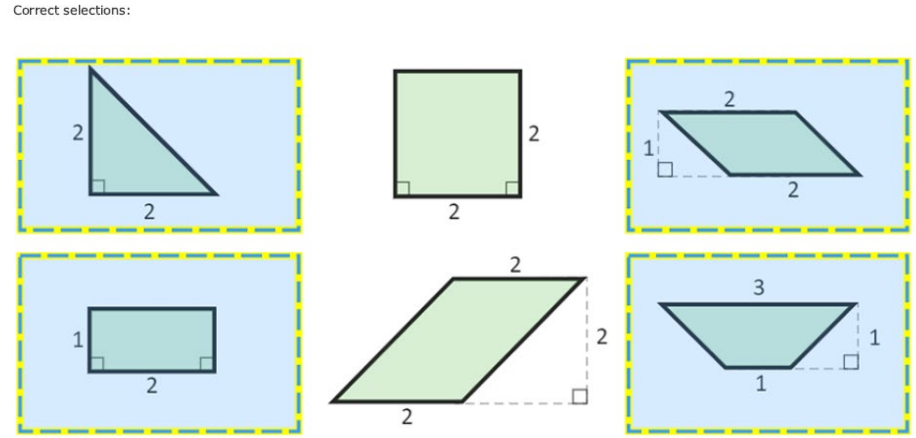
Zone items involve the selection of a graphic or graphics or the selection of a location or locations on a graphic. The zone item format can take the place of some drawing activities, such as plotting a point on a number line, encountered on some paper-and-pencil assessments. As with matching items, writers should consider the number of student actions required in light of the measurement intent of the item. When developing an item that requires the selection of graphics, consideration should be given to the number of graphics presented and the number of correct graphics. When developing an item that requires the selection of a location or locations on a graphic, consideration should be given to the size and clarity of the graphic, the number of locations that are selectable, and the number of correct locations. For zone items, the selectable locations should be purposeful and clearly defined.

The item in Illustration 4.13a presents a set of six graphics from which students choose to answer the question. Since two of the six graphics are correct, this item is comparable to a multiple-selection multiple-choice item. Note that side lengths and right-angle markings were used to clearly convey the size and shape of each figure.

Illustration 4.13a. Selected-Response Example: Zone Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Geometry	Other	Geom – 2.d	SR – zone



Scoring Information	
Key	
Correct	Four correct selections and no incorrect selections
Partial	Four correct selections and one incorrect selection OR Three correct selections and no incorrect selections
Incorrect	Four correct selections and more than one incorrect selection OR Fewer than four correct selections and one or more incorrect selections OR Fewer than three correct selections

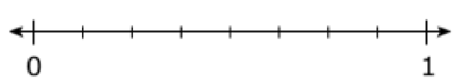
This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #12 M3814EM.

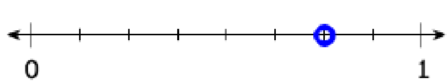
The item in Illustration 4.13b presents a number line on which students can select a point. Although information regarding the number and location of the zones is not provided, it is likely that each of the hash marks on the number line is a zone. This placement of the zones allows students to select any eighth without concern over selection of a zone between two hash marks, approximating an equivalent fraction with a denominator other than 8, or concerns over student dexterity when selecting a zone.

Illustration 4.13b. Selected-Response Example: Zone Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties & Operations	Representing	Num – 1.h	SR – zone

Select a point on the number line to plot a point that is equivalent to $\frac{3}{4}$.



Scoring Information	
Key	

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 PARCC as Item ID VF889661.

Grid

Grid items involve the selection of cells in a table to indicate a response. The rows of the table contain stimuli to be considered. The stimuli should be mathematically related. The first cell in each column of the table lists the options from which students choose. The options should be plausible for each stimulus. As with previously discussed item formats, writers should consider the number of student actions required in light of the measurement intent of the item – that is, how much information students need to provide to demonstrate evidence of understanding of the assessed objective. This should inform the number of rows and columns included in an item.

The item in Illustration 4.14 presents a set of four measurements as stimuli and two comparisons as choices. With the comparison of measurements assessed by this item, similar thinking can be applied for each stimulus. However, the nature of the stimuli chosen requires consideration for each case, as each stimulus is independent of the others.

Illustration 4.14. Selected-Response Example: Grid Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Measurement	Other	Meas – 2.b	SR – grid

2 pints = 1 quart
4 quarts = 1 gallon

There are 2 pints in a quart and there are 4 quarts in a gallon.

Which of the following measures are less than 5 quarts and which are more than 5 quarts?

Make one selection for each measure to show your answer.

Measures	Less than 5 Quarts	More than 5 Quarts
5 pints	O	O
5 gallons	O	O
1 gallon and 1 pint	O	O
1 gallon and 5 pints	O	O

Clear Answer

Scoring Information

Key

Correct selections:

Measures	Less than 5 Quarts	More than 5 Quarts
5 pints	X	
5 gallons		X
1 gallon and 1 pint	X	
1 gallon and 5 pints		X

Correct	Four correct selections
Partial	Three correct selections
Incorrect	Fewer than three correct selections

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #5 M3838MS.

In-line choice

In-line choice items require students to select text that correctly completes a statement. Typically, the item stem presents information relevant to the completion of one or more statements. The statements are written beneath the stem with drop-down menus that present plausible options for sentence completion. Item writers should take care when determining the number of options for each drop-down menu, as the total number of response options has the potential to impact the amount of reasoning required for students to complete the item. Additionally, a student taking the test with a screen reader must listen to **every** potential answer, so the number of options in each drop-down menu impacts the number of combinations that must be heard and processed.

The item in Illustration 4.15 provides information about two functions. Following the information, two statements containing drop-down menus are given. The first statement asks students to compare the slopes of the two functions. The second statement asks students to compare the *y*-intercepts of the two functions. In this example, the option that completes one statement is independent of the other.

Illustration 4.15. Selected-Response Example: In-line Choice Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Algebra	Representing	Alg – 2.b	SR – IC

Function A is a linear function that passes through the points shown in the table.

Function A

x	y
6	7
−3	4

Function B is a linear function defined by the equation $y = 2x + 1$.

Select from the drop-down menus to correctly complete each sentence.

The slope of Function A is [Choose... : greater than / less than / equal to] the slope of Function B.

The *y*-intercept of Function A is [Choose... : greater than / less than / equal to] the *y*-intercept of Function B.

Scoring Information

Key	The slope of Function A is [less than] the slope of Function B. The *y*-intercept of Function A is [greater than] the *y*-intercept of Function B.

This item was revised for the purpose of this document. The original version of this item appeared in the 2019 PARCC as Item ID VH139356.

Constructed-response items for the NAEP Mathematics Assessment also include a variety of formats, including those listed below. Any combination of constructed-response item formats or selected-response formats with at least one constructed-response format in a single item constitutes a composite constructed-response item.

Constructed Response

- Short constructed response: Students respond by giving either a numerical result or the correct name or classification for a group of mathematical objects, or possibly by writing a brief explanation for a given result.
- Extended constructed response: Students respond by giving a description of a situation, analysis of a graph or table of values or an algebraic equation, or a computation involving specific numerical values. These items require students to consider a situation that requires more than a numerical response or a short verbal communication.
- Object-based responses: Students respond by manipulating or using an object. The state of the object upon item completion is the response (see page 201 for additional details).

The table in Illustration 4.16 lists and describes constructed-response item formats, indicates other names by which an item format might be known, and provides the location of exhibits and illustrations within the *Assessment and Item Specifications* of examples and nonexamples. At the beginning of the table are guidelines to assist with the development of constructed-response items.

Illustration 4.16 Constructed-Response Item Information

<table>
<tr><td colspan="4">Constructed Response (CR)
Best used when student communication of the correct response and/or support for a response provides greater evidence than use of other item types.
Examples of item structures or response requirements for which CR items are appropriate are
• computational fluency,
• writing an equation to model a situation, and
• justifying a mathematical conjecture.</td></tr>
<tr><th>NAEP Item Formats</th><th>Abbreviations</th><th>Student Interaction</th><th>Example Item(s)</th></tr>
<tr><td>short constructed response</td><td>SCR</td><td>Students respond by entering a short text, an integer, or a decimal in a response box that consists of a single line.</td><td>Exhibit 3.7
Illustration 4.17a
Illustration 4.17b
Illustration 4.17c</td></tr>
<tr><td>extended constructed response</td><td>ECR</td><td>Students respond by entering extended text in a response box that consists of multiple lines.</td><td>Illustration 3.6
Illustration 4.18a
Illustration 4.18b</td></tr>
</table>

Discussion of Constructed-Response Item Examples

Every constructed-response item has a scoring guide that defines the criteria used to evaluate students' responses. Some short constructed-response items can be scored according to guides that permit partial credit, while others are scored as either correct or incorrect. All constructed-response scoring guides are refined from work with a sample of actual student responses gathered during pilot testing of items. Students are provided information on elements required for a complete response in some of the individual discrete item stems and/or in overviews of composite items. This provides all students with greater access to the item and defines the parameters for their response, honoring their time and energy as they engage in the work.

The type of constructed-response item, short or extended, that is written should depend on the mathematical construct that is being assessed – the content of the objective, the NAEP Mathematical Practice(s) addressed, and the rigor involved in determining and constructing a solution. 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, even before pilot use. Similarly, a mathematical justification or explanation for each rubric category description is needed. Doing so scaffolds development of a clear scoring rubric and provides guidance for those scoring the item. Item writers should refer to additional directions for developing scoring guides provided by National Assessment Governing Board policy and the assessment development contractor when constructing scoring information for an item.

Short Constructed Response

To provide more reliable and valid opportunities for extrapolating about students' approaches to problems, NAEP assessments include items often referred to as short constructed-response (SCR) items. These are short-answer items that require students to give either a numerical result or the correct name or classification for a group of mathematical objects, draw an example of a given concept, or possibly write a brief explanation for a given result. Short constructed-response items may be scored correct, incorrect, or partially correct, depending on the nature of the problem and the information gained from students' responses.

Fill-in-the-blank (FIB) items with one response box are SCR items that require students to enter a numerical or short verbal text (e.g., a name). Some FIBs are written to be scored dichotomously, that is, with two scoring categories: correct or incorrect. FIBs with two scoring categories should measure knowledge and skills in a way that multiple-choice items cannot or be designed to elicit greater evidence of students' understanding. Such FIBs might be appropriate for measuring 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 or when there are different ways to display an answer. Item writers should take care that FIB items would not be better or more efficiently structured as multiple-choice items; there should be a purpose for the use of the item type based on the measurement intent of the item.

Item writers should draft a scoring rubric for each FIB. A writer will not necessarily need to determine the scoring categories for an item, as this depends on the robustness of the item as determined in an iterative item development process.

For dichotomous items, the rubrics should define the following two categories: Correct and Incorrect**.** The item in Illustration 4.17a requires students to perform a calculation. Since this item assesses computational skills, the use of the FIB format is appropriate. The scoring information provided defines a correct result, indicating what is required for a correct response and an incorrect response.

Illustration 4.17a. Short Constructed-Response Example: Fill-in-the-Blank Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Other	Num – 3.c	SCR – FIB

Divide.

228 ÷ 4 = ☐

Scoring Information	
Key	57
Correct	Answer of 57
Incorrect	Incorrect response

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-4M1 #4 M367801.

Some FIBs are written to be scored on a three-category scale. These items should measure knowledge and skills that require students to go beyond giving a viable answer, allowing 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.

For items with three score categories, the rubrics should define the following categories: Correct, Partial, and Incorrect. The item in Illustration 4.17b is a fill-in-the-blank item that asks students to complete the cells of a table. The use of the FIB format allows this item to occupy less space than it would have had students been required to select one of four tables presented as response options. This item was developed with three score categories. A correct response requires that all of the cells be completed correctly, and a partial score is presented for an answer that demonstrates some understanding of how to extend the relationship given.

Illustration 4.17b. Short Constructed-Response Example: Fill in Multiple Cells in a Table

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Algebra	Other	Alg – 1.a	SCR – FIB

Melissa saves money for six weeks to buy a sweater.

She records her weekly savings.

She saves $2.50 the first week.

Each week, she saves $1.25 more than she saved the previous week.

Complete the table to show how much Melissa saves each week.

MELISSA'S SAVINGS BY WEEK

Week	Money Saved ($)
1	2.50
2	
3	
4	
5	
6	

<table>
<tr><th colspan="2">Scoring Information</th></tr>
<tr><td>Key</td><td>
<table>
<tr><th colspan="2">MELISSA'S SAVINGS BY WEEK</th></tr>
<tr><th>Week</th><th>Money Saved ($)</th></tr>
<tr><td>1</td><td>2.50</td></tr>
<tr><td>2</td><td>3.75</td></tr>
<tr><td>3</td><td>5.00</td></tr>
<tr><td>4</td><td>6.25</td></tr>
<tr><td>5</td><td>7.50</td></tr>
<tr><td>6</td><td>8.75</td></tr>
</table>
Note: Accept equivalent values</td></tr>
<tr><td>Correct</td><td>Correct response</td></tr>
<tr><td>Partial</td><td>4 of 5 terms correct
OR
Rule applied correctly to all but one term
OR
Response shows a correct cumulative total for each week</td></tr>
<tr><td>Incorrect</td><td>Incorrect response</td></tr>
</table>

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item NAEP Item ID 2017-8M9 #7 M3553E1.

SCR items require students to enter more than one or two words into a provided answer block (e.g., a brief explanation for a given result). The item in Illustration 4.17c was previously introduced in Chapter 2 (Illustration 2.3, p. 25). This item is being presented again here to include scoring information. Note that, similarly to the item in Illustration 4.17b, this item was developed with three score categories.

Illustration 4.17c. Short Constructed-Response Example

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Other	Data – 4.e	SCR – composite

Al, Bev, and Carmen are going on a ride at the park. Only 2 people can go on the ride at a time. They can pair up 3 different ways, as shown below.

Al and Bev

Al and Carmen

Bev and Carmen

Derek decides to join the group. How many different ways can the 4 students pair up?

Answer: ___________________

Show your work or explain how you got your answer.

Scoring Information	
Key	6 ways: Al and Bev Al and Carmen Al and Derek Bev and Carmen Bev and Derek Carmen and Derek The supporting work or explanation should show or explain how the pairings of people were obtained; this may include drawings only, words only, or a combination of both.
Correct	Correct response 6 different ways with justification that demonstrates how the four people would be paired. It is possible to justify the answer of 6 without explicitly stating the 6 pairs by name, but the justification needs to be clear.
Partial	Partially correct response Response contains the 6 different ways, but the justification is either missing or is partially correct or partially complete. The partial justification may demonstrate that Derek can be paired with more than just one of the remaining people, but the justification falls short of complete, as long as the work shown does not demonstrate that 6 was obtained via invalid reasoning, should also be placed here. OR Response does NOT obtain 6 ways but does demonstrate in some way that Derek can be paired with more than just one of the remaining people.
Incorrect	Incorrect response

This item was revised for the purpose of this document. The original version of this item appeared in the 2013 NAEP Mathematics Assessment with NAEP Item ID 2013-4M6 #14 M136901.

Extended Constructed Response

Extended constructed-response items require a greater amount of mathematical rigor than short constructed-response items. In general, extended constructed-response items ask students to solve a problem by applying and integrating mathematical concepts and require students to analyze a mathematical situation and explain a concept, or both. These 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. Extended constructed-response items often have five scoring categories: Extended, Satisfactory, Partial, Minimal, and Incorrect.

The items in Illustrations 4.18a and 4.18b are extended constructed-response items. The item in Illustration 4.18a asks students to read and interpret two graphical representations of the same data. The item consists of two parts: a multiple-selection multiple-choice item part and a fill-in-the-blank item part. The scoring rubric for this item consists of five scoring categories. For Extended credit, a complete and correct response must be provided for both item parts, while Satisfactory credit allows for a minor error. Responses scored as Partial, Minimal, and Incorrect show decreasing levels of correctness.

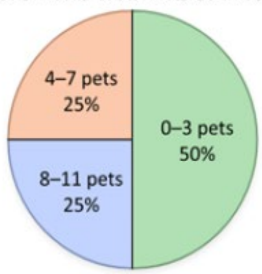
Illustration 4.18a. Extended Constructed-Response Example: MS and FIB Item Parts

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Representing	Data – 1.c	ECR – composite

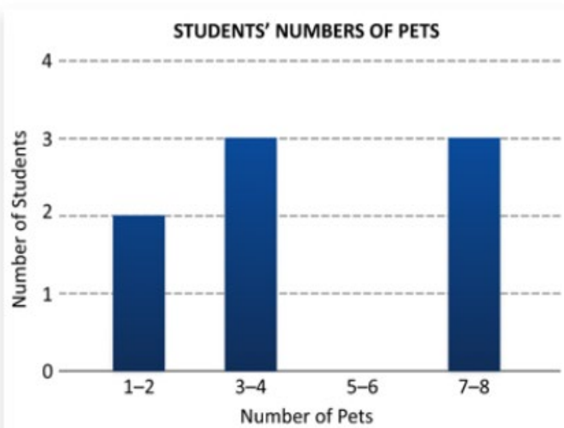
A teacher surveyed 8 students in a class to find out how many pets they each have.

Boyd and Jenny represented the results of the survey in two different ways.

Boyd made a circle graph (pie chart) to show the results.



Jenny made a bar graph to show the results.



Based on the circle graph, which of the following sets of data could be the results of the survey?

Select all the correct answers.

- A ☐ 0, 1, 2, 2, 3, 3, 5, 10
- B ☐ 0, 1, 2, 3, 5, 6, 9, 11
- C ☐ 0, 2, 2, 3, 4, 8, 9, 10
- D ☐ 1, 1, 1, 1, 4, 4, 8, 8
- E ☐ 1, 2, 4, 5, 6, 7, 8, 10

Clear Answer

Based on both Boyd's circle graph and Jenny's bar graph, what could be the results of the survey?

Enter a number in each box to show how many pets each student could have.

☐ , ☐ , ☐ , ☐ ,
☐ , ☐ , ☐ , ☐

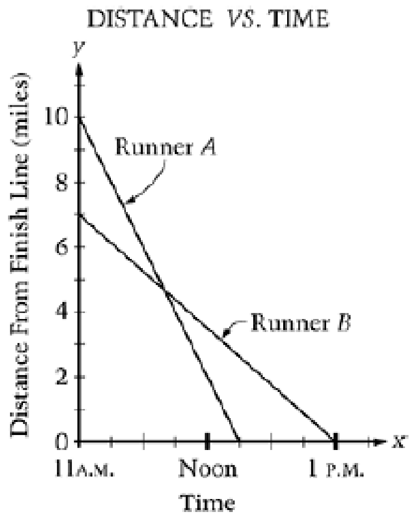
Scoring Information	
Key	(a) Correct selections: B. 0, 1, 2, 3, 5, 6, 9, 11 D. 1, 1, 1, 1, 4, 4, 8, 8 (b) Answer: 1, 1, 3, 3, 4, 7, 8, 8 OR 1, 2, 3, 3, 4, 7, 8, 8 OR 2, 2, 3, 3, 4, 7, 8, 8
Extended	Two correct selections and no incorrect selections for part (a) with a correct data set for part (b)
Satisfactory	Two correct selections and one incorrect selection for part (a) with a correct data set for part (b) OR One correct selection and no incorrect selections for part (a) with a correct data set for part (b)
Partial	Two correct selections and no incorrect selections for part (a) with an incorrect data set for part (b) OR One correct selection and one or more incorrect selections OR Two correct selections and more than one incorrect selection for part (a) with a correct data set for part (b)
Minimal	Two correct selections and one incorrect selection for part (a) with an incorrect data set for part (b) OR One incorrect selection and no incorrect selections for part (a) with an incorrect data set for part (b)
Incorrect	One correct selection and one or more incorrect selections OR Two correct selections and more than one incorrect selection for part (a) with an incorrect data set for part (b)

This item was revised for the purpose of this document. The original version of this item appeared in the 2017 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #13 M3859CL.

The item in Illustration 4.18b asks students to interpret three characteristics of a graph. Unlike the example above, scoring for this item is by characteristic. That is, there are three score categories for each characteristic. Since the item requires words and numbers for a complete response, partial credit addresses scoring for a response that includes only words or only numbers.

Illustration 4.18b. Extended Constructed-Response Example: Extended Text, Multi-Response

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Algebra	Representing	Alg – 4.d	ECR



The graph above shows distance *versus* time for a race between runners *A* and *B* . The race is already in progress, and the graph shows only the portion of the race that occurred after 11 A.M.

The table on the next page lists several characteristics of the graph. Interpret these characteristics in terms of what happened during this portion of the race. Include times and distances to support your interpretation. (A sample interpretation of the *y* -intercepts is given in the table.)

Characteristic of Graph	Interpretation in Terms of the Race
y-intercepts	At 11 A.M. Runner *A* is 10 miles from the finish line and Runner *B* is 7 miles from the finish line.
Slopes	
Point of intersection	
x-intercepts	

<table>
<tr><th colspan="2">Scoring Information</th></tr>
<tr><td>Key</td><td>There are three components in this response.
<table>
<tr><th>Characteristic of Graph</th><th>Interpretation in Terms of the Race</th></tr>
<tr><td>y-intercepts</td><td>At 11 A.M. Runner A is 10 miles from the finish line and Runner B is 7 miles from the finish line.</td></tr>
<tr><td>(a) Slopes</td><td>Runner A's speed is 8 mph and Runner B's speed is 3.5 mph.</td></tr>
<tr><td>(b) Point of intersection</td><td>Runner A and Runner B are both $4\frac{2}{3}$ miles from the finish line at 11:40 A.M.
NOTES:
• Accept distances from 4.5 to 5 miles, inclusive and times from 11:30 to 11:45, inclusive.
• This can also be interpreted as the time and distance when Runner A overtakes Runner B.</td></tr>
<tr><td>(c) x-intercepts</td><td>Runner A finishes the race at 12:15 P.M. and Runner B finishes the race at 1 P.M.</td></tr>
</table></td></tr>
<tr><td>Part A Correct</td><td>Acceptable response</td></tr>
<tr><td>Part A Partial</td><td>Acceptable interpretation without numerical values for slopes
OR
Numerical values without acceptable interpretation for slopes
Examples:
<table>
<tr><th>Characteristic of Graph</th><th>Interpretation Without Values</th><th>Values Only for Interpretation</th></tr>
<tr><td>Slopes</td><td>The speed of each runner.</td><td>A = 8 mph
B = 3.5 mph</td></tr>
</table></td></tr>
<tr><td>Part A Incorrect</td><td>Incorrect response</td></tr>
<tr><td>Part B Correct</td><td>Acceptable response</td></tr>
<tr><td>Part B Partial</td><td>Acceptable interpretation without numerical values for point of intersection
OR
Numerical values without acceptable interpretation for point of intersection
Examples:
<table>
<tr><th>Characteristic of Graph</th><th>Interpretation Without Values</th><th>Values Only for Interpretation</th></tr>
<tr><td>Point of intersection</td><td>The point where or when Runner A overtakes Runner B.
OR
The time when both runners are at the same distance from the finish line.</td><td>11:40AM
$4\frac{2}{3}$ miles</td></tr>
</table></td></tr>
<tr><td>Part B Incorrect</td><td>Incorrect response</td></tr>
<tr><td>Part C Correct</td><td>Acceptable response</td></tr>
<tr><td>Part C Partial</td><td>Acceptable interpretation without numerical values for x-intercepts
OR
Numerical values without acceptable interpretation for x-intercepts
Examples:
<table>
<tr><th>Characteristic of Graph</th><th>Interpretation Without Values</th><th>Values Only for Interpretation</th></tr>
<tr><td>x-intercepts</td><td>The end of the race.
OR
The time that each runner finishes the race.</td><td>A = 12:15 PM
B = 1:00 PM</td></tr>
</table></td></tr>
<tr><td>Part C Incorrect</td><td>Incorrect response</td></tr>
</table>

These items were revised for the purpose of this document. The original versions of these items appeared in the 2009 NAEP Mathematics Assessment with NAEP Item IDs 2009-12M2 #9 M1809CL, M180901, M180902, M180903.

Object-Based Responses

The digitally based NAEP Mathematics Assessment already incorporates use of virtual tools in tool-based responses (e.g., on-screen rulers). A new item type for NAEP Mathematics Assessments in and beyond 2025 is object-based responses. There is a growing ability to capture how students use manipulatives, both digital on screen and with “smart” physical objects off screen that can monitor activity and be connected to the digital assessment. Here there are at least two opportunities to be forward-thinking. First, further inquiry is warranted into ways to incorporate physical manipulatives that can collect data mapped to assessed constructs. The advances in smart tool technology are particularly suited to directly capture the NAEP Mathematical Practices outlined in Chapter 3. Second, further work is needed to align the data collected from tasks to valid measures of a construct. For example, one could imagine students manipulating a digital or physical object, and the solution states that they come up with at different points in time (since activity is monitored continuously) could provide strong differentiating information about mathematical modeling. A solution state of the physical orientation of an object would be the answer (versus a discrete selection or clicking a multiple-choice option). These – and other opportunities – will help NAEP move toward the ultimate goal of using tasks in the assessment in ways that capture the variety of ways students know and do mathematics.

As noted previously, the state of the object defines an object-based response. To collect evidence about the content being assessed by an item involving an object, the response provided by the state of the object must indicate enactment of the mathematics in the content objective. For example, consider an item that aimed to assess angle measurement where students had a physical protractor. A response indicated by the protractor aligned correctly to measure an angle would not provide sufficient evidence that the student can read the protractor to determine the angle measurement. Therefore, this would not be an object-based response item (though, if the protractor were virtual, the item could be a digital tool-based response item such as some items currently used on the assessment). In contrast, an item that asks students to represent the number 126 with base 10 blocks, where students manipulate physical “smart” base ten blocks, would collect evidence that the student can represent a number in base 10. The submitted state of the base ten blocks would be an object-based response. Potential objects for use on future NAEP Mathematics assessments, should they be developed as smart objects, are blocks or tiles for representing bases other than base 10, fraction strips or bars, integer chips, and algebra tiles. Additional smart objects might also be considered as the technology of the assessment evolves.

With the rapid advances in natural language processing, in the future there may be potential for mathematical collaboration to be assessed more effectively in open-ended constructed response formats. For example, the assessment might ask for and then automatically code responses where students are asked to explain their thinking or justify a contribution to collaborative mathematics. While not available at the time of the 2025 Framework revision, such technology may become available for future administrations of the NAEP Mathematics Assessment and may increase accessibility. The assessment might ask students to input their thinking or dialogue via voice (with automatic transcription into text for coding and analysis), which would dramatically open up ways for students to demonstrate what they know and can do. Similarly, pairs of students might be asked to turn on an audio documentation (e.g., a recording device) as they work together on a modeling task. The record of discourse would be part of assessment response, measurable evidence of students creating representations, making conjectures, critiquing and debating, revoicing, or justifying their solutions to one another. Considerable research and development work is needed around the technology for natural language processing and related domains, combined with careful mapping to constructs and measurement needs, to realize the aspirational goal of opening up such ways for students to show what they can do mathematically. Also, special attention must be paid to issues of consent and privacy when considering voice recording.

Additional Scoring Guide Development Information

NAEP scoring guides will be developed in accordance with recommended practice and the Governing Board Item Development and Review Policy (2002). See the Board's policy for principles about scoring guides that apply to all NAEP assessments.

Composite Items

Composite items are composed of two or more item parts. Any item format, selected response or constructed response, can be used in a composite item. Some examples of composite items from this chapter are located in Illustration 4.9, p. 177, which utilizes the fill-in-the-blank item format in each of the three parts; and Illustration 4.18a, p. 195, which utilizes multiple-selection multiple-choice and fill-in-the-blank item formats.

Response Data and Process Data for Future NAEP Mathematics Assessments

A key challenge is the need to capture enough information about mathematics content and practices for a reliable and valid assessment. When this happens, within the context of scenario-based tasks which require more time for engagement and completion, data may be available from fewer items per student.

An opportunity for future NAEP Mathematics Assessments is to develop validated measures from process data, which is generated based on student interaction with the tools and systems in the scenario-based tasks (e.g., clickstreams or activity logs). The data are different from what might be generated in a non-digital format, so it is necessary to describe how the additional data might be handled.

Conventional items always involve the student in a direct response, which generates *response data*. For example, after being presented with information in a table, the student is asked a text-based question and given a limited set of choices from which to select an answer. Student direct responses can also be used in scenarios. Direct response data can include selection from a set of choices (e.g., multiple choice, checking all boxes that apply, or providing a constructed response). Scoring methods for such response data are well established.

By contrast, *process data* reflects interactions in which the student engages in and may provide relevant evidence about whether the student possesses a skill that is an assessment target. Thus, process data can be captured, measured, and interpreted to generate a score. Clickstream data, activity logs, text, and transcribed voice responses are among the ways to capture the state of student activity as they work through a problem. These types of data hold potential power to measure student interactivity in modeling and collaborative mathematics, as well as levels of any mathematical practice (e.g., capturing frequency, density, and intensity of engagement with a mathematical practice or identifying and comparing novice to expert levels of a practice through process data). While this capability is powerful in theory, moving from big data sources to carefully constructed and validated measures is difficult to achieve in practice. A special study in the area of mathematics assessment is needed to explore and fully realize the potential of process data within digital scenario-based tasks.

NAEP Mathematics Tools

The preceding sections provide an overview for thinking through – and developing – diverse ways to show what students know and can do mathematically. Each response type requires related system tools and, at times, mathematics tools. In a digitally based environment, for example, students will require tools to enter mathematical expressions, to draw, highlight, and erase on the screen, to measure the length of virtual objects, to plot points on number lines or in coordinate planes, to graph lines and functions, and to create and modify graphical representations. Additionally, the testing environment will need to provide computational tools equivalent to a four-function calculator at grade 4, a scientific calculator at grade 8, and a graphing calculator at grade 12. Continuing a practice that began with the 2017 NAEP Mathematics Assessment, before the assessment, students complete a brief interactive tutorial designed to orient them to the relevant mathematics tools they will use during the assessment. The 2019 tutorials for each grade level mathematics can be found on the Internet at the links below.

English: https://enaep-public.naepims.org/2019/english.html
Spanish: https://enaep-public.naepims.org/2019/spanish.html

The digitally based environment of the 2025 NAEP Mathematics Assessment provides the majority of these mathematics tools digitally. All digital NAEP assessments include system tools, which are always available and common across all NAEP assessments. There are also mathematics tools, which are specific to and only available for certain items on NAEP Mathematics Assessments. The materials and accompanying tasks need to be carefully chosen to cause minimal disruption of the administration process, and typically only provided when relevant to solving the item. Continuing the calculator policy established for the 2017 digital administration, students will have access to a calculator emulator in blocks of items designated as "calculator blocks." New in 2025 will be the availability of a graphing emulator for grade 12, since high school students typically use graphing calculators or online emulators and not scientific calculators (Crowe & Ma, 2010).

Calculators

Calculator use has been recommended or mandated in high school mathematics in every U.S. state for more than 20 years and research has explored the social, personal, civic, and economic consequences of such policies for nearly as long (see, e.g., Coiro, Knobel, Lankshear, & Leu, 2008; Voogt, & Knezek, 2008). To date, most surveys of students and schools ask about types of calculators used, not about types of emulators or digital graphing environments. There is not yet a national data source on student access to graphing emulators. However, prevalence of use is indicated by the increasing use of textbooks at the high school level that include graphing emulator-embedded items in online homework problem sets and by the inclusion of graphing emulator items on state and multi-state consortia assessments (examples include the TI graphing calculator emulator on PARCC and Desmos software on SBAC).

New for the 2025 NAEP Mathematics Assessment at grade 12, "calculator" refers also to the use of a digital emulator for graphing such as can be found on most state assessments. The assessment developer will propose additional 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.

Allowance of a calculator during assessment administration should be taken into consideration when developing an item so that the presence or absence of a calculator does not interfere with the measurement intent. For example, items assessing computational fluency should not allow for use of a calculator as a calculator computation does not provide evidence of student computational skill (see Illustration 4.17a, p. 191). In contrast, allowing for the use of a calculator when solving a multi-step item in context can improve the reliability of the evidence of student knowledge and skills associated with the intended construct and avoid unintended assessment of a computational skill (see Exhibit 3.18, p. 125).

On-screen Math Keyboard

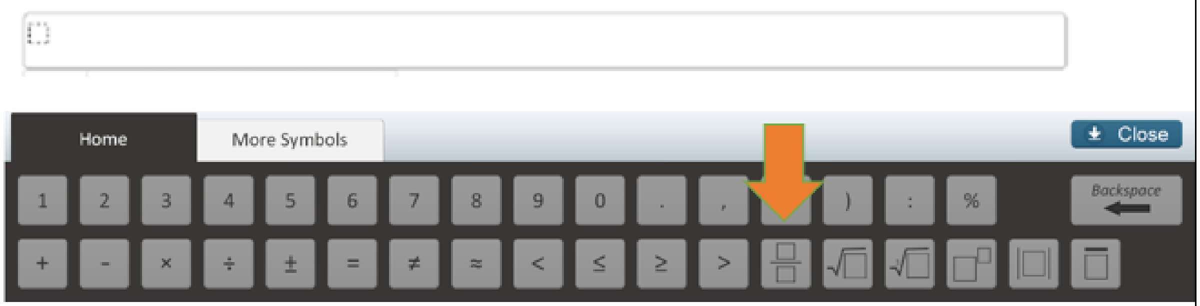
The item in Illustration 4.19 asks students to determine a probability and write their response as a fraction. The need to write the answer as a fraction allows for the use of the NAEP on-screen math keyboard, which has a built-in functionality that allows students to choose a fraction shell and enter the numerators and denominators into response boxes within the fraction shell. FIB items that require a fractional answer or for which a common mathematical error could lead to a fractional answer should allow for use of the math keyboard so that the determined answer can be entered without indicating the number type for the correct response. The on-screen math keyboard available at each grade contains symbols appropriate for that grade, so not all symbols available at one grade are available at another. However, the fraction shell is located on the on-screen math keyboard at all three grade levels.

Illustration 4.19. Short Constructed-Response Example: Use of On-screen Math Keyboard

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Other	Data – 4.d	SCR – FIB

A standard number cube, numbered 1 through 6 on each side, is rolled three times. What is the probability of rolling a 2 on all three rolls? Express your answer as a fraction.

Enter your answer as a fraction in the space provided.



Scoring Information	
Key	1/216 or equivalent

This item was revised for the purpose of this document. The original version of this item appeared in the 2016 PARCC as Item ID M20834.

Future Digital Tools

Examples of future digital mathematics tools for the 2025 NAEP Mathematics Assessment may include number tiles, spreadsheets, symbolic algebra manipulators, graphing tools, simulations, and dynamic geometry software. Continued development of mathematics tools (digital, physical, and other) can serve to achieve the goals of more authentic tasks for students and more diverse ways for students to demonstrate their knowledge and skills. Tools can allow for formal mathematics representations and symbols, and they can also allow students to create and share their own ways of thinking with their own representations. For example, some statistical tools allow students to construct their own graphical representations of data and create their own probability simulators. Considering what tools are needed for new items and the time it will take students to use them, is an integral part of the assessment design process.

Attention to Universal Design

The mathematics assessments should be developed to allow for the participation of the widest possible range of students, so that interpretation of scores leads 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.

To this end, item writing should follow the principles of universal design and sound testing practices as recommended by the National Center on Educational Outcomes (Thompson, Johnstone, Anderson, & Miller, 2005). These include attention to the population being assessed, precise definition of the constructs being assessed, review for fairness and accessibility of item content, clarity of the language and graphics used throughout the assessments, and the provision of accommodations without changing the constructs being assessed.

Although application of universal design principles to the item development process considers the ways in which the population being assessed can demonstrate learning, the use of such principles does not remove the need for accommodations altogether. With this in mind, items should be written to allow for necessary accommodations, including the use of online tools available to students during test administration, without changing the constructs being assessed, providing a necessary feature for valid and reliable assessments.

Accessibility

The NAEP Mathematics Assessment is designed to measure student achievement across the nation. Consequently, NAEP incorporates inclusive policies and practices into every aspect of the assessment, including selection of students, participation in the assessment administration, and valid and effective accommodations. NAEP is administered to a sample of students who represent the student population of the nation, regardless of race/ethnicity, socioeconomic status, disability, status as an English language learner, or any other factors. Similarly, for state-level results and results for the NAEP Trial Urban District Assessment, NAEP is administered to a sample of students who represent the jurisdiction. Therefore, the NAEP Mathematics Assessment provides an opportunity for participating students to demonstrate mathematical knowledge and skill, including students who have learned mathematics in a variety of ways, followed different curricula and used different instructional materials; students who have mastered mathematics content and practices to varying degrees; students with a variety of disabilities; and students who are English language learners. The related design issue is the development of a large-scale assessment that measures mathematics achievement of students who come to the assessment with different experiences, strengths, and challenges; who approach mathematics from different perspectives; and who have different ways of displaying their knowledge and skill.

NAEP uses two methods to design an accessible assessment program that provides accommodations for students with special needs. The first is addressed by careful item and delivery design with the full consideration of the range of participating students. For many students with disabilities and students whose native language is not English, the standard administration of the NAEP assessment will be most appropriate. For other students with disabilities (SD) and some English language learners (ELL), NAEP allows for a variety of accommodations, which can be used alone or in combination.

Some accommodations are built-in features, called Universal Design Elements, of the NAEP system tools that are available to all students. Other accommodations, such as additional assessment time, are offered for specific eligible students. Available accommodations fall into four categories:

- Standard NAEP Practice, available in almost all NAEP assessments for SD and ELL students.
- Other accommodations for SD students that require special presentation, such as Braille or sign language.
- Other accommodations for ELL students.

- Universal Design Elements that are built-in features of the computer-based assessments available to all students.

For more information about accommodations, see the Governing Board's (2014a) Policy on NAEP Testing and Reporting of Students with Disabilities and English Language Learners at https://www.nagb.gov/content/nagb/assets/documents/policies/naep_testandreport_studentswithdisabilities.pdf.

Matrix Sampling

The design of NAEP uses matrix sampling to enable a broad and deep assessment of students' mathematical knowledge and skill that also minimizes the time burden on schools and students. Matrix sampling is a sampling plan in which different samples of students take different samples of items. Students taking part in the assessment do not all receive the same items. Matrix sampling greatly increases the capacity to obtain information across a much broader range of the objectives than would otherwise be possible.

Balance of the Assessment

As mentioned earlier, the goal is to create an authentic assessment, one based on the experiences of students that will diversify the ways that students can show what they know and can do in mathematics. This vision for the 2025 NAEP Mathematics Assessment requires a significant change from the 2017 NAEP Mathematics Assessment. Specifically, scenario-based tasks require more time than discrete items. Likewise, the emphasis placed on NAEP Mathematical Practices in the Framework increases interdependence since multiple practices may be assessed simultaneously in the context of one item. The expansion of item types to include scenario-based tasks also complicates the assessment design.

Having introduced the balance of content and practices in Chapters 2 and 3 respectively, here is a summary of all three balance dimensions.

- Balance by Mathematics Content
    - Number Properties and Operations
    - Measurement
    - Geometry
    - Data Analysis, Statistics, and Probability
    - Algebra
- Balance by Mathematical Practice
    - Representing
    - Abstracting and Generalizing
    - Justifying and Proving
    - Mathematical Modeling
    - Collaborative Mathematics
- Balance by Response Type
    - Selected response
    - Constructed response (short and extended)

Balance of Mathematics Content

Each NAEP Mathematics Assessment item or item part is developed to measure one content objective. Exhibit 4.2 has the distribution of items by grade and content area. See Chapter 2 for further details.

Exhibit 4.2. Approximate 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>45*</td><td>20</td><td>10</td></tr>
<tr><td>Measurement</td><td>20</td><td>10</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>5</td><td>20</td><td>25</td></tr>
<tr><td>Algebra</td><td>15</td><td>30</td><td>35</td></tr>
</table>

* Note: Increased attention to assessing fraction content – at least one-third of grade 4 Number Properties and Operations items should assess fraction content.

Balance of Mathematical Practices

The target percentage range of items for each NAEP Mathematical Practice is given in Exhibit 4.3. Most NAEP Mathematics Assessment items will feature one of the five NAEP Mathematical Practices (55 to 85 percent). The balance of items (15 to 45 percent), those in the "Other" category, will assess knowledge of content without calling on a particular NAEP Mathematical Practice. Because of the matrix sampling used on the NAEP Mathematics Assessment, the proportions in Exhibit 4.3 are for the entire pool of items used and do not represent the experience of each student. See Chapter 3 for further details about the NAEP Mathematical Practices.

Exhibit 4.3. Approximate Percentage Distribution of Items by NAEP Mathematical Practice

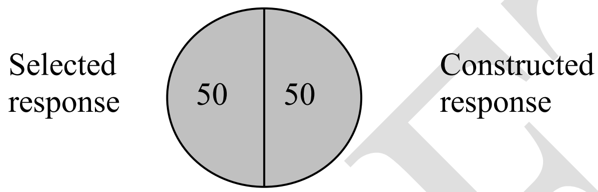
NAEP Mathematical Practice Area	Percentage of Items
Representing	10–15
Abstracting and Generalizing	10–15
Justifying and Proving	15–25
Mathematical Modeling	10–15
Collaborative Mathematics	10–15
Other	15–45

Certain formats are likely to be especially valuable in eliciting particular NAEP Mathematical Practices. As illustrated in Chapter 3, discrete items are useful measures of NAEP Mathematical Practices such as Representing, Abstracting and Generalizing, and Justifying and Proving. Also, as noted in Chapter 3, Mathematical Modeling and Collaborative Mathematics are more appropriately measured by scenario-based tasks.

Balance by Response Type

Items include selected-response and constructed-response types, and these response types may also occur within scenario-based tasks. Selected response includes traditional single-selection multiple choice, as well as other selected-response types such as matching, zones, in-line choice, grid, and discourse limited-option responses. Constructed response includes short and extended constructed response. Types of constructed-response items may include item types such as fill in the blank, extended text, digital tool-based, and object-based constructed responses, as well as discourse and collaboration responses. Testing time on NAEP is divided evenly between selected-response items and constructed-response items as shown in Exhibit 4.4.

Exhibit 4.4. Percent of Testing Time by Response Type



REPORTING RESULTS OF THE NAEP MATHEMATICS ASSESSMENT

NAEP provides the nation with a snapshot of what U.S. students know and can do in mathematics. Results of the NAEP Mathematics Assessment administrations are reported in terms of average scores for groups of students on the NAEP 0–500 scale and as percentages of students who attain each of the three achievement levels (*NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*). This is an assessment of overall achievement, not a tool for diagnosing the needs of individuals or groups of students. Reported scores are always at the aggregate level; by law, scores are not produced for individual schools or students. Results are reported for the nation as a whole, for regions of the nation, for states, and for large districts that volunteer to participate in the NAEP Trial Urban District Assessment (TUDA). The NAEP results are published in an interactive version online as *The Nation's Report Card* (National Assessment Governing Board, n.d.). The online resource provides detailed information on the nature of the assessment, the demographics of the students who participate, the assessment results, and the contexts in which students are learning.

Legislative Provisions for NAEP Reporting

Under the provisions of the Every Student Succeeds Act (ESSA), states receiving Title I grants must include assurance in their state plans that they will participate in the reading and mathematics state NAEP at grades 4 and 8. Local districts that receive Title I funds must agree to participate in biennial NAEP reading and mathematics administrations at grades 4 and 8 if they are selected to do so as part of the NAEP sample. Their results are included in state and national reporting. Participation in NAEP will not substitute for the mandated state-level assessments in reading and mathematics at grades 3 to 8. An important development over the last 20 years has been an evolving understanding of how NAEP complements state assessments, which are tightly aligned with state standards.

In 2002, NAEP initiated TUDA in five large urban school districts that are members of the Council of the Great City Schools (the Atlanta City, City of Chicago, Houston Independent, Los Angeles Unified, and New York City Public Schools districts). In 2003, additional large urban districts began to participate in these assessments, growing to a total of 27 districts by 2017. TUDA is administered biennially in odd-numbered years in tandem with NAEP state-level assessments. Sampled students in TUDA districts are assessed in the same subjects and use the same NAEP field materials as students selected as part of national main or state samples. TUDA results are reported separately from the state in which the TUDA is located, but results are not reported for individual students or schools. With student performance results reported by district, participating TUDA districts can use results for evaluating their achievement trends and for comparative purposes. Here too the complementarity of NAEP with state and local assessments is important to support so as to avoid unnecessary additional testing and to maximize useful information for educators and policymakers to use.

Reporting Scale Scores and Achievement Levels

The NAEP Mathematics Assessment is reported in terms of percentages of students who attain each of the three achievement levels—*NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*. Reported scores are always at the aggregate level. The Framework calls for NAEP results to continue to be reported in terms of sub-scores as well, for each content domain. Cut scores represent the minimum score required for performance at each NAEP achievement level. Cut scores are reported along with the percentage of students who scored at or above the cut score.

The Framework calls for reporting on NAEP Mathematical Practices. Since these practices are fundamentally intertwined with NAEP mathematics content areas, there will not be separate reporting scales for each NAEP Mathematical Practice. Options for measuring and reporting on NAEP Mathematical Practices are described in Appendix B.

Reporting on achievement levels is one way in which NAEP results reach the general public and policymakers. Since 1990, the Governing Board has used achievement levels for reporting results on NAEP assessments; achievement level results indicate the degree to which student performance meets the standards set for what students should know and be able to do at the *NAEP Basic, NAEP Proficient,* and *NAEP Advanced* levels. Descriptions of achievement levels articulate expectations of performance at each grade level (see Exhibit 5.1). They are reported as percentages of students within each achievement level range, as well as the percentage of students at or above *NAEP Basic* and at or above *NAEP Proficient* ranges. Students performing at or above the *NAEP Proficient* level on NAEP assessments demonstrate solid academic performance and competency over challenging subject matter.

It should be noted that the *NAEP Proficient* achievement level does not represent grade level proficiency as determined by other assessment standards (e.g., state or district assessments) and there are significant differences between achievement in the context of NAEP as compared to the context of state level annual tests. For one, teachers and students are not expected to have studied the NAEP framework or systematically aligned state standards or local curricula with it, nor are students expected to study intensively for the assessment. Furthermore, the NAEP assessment is broader than a typical state grade level test, for NAEP covers multiple years of study and does not focus on specific instructional units and school years.

Results for students not reaching the *NAEP Basic* achievement level are reported as below *NAEP Basic*. As noted, individual student performance cannot be reported based on NAEP results.

Achievement Level Descriptions

Since 1990, the Governing Board has used achievement levels for reporting results on NAEP assessments. The achievement levels represent an informed judgment of "how good is good enough" in the various subjects that are assessed. Generic policy definitions for achievement at the *NAEP Basic, NAEP Proficient,* and *NAEP Advanced* levels describe in very general terms what students at each grade level should know and be able to do on the assessment. Achievement level descriptions specific to the 2025 NAEP Mathematics Framework can be found in Appendix A1. These will be used to guide item development and initial stages of standard setting for the 2025 NAEP Mathematics Assessment, if it is necessary to conduct a new standard setting.

The content achievement level descriptions may be revised for achievement level setting, if additional information is obtained or required. A broadly representative panel of exceptional teachers, educators, and professionals in mathematics will be convened to engage in a standard-setting process to determine cut scores that correspond to the achievement level descriptions. All achievement level setting activities for NAEP are performed in accordance with current best practices in standard setting and the Governing Board's (2018b) policy, *Developing Student Achievement Levels for the National Assessment of Educational Progress*. The Board policy does not extend to creating achievement level descriptions for performance below the *NAEP Basic* level.

Exhibit 5.1. Generic Achievement Level Policy Definitions for NAEP

Achievement Level	Definition
NAEP Advanced	This level signifies superior performance beyond *NAEP Proficient.*
NAEP Proficient	This level represents solid academic performance for each NAEP assessment. 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.
NAEP Basic	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the *NAEP Proficient* level.

Contextual Variables

NAEP law (The NAEP Law, 2017) requires reporting according to various student populations [see section 303(b)(2)(G)], including:

- Gender,
- Race/ethnicity,
- Eligibility for free/reduced-price lunch,
- Students with disabilities, and
- English language learners.

At times, people presume that the categories used to report data are related to causal explanations for observed differences, for example, that gender accounts for performance. Although differences in student achievement are often referred to as "achievement gaps," scholars have long found that these differences also represent gaps in students' opportunities to learn (e.g., Carter & Welner, 2013; Flores, 2007; Martin, 2009; Schmidt et al., 2015), as discussed in Chapter 1. When results are interpreted in ways that emphasize achievement gaps without attending to opportunity gaps, score differences across subgroups of students can be misinterpreted as differences in student ability, rather than differences due to unequal and inadequate educational opportunities.

The *Standards for Educational and Psychological Testing* (American Educational Research Association, [AERA], American Psychological Association [APA], and National Council on

Measurement in Education [NCME], 2014) recommends that reports of group differences in assessment performance be accompanied by relevant contextual information, where possible, to both discourage erroneous interpretation and enable meaningful analysis of the differences. That standard reads as follows:

> Reports of group differences in test performance should be accompanied by relevant contextual information, where possible, to enable meaningful interpretation of the differences. If appropriate contextual information is not available, users should be cautioned against misinterpretation. (AERA, APA, & NCME, 2014, Standard 13.6)

Contextual data about students, teachers, and schools are needed to fulfill the statutory requirement that NAEP include information, whenever feasible, for these groups which promotes meaningful interpretation. The important components of NAEP reporting are summarized in Exhibit 5.2.

Exhibit 5.2 Components of NAEP Reporting

Component	Key Characteristics
How Information Is Reported	Elements released to the public include: ● Results published mainly online with an interactive report card ● Dedicated website: Performance of various subgroups at the national level published online ● Online data tools with sample questions, performance associated with all collected contextual variables, item maps, and profiles of states and TUDA districts
What Is Reported	NAEP data are reported by: ● Percentage of students attaining achievement levels ● Scale scores ● Sample responses to illustrate achievement level definitions ● Contextual information from NAEP questionnaires

Contextual variables are selected to be of topical interest, timely, and directly related to academic achievement and current trends and issues in mathematics. In the past, a range of information has been collected as part of NAEP. In one analysis, Pellegrino, Jones, and Mitchell (1999) identified five existing categories of indicators: (1) student background characteristics; (2) home and community support for learning; (3) instructional practices and learning resources; (4) teacher education and professional development; and (5) school climate.

Contextual variables for the 2025 NAEP Mathematics Assessment will build on two broad categories: student factors and opportunity to learn factors. Student factors have been described as skills, strategies, attitudes and behaviors that are distinct from content knowledge and academic skills. Opportunity-to-learn factors have been described as whether students are exposed to opportunities to acquire relevant knowledge and skill in or out of school. These are described below.

Mathematics-Specific Contextual Variables

As noted in Chapter 1, research has informed an expanded view of the factors that shape opportunities to learn, including time, content and practices, instructional strategies (e.g., how students are grouped for learning; the mathematical tasks they engage in; the opportunities students have to reason, model, and debate ideas), and instructional resources (e.g., human, material, and social resources that shape student access to mathematics).

For example, research has demonstrated that what students learn is shaped by the availability of various mathematics programs, curricula, extracurricular activities geared toward mathematics, the percentage of teachers certified in mathematics, teacher years of experience, percentage of mathematics teachers on an emergency license or vacancies/substitute teachers in the school, and number of teachers with mathematics degrees, among other factors. Teachers' and administrators' beliefs about what mathematics is, how one learns mathematics, and who can learn mathematics also affect student learning. What students learn is shaped by their sense of identity and agency. Students who see themselves and who are seen by others as capable mathematical thinkers are more likely to participate in ways that further their learning; students who do not see themselves and are not seen by others as capable mathematical thinkers are likely to be disengaged. Steele, Spencer, and Aronson (2002), for example, found that even passing reminders that a student is a member of one group or another – often in this case a group that is stereotyped as intellectually or academically inferior – can undermine student performance.

There are countless factors that shape what and when students learn. The NAEP Mathematics student, teacher, and administrator surveys cannot possibly cover all such factors. Even though it would be helpful to ask students and teachers the same questions, that too is not possible given time constraints. Furthermore, questions about some factors may not be appropriate in the NAEP context. Given the constraints, not all of the topics proposed above can be addressed.

To support prioritization and ensure that NAEP results have appropriate context for interpretation, the Framework set the following topics receive the greatest emphasis in the 2025 NAEP Mathematics Assessment's contextual questionnaires (in order of priority).

1. *Mathematics content and practices*. The 2025 NAEP Mathematics Framework conceptualizes mathematics as both content and practices. Therefore, contextual variables related to mathematics content are expanded to include reference to mathematical practices as well. Interpreting students' achievement requires a basic understanding of what mathematics content and practices students have engaged with. Given variation across states in standards and frameworks, this information is crucial.
2. *Teacher factors*. Research demonstrates that teacher quality is a critical in-school factor in predicting student achievement. The Framework prioritizes the collection of data on teacher preparation and professional development, as well as teacher mathematical knowledge for teaching.
3. *Student mathematical identity*. Research demonstrates that students' perceptions of their mathematical identity directly relates to their mathematics learning. The Framework prioritizes gathering information about students' mathematical identities through questions that address student participation in activities such as discussion of mathematical ideas or evaluation of how a mathematics problem is framed.

4. *Instructional resources*. A range of resources influence instruction, including school climate, instructional leadership, additional instructional personnel, time, technology, curriculum, and materials. The Framework prioritizes gathering information about school resources that can inform the interpretation of results, including students' exposure to different types of technology, the time devoted to mathematics teaching and learning in school, and the curricular and instructional materials at teachers' and students' disposal to support learning. In terms of technology, questionnaires will emphasize what technology is *available* to support mathematics teaching and learning.
5. *Instructional organization and strategies*. Interpreting student achievement levels will also depend on understanding the instructional strategies used in mathematics class, including collaborating in small group work, engaging in mathematical discussions, and using a range of tools to represent and model mathematics. The Framework prioritizes gathering information both on the organization of classrooms and the instructional routines and approaches that teachers use. It also includes what technologies and formative assessments are *used* in instruction.

Conclusion

As the Nation's Report Card, NAEP reports on student achievement over time, presenting an analysis of national trends in students' mathematical competence. The NAEP Mathematics Assessment is designed to assess the achievement of groups of students through robust and challenging assessments that are well aligned with current understanding of the mathematics content and practices to be learned and that use technology in ways that maximize both student engagement and accessibility. The results of the assessment are informed by data on contextual variables that illuminate potential differences in opportunities to learn for students.

Based on current research, policy, and practice, the *NAEP Mathematics Framework* visioning and development process articulated several major goals: to expand attention to student engagement in reasoning about and doing mathematics, to adjust NAEP's mathematical domains and competencies, to leverage interactive multimedia scenario-based tasks as a way to provide more authentic tasks for students to complete and to increase the assessment's accessibility, and to develop an expansive conception of opportunities to learn that would inform the collection and use of contextual information. Accordingly, Chapters 2 and 3 describe the content and practices of mathematics that students should be measured on the 2025 NAEP Mathematics Assessment as the Nation's Report Card. Chapter 4 describes the expansion of the assessment in ways that prudently leverage technology's potential to increase authenticity and accessibility. Chapters 1 and 5 describe an expansive understanding of opportunities to learn, and the role that contextual information plays in meaningful interpretation of the results from future NAEP Mathematics Assessments based on the Framework.

The ultimate goal of our nation's schools is to ensure that every student has access to learning high quality mathematics. NAEP plays an important role in providing a broad picture of students' knowledge and skills in mathematics to the nation. NAEP scores, illuminated by relevant contextual information, can provide the public, families, students, and schools useful data on student performance that complements information provided by state tests that are more tightly aligned with specific state standards. As a view of present trends, it provides invaluable data to inform policy and practice in the future.

GLOSSARY

Abstracting and Generalizing: A NAEP Mathematical Practice involving decontextualizing, identifying commonality across cases, items, problems, or representations, and extending one's reasoning to a broader domain appropriate for the grade level and the mathematics being assessed.

ALDs: Achievement Level Descriptions. Descriptions of student performance at achievement levels (basic, proficient, and advanced), detailing what students should know and be able to do in terms of the mathematics content areas and practices.

Clickstream: Response and process data generated based on student interactions with tools and systems in scenario-based tasks.

Cognitive complexity: The state or quality of a thought process that involves numerous constructs, with many interrelationships among them. Such mental processing is often experienced as difficult or effortful.

Construct: An image, idea, or theory, especially a complex one formed from a number of simpler elements, and often embedded in a web of related ideas.

Context: The physical, temporal, historical, cultural, or linguistic setting for an event, performance, statement, or idea, and in terms of which such events or statements can be fully understood and assessed.

CCSS-M: Common Core State Standards: Mathematics

Collaborative Mathematics: A NAEP Mathematical Practice defined as the social enterprise of doing mathematics with others through discussion and collaborative problem solving whereby ideas are offered, debated, connected, and built-upon toward solution and shared understanding. Collaborative mathematics involves joint thinking among individuals toward the construction of a problem solution.

Constructed response: An open-ended, text-based response. Every constructed response item has a scoring guide that defines the criteria used to evaluate students' responses.

Contextual variable: Student, teacher, administrator, and school factors that shape students' opportunities to learn, including time, content, instructional strategies, and instructional resources.

Conversational responses: A response within a discourse-based or collaborative task in which students respond by selecting from two or more choices that reflect a conversation between characters described in the task.

Culturally responsive teaching: A pedagogy that recognizes the importance of including students' cultural and linguistic knowledge and experience in all aspects of learning (Ladson-Billings, 1994) so as to both affirm their identities and allow them to access their knowledge in service of learning; a broad category of instructional approaches that include culturally-relevant and culturally sustaining pedagogies, which share similar commitments to the inclusive use of students' cultural backgrounds, communities, and experience.

Deduction: Reasoning that makes a logical argument, draws conclusions, and applies generalizations to specific situations.

Discourse: Denotes written and spoken communications or "language-in-use" (Gee, 1999). Discourse can also refer to the totality of codified language used in a given field of intellectual enquiry and of social practice.

Discrete items: Stand-alone assessment items.

English language learner: Active learners of the English language who may benefit from various types of language support programs; students from a diverse set of backgrounds who often come from non-English-speaking homes and backgrounds, and who typically require specialized or modified instruction in both the English language and in their academic courses.

Funds of knowledge: The strengths students bring with them to the classroom, including academic and personal background knowledge, accumulated life experiences, skills and knowledge used to navigate everyday social contexts, and world views structured by broader historically and politically influenced social forces (Civil, 2016; González, Moll, & Amanti, 2005).

GAIMME: *Guidelines for Assessment and Instruction in Mathematical Modeling Education*. A report issued by a collaboration between the Society for Industrial and Applied Mathematics and the Consortium for Mathematics and Its Applications (U.S.), National Council of Teachers of Mathematics.

Generalization: The act of identifying a property that holds for a larger set of mathematical objects or conditions than the number of individually verified cases.

Induction: Reasoning that begins with specific observations to develop generalizations and conclusions; looking for patterns and making generalizations.

In-line choice items: Items in which students respond by selecting one option from one or more drop-down menus that may appear in various sections of an item.

Instructional practice: Teaching methods that guide interaction in the classroom.

Joint thinking: Working and thinking together on a shared goal, including sharing ideas with others; attending to and making sense of the mathematical contributions of others; evaluating the merit of others' ideas through agreement or disagreement; and productively responding to others' ideas through building on or extending ideas and connecting or generalizing across ideas.

Justifying and Proving: A NAEP Mathematical Practice that involves creating, evaluating, showing, proving, or refuting mathematical arguments/suppositions in developmentally and mathematically appropriate ways.

Mathematical argumentation: The action or process of reasoning systematically in support of an idea, action, or theory.

Mathematical knowledge for teaching: The specialized knowledge mathematics teachers need to support their students' learning that goes beyond the mathematics that any educated adult might need; the mathematics-specific knowledge of content, pedagogy, and students that is needed to perform the recurrent tasks of teaching mathematics to students (Ball, et al., 2008).

Mathematical justification: A critical aspect of the NAEP Mathematical Practice of Justifying and Proving that includes creating arguments, explaining why conjectures must be true or demonstrating that they are false, exploring special cases or searching for counterexamples, understanding the role of definitions and counterexamples, and evaluating arguments.

Mathematical literacy: The application of numerical, spatial, or symbolic mathematical information to situations in a person's life as a community member, citizen, worker, or consumer.

Mathematical modeling: A NAEP Mathematical Practice that involves making sense of a scenario, identifying a problem to be solved, mathematizing it, and applying the mathematization to reach a solution and checking the viability of the solution.

Mathematical practice: The working methods of *doing* mathematics, including the NAEP Mathematical Practices of Representing, Abstracting and Generalizing; Justifying and Proving; Mathematical Modeling; and Collaborative Mathematics.

Mathematical proof: A formal proof is a specific type of argument "consisting of logically rigorous deductions of conclusions from hypotheses" (NCTM, 2000, p. 55). The form used to represent a mathematical proof is valid as long as it communicates the essential features of the proof, that is, it contains logically connected mathematical statements that are based on valid definitions and theorems.

Mathematical problem solving: Completing mathematical tasks where the task contexts may range from the purely mathematical to those that are experientially concrete or real to students.

Mathematical reasoning: A skill that involves using other mathematical skills, including evaluating situations, selecting problem-solving strategies, drawing logical conclusions, developing and describing solutions, and recognizing how those solutions can be applied.

Mathematical reasoners are able to reflect on solutions to problems and determine whether or not they make sense.

Object-based responses: Assessment responses that involve manipulating or using a physical object.

Opportunity gap: Relates to the inputs, the unequal or inequitable distribution of resources and opportunities that contribute to and perpetuate lower educational achievement and attainment based on race, ethnicity, socioeconomic status, English proficiency, community wealth, familial situations, or other factors.

Opportunity to learn: Inputs and processes that enable student achievement of intended outcomes.

PISA: The Program for International Student Assessment, an international assessment that measures 15-year-old students' reading, mathematics, and science literacy every three years.

Representing: A NAEP Mathematical Practice that involves recognizing, using, creating, interpreting, or translating among representations appropriate for the grade level and the mathematics being assessed.

Revoicing: A method of communication that can be used by students or teachers to "re-utter another's contribution through the use of repetition, expansion, or rephrasing" (Enyedy, et al, 2008, p. 135).

Scenario-based task: Assessment tasks that have both context and extended storylines to provide opportunities to demonstrate facility with mathematical practices.

Selected response: Assessment responses that involve a student selecting one or more response options from a given, limited set of choices.

Single selection multiple choice: Assessment items in which students respond by selecting a single choice from a set of given choices.

Student identity: A person's evolving view of self in a given social context influenced by their experiences, personal history, and other events. Students' mathematical identity is how they see themselves in relation to mathematics and mathematics learning (Bishop, 2012).

Tool-based responses: Assessment responses that involve manipulating or using a virtual tool on-screen (e.g., an on-screen ruler).

APPENDIX A: NAEP MATHEMATICS ACHIEVEMENT LEVELS DESCRIPTIONS

The Achievement Level Descriptions (ALDs) in this appendix provide examples of what students performing at the *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* achievement levels should know and be able to do in terms of the mathematics content areas and practices identified in the Framework. The intended audiences for these ALDs are the NAEP assessment development contractor and item writers; the ALDs help ensure that a broad range of items is developed at each assessed grade.

The ALDs in the 2025 NAEP Mathematics Framework have changed, relative to ALDs presented in the 2009–2017 Framework. The differences reflect not only changes to the mathematics knowledge, skills, and abilities assessed (math content areas and math practices) but also an effort to develop ALDs that provide explicit guidance for item developers. Specifically, across grade levels, the 2025 Framework ALDs have changed in the following ways:

- Updates to the grade-level objectives in Chapter 2 of the Framework are reflected in the content foci described through the levels at each grade level.
- Mathematical Practices are new to the 2025 Framework and are made explicit at every achievement level in every grade in these ALDs. The mathematical practices absorbed much of the reasoning and problem-solving language from the 2017 Framework ALDs. As noted in Chapter 3, some NAEP Mathematics items will not assess a NAEP Mathematical Practice. Thus, some elements of the NAEP Mathematics ALDs are not linked to a NAEP Mathematical Practice. Instead, they are associated with other activity such as knowledge of mathematical facts, procedural fluency, and mathematical practices that are not included in the five identified for the NAEP Mathematics Assessment.
- Although Chapter 4 of the Framework provides examples of assessment technology (e.g., graphing tools) that may be common in 2025 and beyond, these ALDs have reduced the focus on technology-specific descriptions of the mathematics students should know and be able to do on the NAEP Mathematics Assessment.
- To provide specific and unambiguous guidance to item developers, these ALDs provide more explicit elaborations of the knowledge and skills students should demonstrate and the actions they should perform at each grade level and within each achievement level.

Within each grade level, the shifts from one achievement level to the next share commonalities and the content of each achievement level can be described generally. Descriptions at each achievement level for all grade levels are described below.

- Descriptions at the *NAEP Basic* level focus on emerging understanding of on-grade-level concepts and introductory engagement with mathematical practices.
- Descriptions at the *NAEP Proficient* level focus on application of on-grade-level concepts and skillful engagement with mathematical practices.
- Descriptions at the *NAEP Advanced* level focus on extension of on-grade-level concepts and expert engagement with mathematical practices.

Text that elaborates on these statements is included within the ALD tables.

Furthermore, to add clarity and specificity, the 2025 *Framework* ALDs include example items targeting each achievement level within each grade level. Following the ALDs presentation, in Appendix A2, three sets of items (one set each for grades 4, 8, and 12) illustrate the knowledge and skills required at different NAEP achievement levels. The items are not intended to represent the entire set of mathematics content areas or practices, nor do the items imply priority or importance of some content areas or practices above others.

Finally, to guard against misinterpretations, it is important to clarify the intended meaning of the term *routine*, which is used frequently in the ALDs. For the purposes of the ALDs, *routine* is defined as *having a readily available solution method*.

Mathematics Achievement Level Descriptions for Grade 4

NAEP Basic	**Grade 4 students performing at the *NAEP Basic* level should show evidence of emergent understanding of mathematics concepts and procedures in the five NAEP content areas. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed.** Grade 4 students performing at the *NAEP Basic* level should be able to estimate and perform paper and pencil computations with whole numbers (e.g., addition and subtraction within 1,000; multiplication and division within 100); understand the meaning of fractions and decimals, but not necessarily the relations between fractions and decimals; compare numbers to familiar benchmarks such as 0, ¼, ½, ⅔, ¾, and 1; identify or measure attributes of simple plane figures (e.g., triangles, rectangles, squares, and circles) and simple solid figures (e.g., cubes, spheres, and cylinders), choosing appropriate measuring tools and units of measure; and solve problems involving these concepts and procedures. Students should be able to represent whole numbers, fractions, and decimals using visual representations; draw or sketch simple plane figures from a written description; create a visual, graphical, or tabular representation of a given set of data; and recognize, describe (in words or symbols), or extend numerical and visual patterns. They should be able to explain or defend their strategy or solution (e.g., justify solutions to word problems through numeric representations and operations); make mathematical sense of a problem scenario, selecting and using visual, physical, or symbolic representations, as needed, to lead to a solution; and share ideas and revoice the ideas of others.
NAEP Proficient	**Grade 4 students performing at the *NAEP Proficient* level should be able to recognize when particular concepts, procedures, and strategies are appropriate, and select, integrate, and apply them to represent or model situations mathematically and solve problems requiring more than the application of a known procedure or strategy. Students should be able to reason about relationships involving the domains of number, space, or data. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed.**

	Grade 4 students performing at the *NAEP Proficient* level should be able to estimate and compute with whole numbers (within the guidelines set by the NAEP objectives) and determine whether and explain why the results are reasonable; identify, represent, compare, add, and subtract fractions and decimals, using visual representations to compare numbers and as tools to solve problems; identify or draw angles; draw or sketch simple plane and solid figures from a written description; read and interpret a single set of data, including the interpretation of graphical or tabular representations of data; extend their understanding of patterns to create a different representation of a pattern or sequence; and create, use, and defend visual representations of problem situations involving these concepts and procedures. In all content areas, students should be able to abstract or de-contextualize and re-contextualize ideas in routine problems using written and symbolic structures; create and evaluate mathematical arguments; explain why conjectures must be true or demonstrate that they are false; explore with examples or search for counterexamples and understand the role of counterexamples in mathematical arguments; determine assumptions, pose answerable questions, and determine tools to use as they interpret and solve problems; and make sense of and evaluate the mathematical contributions of others through expressing and defending agreement or disagreement.
NAEP Advanced	**Grade 4 students performing at the *NAEP Advanced* level should be able to apply conceptual understanding and procedural knowledge in non-algorithmic ways to complex and non-routine mathematical or real-world problems in the five NAEP content areas. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed.** Grade 4 students performing at the *NAEP Advanced* level should be able to solve complex and non-routine real-world problems in all NAEP content areas. These students should be able to draw logical conclusions from the results of a solution process; justify answers and solution processes by explaining how and why they were achieved; and use words or symbols to generalize a pattern appearing in a sequence or table. Students should be able to build on, analyze, and justify representations or mathematical models created by others; use structures and patterns to generate a rule and investigate conditions under which the rule applies; use a variety of grade-appropriate methods to justify or refute a mathematical statement using valid definitions, statements, or counterexamples; determine and use a series of processes to mathematize a complex or non-routine situation and evaluate the results obtained; and extend, connect, or generalize across the ideas of others.

Mathematics Achievement Level Descriptions for Grade 8

NAEP Basic	**Grade 8 students performing at the *NAEP Basic* level should show evidence of emergent understanding, recognition, and application of concepts and procedures in the five NAEP content areas. Students should show evidence of engagement in the five NAEP Mathematical Practices as detailed.** Grade 8 students performing at the *NAEP Basic* level should be able to estimate and perform paper-and-pencil computations with rational numbers, including integers; solve linear equations or inequalities; choose appropriate measuring tools and units of measure; and solve problems involving strategic reasoning with these concepts and procedures, including using proportional reasoning to represent and solve routine problems. Students should be able to visually represent rational numbers, including decimals and integers, and use these representations as tools to solve problems; draw or sketch polygons, circles, or semicircles from a written description; create a visual, graphical, or tabular representation of a given set of data; and recognize, describe (in words or symbols), or extend numerical and visual patterns. They should be able to explain or defend a strategy or solution (e.g., justify solutions to word problems through numeric representations and operations); make mathematical sense of a problem scenario, selecting and using visual, physical, or symbolic representations, as needed, to lead to a solution; and share ideas and revoice the ideas of others.
NAEP Proficient	**Grade 8 students performing at the *NAEP Proficient* level should show evidence of recognizing and applying concepts and procedures to solve problems requiring more than routine application of a known process or result in the five NAEP content areas. They should recognize when particular concepts, procedures, and strategies are appropriate and select, integrate, and apply them to represent or model situations mathematically. Students should be able to reason about relationships involving the domains of number, space, or data. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed.** Grade 8 students performing at the *NAEP Proficient* level should understand the connections among integers, fractions, percents, and decimals and be able to work across these sets of numbers to examine proportional and linear relationships; expand their understanding of algebraic relationships to translate between different representations, compare properties of two relationships each represented differently, identify linear functions, and use the structure of an algebraic expression to solve problems; estimate the size of an object with respect to a given measurement attribute (e.g., length, area, volume, angle measurement, weight, or mass); compare figures or objects with respect to a measurement attribute; identify, describe, and justify relationships of congruence, similarity, and symmetry; organize data in order to make inferences and draw conclusions, interpret data in terms of generalized phenomena (e.g., shape, center, spread, clusters), and make comparisons or explore differences

	within and among sets of data; and interpret and apply probability concepts to routine situations. In all content areas, students should be able to abstract or de-contextualize and re-contextualize ideas in routine problems using written and symbolic structures; create and evaluate mathematical arguments; explain why conjectures must be true or demonstrate that they are false; explore with examples or search for counterexamples and understand the role of definitions and counterexamples in mathematical arguments; determine assumptions, pose answerable questions, and determine tools to use as they interpret and solve problems; and make sense of and evaluate the mathematical contributions of others through expressing and defending agreement or disagreement.
NAEP Advanced	**Grade 8 students performing at the *NAEP Advanced* level should be able to apply conceptual understanding and procedural knowledge in non-algorithmic ways to complex and non-routine mathematical or real-world problems. They should also be able to justify, generalize, and apply concepts and procedures, and be able to synthesize concepts and processes in the five NAEP content areas. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed.** Grade 8 students performing at the *NAEP Advanced* level should be able to solve complex and non-routine real-world problems in all NAEP content areas. They should be able to probe examples and counterexamples in order to shape generalizations from which they can develop mathematical models; use number sense and geometric awareness (e.g., definitions, properties of and relationships between geometric figures, results of transformations) to consider the reasonableness of an answer; and create problem-solving techniques, explaining the reasoning processes underlying their conclusions. Students should be able to use, analyze, and justify representations created by others; use structures and patterns to generate a rule and investigate conditions under which the rule applies; use a variety of grade-appropriate proof methods to justify a mathematical statement using valid definitions, statements, or counterexamples; determine and use a series of processes to mathematize a complex or non-routine situation and evaluate the results obtained; and extend, connect, or generalize across the ideas of others.

Mathematics Achievement Level Descriptions for Grade 12

NAEP Basic	**Grade 12 students performing at the *NAEP Basic* level should show evidence of emergent understanding, recognition, and application of concepts and procedures in the five NAEP content areas. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed.** Grade 12 students performing at the *NAEP Basic* level should be able to estimate and perform computations with real numbers, including irrational numbers; select appropriate units related to representing or measuring an attribute of an object; identify and describe relationships of congruence, similarity, and symmetry; organize data in order to make inferences and draw conclusions; interpret data in terms of generalized phenomena (e.g., shape, center, spread, clusters); make comparisons or explore differences within and among sets of data; interpret and apply probability concepts to routine situations; recognize, identify, and interpret information about functions presented in various forms; and solve problems involving these concepts and procedures, including using the coordinate plane to model and solve routine problems. Students should be able to represent real numbers, including very large and very small numbers, using visual representations and numerical expressions (e.g., scientific notation), and use these representations and expressions as tools to solve problems; draw or sketch plane figures and planar images of three-dimensional figures from a written description; create a visual, graphical, or tabular representation of a given set of data; and recognize, describe, or extend numerical patterns, including arithmetic and geometric progressions. They should be able to explain or defend a strategy or solution (e.g., justify solutions to word problems through numeric representations and operations); make mathematical sense of a problem scenario, selecting and using visual, physical, or symbolic representations, as needed, to lead to a solution; and share ideas and revoice the ideas of others.
NAEP Proficient	**Grade 12 students performing at the *NAEP Proficient* level should be able to recognize when particular concepts, procedures, and strategies are appropriate and to select, integrate, and apply them to represent or model situations mathematically to solve problems requiring more than the application of a known result. Students should be able to reason about relationships involving the domains of number, space, or data. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed**. Grade 12 students performing at the *NAEP Proficient* level should be able to solve complex non-routine items using algebraic and geometric approaches. Students should be able to find, test, and validate geometric and algebraic results and conjectures using a variety of methods. They should be able to design and carry out statistical surveys and experiments and interpret results that are obtained by

	them or by others. Students should also be able to translate between representations of functions (linear and nonlinear, quadratic and exponential), including verbal, graphical, tabular, and symbolic representations. In all content areas, students should be able to abstract or de-contextualize and re-contextualize ideas in routine problems using written and symbolic structures; create and evaluate mathematical arguments; explain why conjectures must be true or demonstrate that they are false; explore with examples or search for counterexamples and understand the role of definitions and counterexamples in mathematical arguments; determine assumptions, pose answerable questions, and determine tools to use as they interpret and solve problems; and make sense of and evaluate the mathematical contributions of others through expressing and defending agreement or disagreement.
NAEP Advanced	**Grade 12 students performing at the *NAEP Advanced* level should demonstrate in-depth knowledge of and be able to reason about mathematical concepts and procedures in the realms of number, algebra, geometry, and statistics. Students should also show evidence of engagement in the five NAEP Mathematical Practices as detailed.** Grade 12 students performing at the *NAEP Advanced* level should be able to defend their solutions to complex non-routine tasks. Students should be able to reason about and with functions and transformations, using properties of functions and transformations to analyze relationships and to determine and construct appropriate representations for solving problems; explain or defend reasoning processes; and understand the role of hypotheses, deductive reasoning, and conclusions in geometric proofs and algebraic arguments made by themselves and others. Students should be able to use, analyze, and justify representations created by others; use structures and patterns to generate rules and investigate the conditions under which rules apply; use a variety of grade-appropriate proof methods to justify a mathematical statement using valid definitions, statements, theorems, or counterexamples; determine and use a series of processes to mathematize a complex or non-routine situation and evaluate the results obtained; and extend, connect or generalize across the ideas of others.

NAEP Basic, NAEP Proficient, and NAEP Advanced Achievement Levels for Grade 4

NAEP Basic, Grade 4

In this item, students are given a representation for a number and asked to choose the number, addressing NAEP Basic level language "identify or measure attributes of simple plane figures."

Subject: Mathematics, Grade: 4, Year: 2013
Content Classifications: Geometry, Low, Type: MC, Difficulty Level: Easy

The correct answer is:

A. Pentagon

Part of a closed shape is shown above. When the shape is completed, which of these could it be?

A. Pentagon

B. Rectangle

C. Square

D. Triangle

NAEP Proficient, Grade 4

In this item, students are presented with a problem situation involving multi-step computation and interpretation within the context of the situation, addressing NAEP Proficient level language “Students should be able to estimate and compute with whole numbers (within the guidelines set by the NAEP objectives)” and “abstract or de-contextualize and re-contextualize ideas in routine problems.”

Subject: Mathematics, Grade: 4, Year: 2017
Content Classifications: Number properties and operations, Moderate, Type: SR, Difficulty Level: Hard

A school will receive between $600 and $900 to spend on art supplies.

The money will be given to three school clubs.

Each school club will get the same amount of money.

Which of the following amounts of money could each school club get?

Select all the correct answers.

- A ☐ $145
- B ☐ $225
- C ☐ $295
- D ☐ $325
- E ☐ $355

Clear Answer

NAEP Advanced, Grade 4

In this item, students are presented with a specific mathematical scenario and asked to generalize the results and provide a justification for the generalization, addressing NAEP Advanced level language “use structures and patterns to generate a rule” and “use a variety of grade-appropriate methods to justify or refute a mathematical statement [the rule] using valid definitions, statements, or counterexamples.”

Subject: Mathematics, Grade: 4, Year: 2011
Content Classifications: Number properties and operations, High, Type: SCR, Difficulty Level: Hard

Mr. Jones picked a number greater than 100.

He told Gloria to divide the number by 18.

He told Edward to divide the number by 15.

Whose answer is greater?

◯ Gloria's ◯ Edward's

Explain how you know this person's answer will always be greater for any number that Mr. Jones picks.

NAEP Basic, NAEP Proficient, and NAEP Advanced Achievement Levels for Grade 8

For all of Items 1 through 4 below, refer to the following three figures:

Figure 1

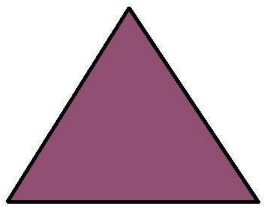


Figure 2

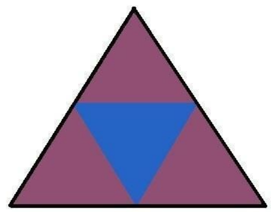
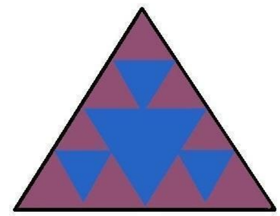


Figure 3



NAEP Basic, Grade 8

Item 1

Figure 1 is an equilateral triangle and *s* is the length of a side of the triangle. *P* is the perimeter of the triangle in Figure 1. Complete the equation for the perimeter, *P*, of Figure 1.

$$P = \square \cdot s$$

This item is an indicator of *NAEP Basic* because students are asked to recognize or apply directly procedures and representations that are routine at grade 8 regarding perimeter of triangles.

Item 2

In Figure 2 the blue triangle has been created by connecting the midpoints of the sides of the original triangle in Figure 1. Indicate if each of the following statements is true or false:

a) The perimeter of the blue triangle is one-fourth the perimeter of the original triangle
b) The perimeter of the blue triangle is one-half the perimeter of the original triangle
c) The area of the blue triangle is one-fourth the area of the original triangle
d) The area of the blue triangle is one-half the area of the original triangle

This item is an indicator of *NAEP Basic* because students are asked to recognize or apply simple relationships regarding area and perimeter of triangles.

NAEP Proficient, Grade 8

Item 3

Figure 1 is an equilateral triangle, and *s* is the length of a side of the triangle. In Figure 2 the blue triangle has been created by connecting the midpoints of the sides of the original triangle. In Figure 3 the smaller blue triangles have been created by connecting the midpoints of the sides of each interior triangle in Figure 2.

1) Express the perimeter of the blue triangle in Figure 2 in terms of *s*.
2) Express the sum of the perimeters of all the blue triangles in Figure 3 in terms of *s*.

Item 3 is an indicator of *NAEP Proficient* because it involves applying a well-known procedure to solve a non-routine problem that should be accessible to grade 8 students and representing the solution using grade-level appropriate algebraic representations.

NAEP Advanced, Grade 8

Item 4

Figure 1 is an equilateral triangle. In Figure 2 the blue triangle has been created by connecting the midpoints of the sides of the original triangle. In Figure 3 the smaller blue triangles have been created by connecting the midpoints of the sides of each interior triangle in Figure 2. Suppose you continue this process of connecting midpoints to obtain subsequent figures (Figure 4, Figure 5, Figure 6, and so on).

1) Express the sum of the perimeters of all the blue triangles in Figure 5 in terms of *s*.
2) Express the sum of the perimeters of all the blue triangles in Figure 10 in terms of *s*.

Item 4 is an indicator of *NAEP Advanced* because it involves generalizing a pattern and using a well-known procedure in the context of the pattern to solve a non-routine problem, and representing the solution using grade-level appropriate algebraic representations.

NAEP Basic, NAEP Proficient, and NAEP Advanced Achievement Levels for Grade 12

NAEP Basic, Grade 12

In this item, students are given pairs of shapes and asked to identify the pair that must always be similar, addressing NAEP Basic level language "identify and describe relationships of congruence, similarity, and symmetry."

Subject: Mathematics, Grade: 12, Year: 2005
Content Classifications: Geometry, Low, Type: MC, Difficulty Level: Medium

The correct answer is:

A. Two equilateral triangles

Which of the following pairs of geometric figures must be similar to each other?

A. Two equilateral triangles

B. Two isosceles triangles

C. Two right triangles

D. Two rectangles

E. Two parallelograms

NAEP Proficient, Grade 12

In this item, students are asked to select the data collection method most appropriate for the question of interest, addressing NAEP Proficient level language "They should be able to design and carry out statistical surveys."

Subject: Mathematics, Grade: 12, Year: 2009
Content Classifications: Data analysis, Statistics, and Probability, Moderate, Type: MC, Difficulty Level: Medium

The correct answer is:

C. Randomly select 25 students from a list of all students at the school.

The principal of a high school would like to determine why there has been a large decline during the year in the number of students who buy food in the school's cafeteria. To do this, 25 students from the school will be surveyed. Which method would be the most appropriate for selecting the 25 students to participate in the survey?

A. Randomly select 25 students from the senior class.

B. Randomly select 25 students from those taking physics.

C. Randomly select 25 students from a list of all students at the school.

D. Randomly select 25 students from a list of students who eat in the cafeteria.

E. Give the survey to the first 25 students to arrive at school in the morning.

NAEP Advanced, Grade 12

In this item, students need to use geometric properties, definitions, and principles to describe a geometric process for finding the center of any circle, addressing NAEP Advanced level language "use a variety of grade-appropriate proof methods to justify a mathematical statement using valid definitions, statements, theorems, or counterexamples."

Subject: Mathematics, Grade: 12, Year: 1996
Content Classifications: Geometry, Problem solving, Type: ECR, Difficulty Level: Hard

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important to show all your work.

Describe a procedure for locating the point that is the center of a circular paper disk. Use geometric definitions, properties, or principles to explain why your procedure is correct. Use the disk provided to help you formulate your procedure. You may write on it or fold it any way that you find helpful, but it will not be collected.

APPENDIX C: SUMMARY OF VISIONING PANEL GUIDELINES

MATHEMATICS

1. EXPANSION OF ATTENTION TO STUDENT REASONING AND MATHEMATICAL PRACTICES
We recommend defining mathematical practice constructs of priority interest in the framework (e.g., representing, abstracting and generalizing, justifying and proving, modeling, mathematical collaboration), providing examples of how they can be assessed (e.g., in the Assessment and Item Specifications), and using these definitions to systematically assess these practices, integrated with content, in 2025.

2. SIGNIFICANT BROADENING OF MATHEMATICAL DOMAINS AND COMPETENCIES
The mathematics content of the preK-12 curriculum has significantly evolved, and these changes need to be reflected in NAEP. **We recommend** a broadening of the content in several ways, including:

- (a) content that reflects research on mathematics teaching and learning that responds to students' diverse experiences, backgrounds, language, and culture;
- (b) a re-examination of statistics, data analysis and probability concepts and skills in light of current scholarship and standards documents;
- (c) attention to a wider range of technological tools available for students;
- (d) highlighting foundational mathematical themes that cut across different areas of content domains (e.g., geometry, algebra) and the grade bands from grades 4 to 8 to 12; and
- (e) consideration of a new cross-cutting theme or content area (at grade 12) that expands on calculus-readiness and statistics to include increasingly relevant applied mathematics important to informed citizenship, to personal financial and other decisions, and a variety of careers.

3. ATTENTION TO THE BALANCE OF COGNITIVE DEMAND
NAEP's current levels of "mathematical complexity" afford a balance between low-level items that ask for recall or demonstration of procedures, medium-level items that require connection-making on multi-step procedures, and high-level items that require analysis, creativity, synthesis, or justification and proof. **We recommend** a NAEP mathematics framework update in terms of relevant research on mathematical complexity and cognitive demand.

TEST DESIGN AND TECHNOLOGY

4. TEST DESIGN
We recommend the integration of content and practice skills through leveraging interactive multimedia scenario-based tasks as a way to provide more authentic tasks for students to complete (e.g., NAEP Technology and Engineering Literacy; see online TEL tasks).

5. STRATEGIC USE OF TECHNOLOGY

We recommend that NAEP revisions leverage technology to increase the assessment's authenticity (allowing students to use the technologies they use in and out of school) and the assessment's accessibility. Given the digital divide, as the NAEP instrument evolves, panels should address known and potential implementation issues and recommend ways to mitigate issues of access and test-taking that could occur in under-resourced communities.

OPPORTUNITIES TO LEARN AND OPPORTUNITIES TO DEMONSTRATE LEARNING

6. EXPANSIVE CONCEPTION OF OPPORTUNITIES TO LEARN

We recommend developing a broad approach to the framework update that scaffolds attention to opportunities to learn mathematics content, processes, and practices. This intent should be woven into the objectives in the framework, the item types and examples, and realized in contextual variables used on surveys.

We recommend updates to contextual variables in surveys that include attention to students' views of mathematics, and of themselves as mathematics learners; students' views of their peers', teachers', and school's beliefs/interest in their progress in mathematics; students' views of mathematics teaching and mathematics assessment (including NAEP); student access to and engagement with the language and culture of the test; teachers' knowledge of what has been taught before NAEP is administered; and teachers' beliefs about mathematics, mathematics teaching, and what their students can do.

7. ACCESSIBLE ASSESSMENTS FOR ALL STUDENTS

We recommend developing authentic assessment items with multiple access points that provide diverse populations of students with opportunities to demonstrate their mathematical knowing and reasoning in creative, authentic ways. This includes improving the accessibility of the assessment through short term goals like reconsidering test time limits, establish testing conditions that are more closely aligned with learning conditions (the use of typical tools, for example, or allowing teachers to be present) as well as longer term efforts to document how the current assessment remains inaccessible. Items should have consequential validity, be engaging to students, reflect guidelines for "low floor, high ceiling" tasks that provide opportunities for multiple approaches, and connect to students' lived experiences and funds of knowledge. Making the testing technologies widely available to students and teachers well before the assessment would also increase access and authenticity. Finally, because some research suggests that using mathematics tasks situated in everyday situations allows students to bring greater meaning to those tasks, we believe the authenticity of assessment items may allow for a more successful assessment of the mathematics students are learning (Boaler, 2002; Tomaz & David, 2015).

APPENDIX D: PRACTICES AND CALCULATOR ACTIVITY BY OBJECTIVE

To assist item writers when coordinating decisions about item content, NAEP Mathematical Practices, and calculator use, the tables in this appendix include the objectives from Chapter 2, the NAEP Mathematical Practice(s) most likely to be assessed by each objective, and the likely calculator designation for items aligning to the content.

- **Inherent Practice(s).** The assignment of NAEP Mathematical Practices inherent to a set of objectives is based on the verb usage in the objectives and their alignment to the description of a particular practice or practices (see Chapter 3 for discussion of the NAEP Mathematics Practices). Some cells in the Inherent Practice(s) column do not name a NAEP Mathematical Practice, and instead contain "Other" or "Variable."
  - Other: As written, the content of this objective lends itself to recall, procedural fluency, or mathematical practice(s) other than NAEP Mathematical Practices.
  - Variable: The direction taken with the development of context, activity, and/or questions in items aligned to one or more of the listed content objectives would determine which, if any, of the NAEP Mathematical Practices also is assessed.
- **Calculator Designation.** The assignment of a calculator designation for a set of objectives is mostly concerned with objectives where the evidence is best demonstrated with or without a calculator (See Chapter 4 for further information on calculator access.)

Number Properties and Operations

Num–1. Number sense			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Identify place value and actual value of digits in whole numbers, and think flexibly about place value notions (e.g., there are 2 hundreds in 253, there are 25 tens in 253, there are 253 ones in 253).	a) Use place value to represent and describe integers and decimals.		Representing
b) Represent numbers using base 10, number line, and other representations.	b) Represent or describe rational numbers or numerical relationships using number lines and diagrams.		Representing
c) Compose or decompose whole quantities either by place value (e.g., write whole numbers in expanded notation using place value: 342 = 300 + 40 + 2 or 3 × 100 + 4 × 10 + 2 × 1) or convenience (e.g., to compute 4 × 27 decompose 27 into 25 + 2 because 4 × 25 is 100, and 4 × 2 is 8 so 4 × 27 is 108).			Representing; Abstracting and Generalizing

Num–1. Number sense (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
d) Write or rename whole numbers (e.g., 10: 5 + 5, 12 – 2, 2 × 5).	d) Write or rename rational numbers.	# d) Represent, interpret, or compare expressions for real numbers, including expressions using exponents and *logarithms.	Representing
e) Connect across various representations for whole numbers, fractions, and decimals (e.g., number word, number symbol, visual representations).	e) Recognize, translate or apply multiple representations of rational numbers (fractions, decimals, and percents) in meaningful contexts.		Representing; Abstracting and Generalizing
	f) Express or interpret large numbers using scientific notation from real-life contexts.	# f) Represent or interpret expressions involving very large or very small numbers in scientific notation.	Representing
	g) Find absolute values or apply them to problem situations.	g) Represent, interpret, or compare expressions or problem situations involving absolute values.	Representing
h) Recognize and generate simple equivalent (equal) fractions and explain why they are equivalent (e.g., by using drawings).	h) Order or compare rational numbers (fractions, decimals, percents, or integers) using various representations (e.g., number line).		Representing
i) Order or compare whole numbers, decimals, or fractions using common denominators or benchmarks.	i) Order or compare rational numbers including very large and small integers, and decimals and fractions close to zero.	i) Order or compare rational or irrational numbers, including very large and very small real numbers.	Other

Num–2. Estimation			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Use benchmarks (well-known numbers used as meaningful points for comparison) for whole numbers, decimals, or fractions in contexts (e.g., ½ and 0.5 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., π) in contexts.		Variable
b) Make estimates appropriate to a given situation with whole numbers, fractions, or decimals.	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.	# 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.	Variable
c) Verify and defend 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 or computer results.	# c) Verify solutions or determine the reasonableness of results in a variety of situations.	Representing
	d) Estimate square or cube roots of numbers less than 150 between two whole numbers.	d) Estimate square or cube roots of numbers less than 1,000 between two whole numbers.	Other

Num–3. Number operations			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Add and subtract using conventional or unconventional procedures (e.g., strategic decomposing and composing): • Whole numbers, or • Fractions and mixed numbers with like denominators.	a) Perform computations with rational numbers.	a) Find integral or simple fractional powers of real numbers.	Other
b) Multiply numbers using conventional or unconventional procedures (e.g., strategic decomposing and composing): • Whole numbers no larger than two digits by two digits with paper and pencil computation, or • Larger whole numbers using a calculator, or • Multiplying a fraction by a whole number.		b) Perform arithmetic operations with real numbers, including common irrational numbers.	Other
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.		c) Perform arithmetic operations with expressions involving absolute value.	Other
	d) Describe the effect of operations on size, including the effect of attempts to multiply or divide a rational number by: • Zero, or • A number less than zero, or • A number between zero and one, or • One, or • A number greater than one.	d) Describe the effect of multiplying and dividing by numbers including the effect of attempts to multiply or divide 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.	Abstracting and Generalizing

Num–3. Number operations (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
e) Interpret, explain, or justify whole number operations and explain the relationships between them.	e) Interpret, explain, or justify rational number operations and explain the relationships between them.	e) *Analyze or interpret a proof by mathematical induction of a simple numerical relationship.	Justifying and Proving
f) Solve problems involving whole numbers and fractions with like denominators.	f) Solve problems involving rational numbers and operations using exact answers or estimates as appropriate.	# f) Solve problems involving numbers, including rational and common irrationals.	Variable

Num–4. Ratios and proportional reasoning			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
	a) Use ratios to describe problem situations.		Representing
	b) Use fractions to represent and express ratios and proportions.		Representing
	c) Use proportional reasoning to model and solve problems (including rates and scaling).	# c) Use proportions to solve problems (including rates of change and per capita problems).	Representing; Abstracting and Generalizing
	d) Solve problems involving percentages (including percent increase and decrease, interest rates, tax, discount, tips, or part/whole relationships).	# d) Solve multistep problems involving percentages, including compound percentages.	Variable

Num–5. Properties of number and operations			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Identify odd and even numbers			Other
b) Identify factors of whole numbers	b) Recognize, find, or use factors, multiples, or prime factorization.		Other
	c) Recognize or use prime and composite numbers to solve problems.	c) Solve problems using factors, multiples, or prime factorization.	Variable
	d) Use divisibility or remainders in problem settings.	# d) Use divisibility or remainders in problem settings.	Variable
e) Apply basic properties of operations.	e) Apply basic properties of operations, including conventions about the order of operations as applied to integers and rational numbers.	e) Apply basic properties of operations, including conventions about the order of operations as applied to real numbers.	Variable
		f) Recognize properties of the number system (whole numbers, integers, rational numbers, real numbers, and *complex numbers) and how they are related to each other and identify examples of each type of number.	Abstracting and Generalizing

Measurement

Meas–1. Measuring physical attributes			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Identify the attribute that is appropriate to measure in a given situation.			Other
b) Compare objects with respect to a given attribute, such as length, area, capacity, 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, area, and volume.	Other
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.	Other
		d) Solve problems of angle measure, including those involving triangles or other polygons or parallel lines cut by a transversal.	Variable
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.		Other
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.	Variable
g) Solve problems involving area of squares and rectangles.			Variable
	h) Solve problems involving volume or surface area of rectangular solids, and volume of right cylinders and prisms, or composite shapes	h) Solve problems by determining, estimating, or comparing volumes or surface areas of three-dimensional figures.	Variable
	i) Solve problems involving rates such as speed or ratios such as population density.	# i) Solve problems involving rates and ratios such as speed, density, population density, or flow rates.	Variable

Meas–2. Systems of measurement			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Select or use an appropriate type of unit for the attribute being measured such as length, angle size, time, or temperature.	a) Select or use an appropriate type of unit for the attribute being measured such as length, area, angle, time, or volume.	a) Select or use an appropriate type of unit for the attribute being measured such as length, angle size, time, or temperature.	Other
b) Solve problems involving conversions within the same measurement system such as conversions involving inches and feet or hours and minutes.	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 the same measurement system such as conversions involving inches and feet or hours and minutes.	Variable
	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: 1 kilometer is approximately 0.6 mile. • Money: U.S. dollars to Canadian dollars. • Temperature: Fahrenheit to Celsius.		Other
d) Determine appropriate unit of measurement in problem situations involving such attributes as length, time, capacity, or weight.	d) Determine appropriate unit of measurement in problem situations involving such attributes as length, area, or volume.	d) Determine appropriate unit of measurement in problem situations involving such attributes as length, time, capacity, or weight.	Other
		# 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.	Variable
	f) Construct or solve problems (e.g., floor area of a room) involving scale drawings.		Variable

Meas–3. Measurement in triangles			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
		# a) Solve problems involving indirect measurement.	Variable
		b) Solve problems using the fact that trigonometric ratios (sine, cosine, and tangent) stay constant in similar triangles.	Variable
		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.	Variable
		d) * Interpret and use the identity $\sin2\theta + \cos2\theta = 1$ for angles θ between 0° and 90°; recognize this identity as a special representation of the Pythagorean theorem.	Variable
		e) * Determine the radian measure of an angle and explain how radian measurement is related to a circle of radius 1.	Abstracting and Generalizing
		f) * Use trigonometric formulas such as addition and double angle formulas.	Variable
		g) * Use the law of cosines and the law of sines to find unknown sides and angles of a triangle.	Variable
		h) * Interpret the graphs of the sine, cosine, and tangent functions with respect to periodicity and values of these functions for multiples of $\pi/6$ and $\pi/4$.	Variable

Geometry

Geom–1. Dimension and shape			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) 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).	a) Identify a geometric object given a written description of its properties.		Representing
b) Identify or draw angles and other geometric figures in the plane.	b) Identify, define, or describe geometric shapes in the plane and in three-dimensional space given a visual representation.	b) Give precise mathematical descriptions or definitions of geometric shapes in the plane and in three-dimensional space.	Other
	c) Draw or sketch from a written description polygons, circles, or semicircles.	c) Draw or sketch from a written description plane figures and planar images of three-dimensional figures.	Representing
		# d) Use two-dimensional representations of three-dimensional objects to visualize and solve problems.	Representing
e) Describe or distinguish among attributes of two- and three-dimensional shapes.	e) Demonstrate an understanding of two- and three-dimensional shapes in the world through identifying, drawing, reasoning from visual representations, composing, or decomposing.	# e) Analyze properties of three-dimensional figures including prisms, pyramids, cylinders, cones, spheres and hemispheres.	Representing; Abstracting and Generalizing; Justifying and Proving

Geom–2. Transformation of figures and preservation of properties			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
	a) Identify lines of symmetry in plane figures or recognize and classify types of symmetries of plane figures.	a) Recognize or identify types of symmetries (e.g., translation, reflection, rotation) of two- and three-dimensional figures.	Other
		b) Give or recognize the precise mathematical relationship (e.g., congruence, similarity, orientation) between a figure and its image under a transformation.	Other
	c) Recognize or informally describe the effect of a transformation (reflection, rotation, translation, or dilation) on two-dimensional figures.	c) Perform or describe the effect of a single transformation (reflection, rotation, translation, or dilation) on two- or three-dimensional geometric figures.	Other
d) Recognize attributes (such as shape and area) that do not change when plane figures are subdivided and rearranged.	d) Predict results of combining, subdividing, and recombining shapes of plane figures and solids (e.g., paper folding, tiling, subdividing and rearranging the pieces).	d) Identify transformations of shapes that preserve the area of two-dimensional figures or the volume of three-dimensional figures.	Justifying and Proving
	e) Justify relationships of congruence and similarity and apply these relationships using scaling and proportional reasoning.	e) Justify relationships of congruence and similarity and apply these relationships using scaling, proportional reasoning, and established theorems.	Justifying and Proving
	f) Apply the relationships among angle measures, lengths, and perimeters among similar figures.	f) Apply the relationships among angle measures, lengths, perimeters and volumes among similar figures.	Variable
		g) Perform or describe the effects of successive (composites of) isometries and/or similarity transformations.	Representing; Justifying and Proving

Geom–3. Relationships between geometric figures			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Analyze or describe patterns in polygons when the number of sides increases, or the size or orientation changes.			Abstracting and Generalizing

Geom–3. Relationships between geometric figures (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
b) Combine simple plane shapes to construct a given shape.	b) Apply geometric properties and relationships in solving problems in two and three dimensions.	b) Apply geometric properties and relationships to solve problems in two and three dimensions.	Justifying and Proving
c) Recognize two-dimensional faces of three-dimensional shapes.	c) Represent problem situations with geometric figures to solve mathematical or real-world problems.	# c) Represent problem situations with geometric models to solve mathematical or real-world problems.	Representing; Abstracting and Generalizing
	d) Use the Pythagorean theorem to solve problems in two-dimensional situations.	# d) Use the Pythagorean theorem to solve problems in two- or three-dimensional situations.	Abstracting and Generalizing; Justifying and Proving
		e) Recall and interpret or use definitions and basic properties of congruent and similar triangles, circles, quadrilaterals, other polygons, parallel, perpendicular and intersecting lines, and associated angle relationships (e.g., in solving problems or creating proofs).	Justifying and Proving; Mathematical Modeling
f) Describe and compare properties of simple and compound figures composed of triangles, squares, and rectangles.	f) Describe, compare or analyze attributes of, or relationships between, triangles, quadrilaterals, and other polygonal plane figures.	f) Analyze attributes or relationships of triangles, quadrilaterals, and other polygonal plane figures.	Abstracting and Generalizing
	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.	Abstracting and Generalizing
		h) Make, test, and validate geometric conjectures using a variety of methods, including deductive reasoning and counterexamples.	Justifying and Proving
		i) * Analyze properties of circles and the intersections of lines and circles (inscribed angles, central angles, tangents, secants, and chords).	Abstracting and Generalizing; Justifying and Proving

Geom–4. Position, direction, and coordinate geometry			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
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 using distance between two points, the midpoint of a segment, or slopes of perpendicular or parallel lines.	Variable
	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.	Abstracting and Generalizing
	c) Visualize or describe the cross section of a solid.	c) Describe or identify conic sections and other cross sections of solids.	Abstracting and Generalizing
	d) Represent geometric figures using rectangular coordinates on a plane.	d) Represent two-dimensional figures algebraically using coordinates and/or equations.	Representing
		e) * Use vectors to represent velocity and direction; multiply a vector by a scalar and add vectors both algebraically and graphically.	Representing
		f) Find an equation of a circle given its center and radius and, given an equation of a circle, find its center and radius.	Other
		g) * Graph or determine equations for images of lines, circles, parabolas, and other curves under translations and reflections in the coordinate plane.	Representing; Justifying and Proving
		h) * Represent situations and solve problems involving polar coordinates.	Abstracting & Generalizing

Data Analysis, Statistics, and Probability

Data–1. Data representation			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
Representations of data are indicated below for each grade level. Objectives in which only a subset of these representations is applicable are indicated in the parenthesis associated with the objective.			
Pictographs, bar graphs, circle graphs, dot plots, tables, and tallies.	Histograms, plots over time, dot plots, scatterplots, box plots, bar graphs, circle graphs, stem and leaf plots, frequency distributions, and tables.	Histograms, plots over time, dot plots, scatterplots, box plots, bar graphs, circle graphs, stem and leaf plots, frequency distributions, and tables, including two-way tables.	
a) Read or interpret a single distribution of data.	a) Read or interpret data, including interpolating or extrapolating from data.	# a) Read or interpret graphical or tabular representations of data.	Representing; Abstracting and Generalizing
b) For a given set of data, complete a graph (limits of time make it difficult to construct graphs completely).	b) For a given distribution of data, complete a graph and solve a problem using the data in the graph (histograms, plots over time, dot plots, scatterplots, bar graphs, circle graphs).	# b) For a given set of data, complete a graph and solve a problem using the data in the graph (histograms, plots over time, dot plots, scatterplots).	Representing
c) Answer statistical questions by estimating and computing within a single distribution of data.	c) Answer statistical questions by estimating and computing with data from a single distribution or across distributions of data.	c) Answer statistical questions involving univariate or bivariate distributions of data.	Other
	d) Given a graphical or tabular representation of a distribution of data, determine whether the information is represented effectively and appropriately (histograms, plots over time, dot plots, scatterplots, box plots, bar graphs, circle graphs).	# d) Analyze, compare and contrast different graphical representations of univariate and bivariate data (e.g., identify misleading uses of data in real-world settings and critique different ways of presenting and using information).	Representing; Justifying and Proving
		# e) * Organize and display data in a spreadsheet in order to recognize patterns and solve problems.	Representing

Data–2. Characteristics of data sets			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
	a) Calculate, use, or interpret mean, median, mode, range or shape of a distribution of data.	# a) Calculate, interpret, or use summary statistics for distributions of data including measures of center (mean, median), position (quartiles, percentiles), spread (range, interquartile range, variance, and standard deviation) or shape (skew, uniform, uni/bi-modal).	Representing
b) Given a distribution of whole number data in a context, identify and explain the meaning of the greatest value, the least value, or of any clustering or grouping of data in the distribution.	b) Describe a distribution of data using its mean, median, mode, range, interquartile range, and shape.	b) Recognize how linear transformations of one-variable data affect mean, median, mode, range, interquartile range, and standard deviation.	Abstracting and Generalizing
	c) Identify outliers and determine their effect on the mean, median, mode, or range.	# c) Determine the effect of outliers on the mean, median, mode, range, interquartile range, or standard deviation.	Abstracting and Generalizing
d) Compare two sets of related data using the greatest value, least value, or any clustering or grouping of 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.	d) Compare data sets using summary statistics (mean, median, mode, range, interquartile range, shape, or standard deviation) describing the same characteristic for two different populations or subsets of the same population.	Other
	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 predictions.	Representing; Justifying and Proving
		# f) Recognize or explain how an argument based on data might confuse correlation with causation.	Justifying and Proving
		g) * Identify and interpret the key characteristics of a normal distribution such as shape, center (mean), and spread (standard deviation).	Abstracting and Generalizing

Data–2. Characteristics of data sets (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
		# h) * Recognize and explain the potential errors that can arise when extrapolating from data.	Justifying and Proving
Data–3. Experiments and samples			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
	a) Given a sample, identify possible sources of bias in sampling.	# a) Identify possible sources of bias in sample survey populations or questions and describe how such bias can be controlled and reduced.	Mathematical Modeling
	b) Distinguish between a random and nonrandom sample.	b) Recognize and describe a method to select a simple random sample.	Mathematical Modeling
		# 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."	Abstracting and Generalizing
	d) Evaluate the design of an experiment.	d) Identify or evaluate the characteristics of a good survey or of a well-designed experiment.	Mathematical Modeling
		e) * Recognize the differences in design and in conclusions between randomized experiments and observational studies.	Justifying and Proving
Data–4. Probability			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
		# a) Determine whether two events are independent or dependent.	Other
	b) Using assumption of randomness, determine the theoretical probability of simple or compound events in familiar contexts.	# b) Using assumptions such as randomness, determine the theoretical probability of simple or compound events in familiar or unfamiliar contexts.	Other

Data–4. Probability (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
	c) Given the results of an experiment or simulation, estimate the probability of simple and compound events in familiar contexts.	# c) Given the results of an experiment or simulation, estimate the probability of simple or compound events in familiar or unfamiliar contexts.	Other
	d) Use theoretical probability to evaluate or predict experimental outcomes in familiar contexts.	# d) Use theoretical probability to evaluate or predict experimental outcomes in familiar or unfamiliar contexts.	Justifying and Proving; Mathematical Modeling
	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.	Representing
	f) Use a sample space to determine the probability of possible outcomes for an event.		Other
	g) Represent the probability of a given outcome using fractions, decimals, and percents.		Other
	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.	Other
		i) Determine conditional probability using two-way tables.	Representing
	j) Interpret and apply probability concepts to practical situations, and simple games of chance.	# j) Interpret and apply probability concepts to practical situations, including odds of success or failure in simple lotteries or games of chance.	Representing
		k) * Use the binomial theorem to solve problems.	Variable

Algebra

Alg–1. Patterns, relations, and functions			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Recognize, describe (in words or symbols), or extend simple numerical and visual patterns.	a) Recognize, describe, or extend numerical and visual patterns using tables, graphs, words, or symbols.	a) Recognize, describe, or extend numerical patterns, including arithmetic and geometric sequences (progressions).	Abstracting and Generalizing
		b) Express linear and exponential functions in recursive and explicit form given a verbal description, table, or some terms of a sequence.	Abstracting and Generalizing
c) Given a description, extend or find a missing term in a pattern or sequence.	c) Examine or create patterns, sequences, or linear functions expressed as a rule numerically, verbally, or symbolically.		Abstracting and Generalizing
d) Create a different representation of a pattern or sequence given a verbal description.			Representing
	e) Identify functions as linear or nonlinear or contrast distinguishing properties of functions from tables, graphs, or equations.	e) Identify or analyze distinguishing properties of linear, quadratic, rational, exponential, or *trigonometric functions from tables, graphs, or equations.	Other
	f) Interpret the meaning of slope or intercepts, or determine the rate of change between two points on a graph of a linear function.		Other
		g) Determine whether a relation, given in verbal, symbolic, tabular, or graphical form, is a function.	Representing
		h) Recognize and analyze the general forms of linear, quadratic, rational, exponential, or *trigonometric functions.	Representing
		i) Determine the domain and range of functions given in various forms and contexts.	Other

Alg–1. Patterns, relations, and functions (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
		j) * Given a function, determine its inverse if it exists and explain the contextual meaning of the inverse for a given situation.	Abstracting and Generalizing

Alg–2. Algebraic representations			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Translate between different representational forms (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.	Representing
	b) Interpret and compare representations of linear relationships expressed in symbols, graphs, tables, diagrams, or written descriptions.	# b) Interpret and compare representations of relationships expressed in symbols, graphs, tables, diagrams (including Venn diagrams), or written descriptions.	Representing
	c) Graph or interpret points represented by ordered pairs of numbers on a rectangular coordinate system.		Representing
	d) Solve problems involving coordinate pairs on the rectangular coordinate system.	d) Perform or interpret transformations on the graphs of linear, quadratic, exponential, and *trigonometric functions.	Representing
		e) Make inferences or predictions using an algebraic model of a situation.	Abstracting and Generalizing; Justifying and Proving; Mathematical Modeling

Alg–2. Algebraic representations (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
	f) Identify or represent functional relationships in meaningful contexts including proportional, linear, and common nonlinear relationships (e.g., compound interest, bacterial growth) in tables, graphs, words, or symbols.	f) Given a real-world situation, determine if a linear, quadratic, rational, exponential, *logarithmic, or *trigonometric function fits the situation.	Representing
		# g) Solve problems involving exponential growth and decay.	Variable
		h) *Identify distinguishing characteristics of exponential, logarithmic, and rational functions (e.g., discontinuity, asymptotes, concavity).	Other

Alg–3. Variables, expressions, and operations			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Use letters and symbols to represent an unknown quantity in a simple mathematical expression.			Representing
b) Express simple mathematical relationships using expressions, equations or inequalities.	b) Write algebraic expressions, equations, or inequalities to represent a situation.	b) Write algebraic expressions, equations, or inequalities to represent a situation.	Representing
	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.	Other
		d) Write equivalent forms of algebraic expressions, equations, or inequalities to represent and explain mathematical relationships.	Representing; Abstracting and Generalizing
		# e) Evaluate algebraic expressions, including polynomials and rational expressions.	Other
		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.	Other
		g) * Determine the sum of finite and infinite arithmetic and geometric series.	Abstracting and Generalizing
		h) Use basic properties of exponents and *logarithms to solve problems.	Other

Alg–4. Equations and inequalities			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
a) Find the unknown(s) in a whole number sentence (e.g., in an equation or simple inequality like [_] + 3 > 7).	a) Solve linear equations or inequalities (e.g., Solve for x in ax + b = c or ax + b = cx + d or ax + b > c).	a) Solve linear, rational, or quadratic equations or inequalities, including those involving absolute value.	Other
b) Interpret "=" as an equivalence between two values and use this interpretation to solve problems.		b) * Determine the role of hypotheses, logical implications, and conclusions in algebraic arguments about equality and inequality.	Other
c) Verify a conclusion using simple algebraic properties derived from work with numbers (e.g., commutativity, properties of 0 and 1).	c) Make, validate, and justify conclusions and generalizations about linear relationships.	c) Use algebraic properties to develop a valid mathematical argument.	Justifying and Proving
	d) 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).	# d) Analyze situations, develop mathematical models, or solve problems using linear, quadratic, exponential, or *logarithmic equations or inequalities symbolically or graphically.	Representing; Mathematical Modeling
	e) Interpret relationships between symbolic linear expressions and graphs of lines by identifying and computing slope and intercepts (e.g., in y = ax + b, know that a is the rate of change and b is the vertical intercept).	e) Solve (symbolically or graphically) a system of equations or inequalities and recognize the relationship between the analytical solution and graphical solution.	Representing
	f) Use and evaluate common formulas (e.g., relationship between a circle's circumference and diameter [C = πd], distance and time under constant speed).	# f) Solve problems involving special formulas such as: $A = P(I + r)^t$ or $A = Pe^{(rt)}$.	Mathematical Modeling
		# g) Solve an equation or formula involving several variables for one variable in terms of the others.	Other
		h) * Solve quadratic equations with complex roots.	Other

APPENDIX E: SPECIAL STUDIES

Three special studies are proposed to support the 2025 NAEP Mathematics Framework. Special studies play a unique and essential role in the NAEP Mathematics ecosystem: advancing the vision set forth in the Framework. Some components of that vision may be aspirational – policies or processes that are essential for valid and reliable assessment of mathematics knowledge and skills, but that require advancements in assessment design, research, or technology to support implementation at scale.

Advancing the Assessment of NAEP Mathematical Practices and Mathematical Literacy

As a group, the NAEP Mathematical Practices introduced in Chapter 3 constitute an ideal topic area for special study. The practices are new to the Framework and a defining feature of the vision it presents. To ensure this vision is executed fully, NAEP should advance several strands of research and development, which correspond with the studies described below. First, NAEP should assess mathematical practices in authentic settings that resemble real-world problems where the practices are often applied. This strand is addressed in Study 1 below. Second, NAEP should signal the value of mathematical practices in the same way most large-scale assessment programs promote the importance of the skills they measure – by reporting results. This strand is addressed in Study 2 below. Third, the Framework emphasizes mathematical literacy as an essential component of mathematics knowledge and skills that NAEP Mathematics items can and should target. Mathematical literacy is not, however, included formally as one of the five content areas or as a one of the five NAEP Mathematical Practices. Therefore, a third strand of research should focus on the extent to which mathematical literacy can be measured and reported – accurately and reliably – under the requirements and constraints of a NAEP Mathematics operational administration. This strand is addressed in Study 3 below.

Study 1: Assessing Mathematical Practice in Context

Overview

Study 1 will examine ways to measure the NAEP Mathematical Practices by leveraging the rich data that scenario-based tasks (SBTs) and discrete items can generate in a digital assessment environment. The first phase of the study will establish a baseline by examining measurement features (e.g., content coverage, discrimination) of SBTs and other context-situated items linked to each NAEP Mathematical Practice. The second phase of the study will collect process data (e.g., activity logs), which will be recorded as students interact with elements of each situation (i.e., scenarios and discrete item contexts). This phase will also explore research-based methods for using process data to generate measures of student performance. Ultimately, this study will help NAEP not only determine the feasibility of capturing new process data through existing item types, but also gauge the measurement value of process data, either compared against or combined with response data.

Rationale

As interest in mathematical practices has grown, so too has the need for assessment approaches that allow students to demonstrate mastery of those practices. Under traditional item-delivery models, the steps that take place between the presentation of the item and the student's response are invisible in the assessment process. If those interim steps are meaningful to the assessed construct – in this case a NAEP Mathematical Practice – then decomposing items into their constituent steps could help sharpen inferences about a practice.

A new family of mathematics assessment approaches has emerged in response to this need, its defining feature being multiple points of measurement (e.g., item collections), often connected to a common stimulus or problem, and in some cases building on previous steps. These multistep approaches are particularly well-suited to the assessment of mathematical practices, because success depends not on the recollection of an isolated fact or theorem, but rather on the skill to draw upon multiple mathematics domains and solve complicated problems requiring multiple steps. These multistep approaches are also particularly well-suited to NAEP Mathematics, since the Framework places the NAEP Mathematical Practices alongside mathematics content as fundamental elements of mathematics assessment. This special study is intended to advance that vision, not only by leveraging current techniques for assessing practices, but also by extending those assessment techniques to learn more about students' response processes.

For NAEP Mathematics, the multistep approach will be accomplished through scenario-based tasks (SBTs) as well as other context-situated digitally-based items. Study 1 will use these items as a starting point and explore extensions that could generate even richer student performance data. For example, the typical SBT on a NAEP Science of NAEP TEL assessment is a cluster of items that ask students to complete a series of steps related to the same underlying scenario. Like any other collection of items, an SBT yields a group of item scores for each student, allowing NAEP to present students with engaging multistep problems while maintaining the same fundamental approaches to item scoring and psychometric scaling that are applied to other items.

Intuitively, SBTs offer a promising avenue for measuring the NAEP Mathematics Practices, so the first phase of Study 1 will involve testing that intuition, examining the measurement information SBTs produce across content areas and across the performance continuum. More importantly, the first phase will provide a baseline, characterizing the information that NAEP items and tasks provide about mathematics practices through *response* data (i.e., students' scored responses to a discrete item or to the standardized group of items administered through an SBT). If students' navigations through and interactions with these item's contexts are summarized only through response data alone, measurement information may be left on the table. Therefore, the second phase of the study will explore what additional information can be gained from *process* data.

In brief, process data are recordings of students' interactions with a digital environment. Clickstream data, activity logs, text, and transcribed voice responses are all examples. Once recorded, process data can be analyzed using a variety of statistical methods to produce measures of mathematical practice according to a standardized set of rules. One potential advantage to collecting and analyzing process data is the insight these data provide about students' actions that are not part of a formal item but are nonetheless relevant to mathematical practice. In addition, process data are collected passively, recording students' interactions rather than pausing a scenario to deliver an item, which could improve time efficiency.

However, process data also present new complexities. Student privacy concerns and available technology could each limit the variety and usefulness of available process data. In addition, process data are by definition untethered to item-writing rules and content targets, so establishing evidence of content validity (e.g., the alignment of a finite set of items to the NAEP Mathematics ALDs) for process data may prove challenging.

Nonetheless, it is reasonable to expect that an initial collection of analysis rules for process data could be developed as a component of Study 1, with each rule specifying what evidence a given analytical procedure (e.g., natural language processing) is designed to elicit and how that evidence maps to the 2025 NAEP Mathematics ALDs, which now include explicit references to the five NAEP Mathematical Practices. These analysis rules would function the same way item-development and scoring-rubric guidelines do for conventional NAEP Mathematics items, providing a clear link between the ALDs and students' response processes (for relatively structured assessment tasks like responding to a multiple-choice item or for relatively unstructured tasks like interacting with a digital environment). In fact, process data may increase the measurement information that can be gleaned from discrete items and SBTs, by generating even richer data at a smaller grain size.

Outcomes

Study 1 will produce three key outcomes in service of the 2025 Mathematics Framework's vision for the NAEP Mathematical Practices:

1) NAEP will characterize the measurement properties (e.g., content coverage, discrimination, potential bias, assessment time relative to measurement information) of items and tasks *as they relate to the NAEP Mathematical Practices*. NAEP currently collects these data for mathematics content areas, but the 2025 Framework is the first to explicitly include attention to five NAEP Mathematical Practices.
2) NAEP will determine the feasibility of collecting process data through different item types. A wide variety of process data have been suggested in the literature, but the variety of process data that can be collected within the constraints of a NAEP administration may be more limited.
3) NAEP will compile preliminary information about the value of process data in comparison to and as a companion to conventional response data, in terms of relevant information about student performance in mathematical practices.

Study 2: Reporting Results for Mathematical Practices

Overview

Study 2 will examine ways to provide information about the NAEP Mathematical Practices to the general public. The first phase will involve researching commonly used approaches for communicating assessment results, conceptualizing a limited set of reporting options, and producing sample reports. The second phase will involve gathering feedback on reporting approaches through focus groups with stakeholders and, if practicable, conducting structured A/B testing. This study is intended to produce feasible ways to provide information about the NAEP Mathematical Practices, under a key constraint: Unlike the content areas, the five NAEP

Mathematical Practices will not be scaled independently. Therefore, given the absence of scale scores for the practices, NAEP should avoid formats for reporting that risk confusion or misinterpretation. Although reporting options for NAEP Mathematical Practices affect any decision to scale practices separately, this study will not address the feasibility of scaling for the NAEP Mathematics Practices.

Rationale

The 2025 NAEP Mathematics Framework defines five NAEP Mathematical Practices, articulates how those practices should be assessed in various content areas, and positions the practices as a core component of assessing student achievement in mathematics – critical information for educators, parents, policymakers, and assessment developers. An important next step, as emphasized in the Framework, will be sharing the results with the general public. Releasing information about NAEP items and the NAEP Mathematical Practices should underscore the practices' fundamental importance in NAEP assessment.

Since NAEP Mathematical Practices are intertwined with NAEP mathematics content areas, the 2025 Framework's Technical Advisory Committee recommended against creating separate reporting scales for each practice. Instead, student performance on items that assess NAEP Mathematical Practices may be communicated descriptively, drawing upon common reporting approaches in large-scale assessment programs. The first phase of Study 2 will involve compiling a list of candidate reporting approaches based on a scan of what is done for other large-scale assessments (including, of course, other NAEP assessment programs). One option for descriptive reporting (item maps) is described next. Note that this example is provided for illustrative purposes only, and not as a suggested reporting tool for the NAEP Mathematical Practices. Any reporting approach would need to be evaluated in terms of its cost, its appeal to stakeholders, and the extent to which it maximizes effective communication and minimizes misinterpretation.

Item Maps Example

NAEP uses item maps to help illustrate what students know and can do in a variety of subject areas, including mathematics. In an item map, items are placed along the NAEP scale in each grade level. An item's position depends on its difficulty, which is estimated empirically using student response data. Items associated with higher scale scores are more difficult, requiring higher levels of knowledge and skills for a correct response.

An example item map for the 2017 NAEP Mathematics Assessment (Retrieved from https://www.nationsreportcard.gov/itemmaps/?subj=MAT&grade=4&year=2017) is presented in Exhibit E.1. Each item's description focuses on the knowledge and skills needed to respond successfully and "content classifications" icons refer to the specific content area being assessed. The same approach could be used to illustrate the relative difficulty of specific practices (by adding five NAEP Mathematical Practice icons).

Exhibit E.1. NAEP Mathematics Grade 4 Item Map

CONTENT CLASSIFICATIONS

Click on a classification to see a description.

● Number Properties and Operations	■ Measurement	▲ Geometry	▼ Data Analysis, Statistics, and Probability	♦ Algebra

500

▲ **333** Identify pairs of congruent figures—Correct (SR)
● **307** Identify the number represented by a set of base ten blocks (calculator available) (MC)
● **301** Identify multiple correct solution methods to an addition problem—Correct (SR)
♦ **286** Determine and apply a rule based on an input-output table (calculator available)—Satisfactory (CR)
● **282** Represent fractions using a model (calculator available)—Correct (SR)

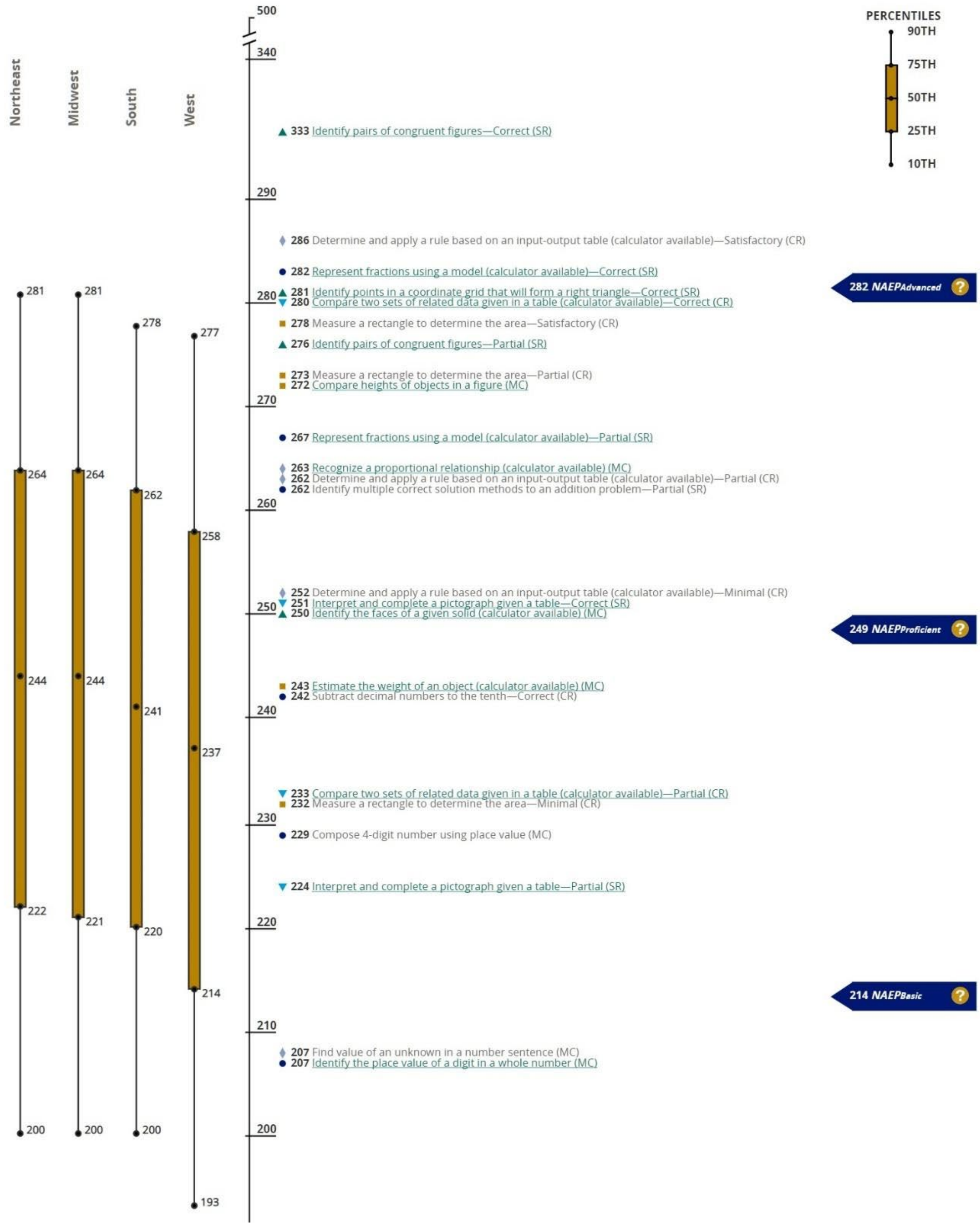
282 *NAEP Advanced*

▲ **281** Identify points in a coordinate grid that will form a right triangle—Correct (SR)
▼ **280** Compare two sets of related data given in a table (calculator available)—Correct (CR)
■ **278** Measure a rectangle to determine the area—Satisfactory (CR)
▲ **276** Identify pairs of congruent figures—Partial (SR)
■ **273** Measure a rectangle to determine the area—Partial (CR)
■ **272** Compare heights of objects in a figure (MC)
● **267** Represent fractions using a model (calculator available)—Partial (SR)

Item maps can be augmented to summarize student performance or enable comparisons across student groups. In Exhibit E.2, four box and whisker plots have been added to summarize student performance for the 2017 NAEP Mathematics Assessment in four U.S. regions (Retrieved from https://www.nationsreportcard.gov/itemmaps/?subj=MAT&grade=4&year=2017&jurisdiction=NT&variable=CENSREG). Another key component of the first phase of Study 2 would be estimating the time investment involved in each potential reporting solution. For example, in Exhibits B.1 and B.2, the addition of NAEP Mathematical Practices information would require new and existing NAEP items to be tagged with the practice(s) they feature but would not require additional scaling or standard-setting procedures.

Adding NAEP mathematical practices to the figures in Exhibits E.1 and E.2 would create a somewhat crowded visual, with not only five content-area icons but also five NAEP Mathematical Practices icons. One alternative could be to create a separate item map for each content area, and then include items within that content area tagged with each practice. To simplify the presentation even further, items could be removed from the maps and NAEP Mathematical Practices could instead be summarized in box and whisker plots. For example, in Exhibit E.2, NAEP mathematical practices could take the place of U.S. regions, and box and whisker plots would summarize the distribution of item difficulties associated with each practice. Again, this would require NAEP items to be tagged with practices but would not require new scaling or standard setting.

Exhibit E.2. NAEP Mathematics Grade 4 Item Map by U.S. Region



After a limited set of sample reports are created representing the candidate reporting solutions, the second phase of Study 2 will involve report field-testing. A plausible first step in field-testing would be to convene geographically diverse stakeholder focus groups to solicit feedback on each report's clarity, simplicity, and any areas that raise the risk of confusion of misinterpretation. Focus group panelists may also be asked to provide their interpretations of the report data (anonymously, to avoid peer influence). Unexpected interpretations in the anonymous feedback may highlight problem areas requiring further evaluation and development. If practicable, A/B testing could be added in the second phase, offering a more formal approach to comparing interpretations. Typical A/B testing involves randomly assigning consumers of visual data to one of two (or more) formats; the accuracy of consumers' interpretations across the two groups can be compared to help identify the format that minimizes confusion and misinterpretation.

Outcomes

Study 2 will produce two key outcomes in service of the 2025 Mathematics Framework vision for mathematical practices:

1) NAEP will research common reporting approaches and then assess the viability of replicating or adapting those approaches in the NAEP context. This feasibility study will generate or supplement a set of practical considerations NAEP can use when considering the adoption of other large-scale programs' reporting methods. In addition, this study will produce one or more candidates for consideration by other NAEP programs seeking to report on domains without scale scores.
2) NAEP will determine a useful and appropriate reporting formats for the NAEP Mathematical Practices. This will allow NAEP to signal the value of mathematical practices (an essential element of the Framework vision) without disseminating reports that risk widespread misinterpretation.

Study 3: Investigating Options for Assessing and Reporting Mathematical Literacy

Overview

Study 3 will focus on the extent to which mathematical literacy can be assessed and potentially reported via collections of NAEP Mathematics items and content objectives in grade 12. The first phase of the study will focus on the mathematical literacy construct itself; empirical analyses will help NAEP determine the precision and accuracy with which mathematical literacy can be measured and whether student performance in mathematical literacy constitutes a new dimension separate from the existing content areas and practices. Provided students' mathematical literacy skills are separable from other content knowledge and practices, the second phase of the study will investigate options for reporting on mathematical literacy. The second phase, therefore, may share many design features and decision points with Study 2 (reporting on mathematical practices). Ultimately, this study will help NAEP determine the feasibility assessing mathematical literacy and identify potential item-development or psychometric issues that would need to be addressed in order to do so.

Rationale

Relative to previous NAEP Mathematics frameworks, the 2025 Framework increases the focus on the assessment of mathematical literacy, in particular in grade 12. First, the Framework provides a definition of mathematical literacy, drawing on the definition introduced in PISA assessments:

> *Mathematical literacy is the application of numerical, spatial, or symbolic mathematical information to situations in a person's life as a community member, citizen, worker, or consumer.*

As noted in the 2025 Framework, a variety of NAEP items assess student actions and knowledge that could be viewed as requiring mathematical literacy (e.g., making decisions about personal finances; understanding quantitative information in print and visual media; making the accurate measurements in order to prepare a meal). Mathematical literacy can be found in the objectives in grades 4 and 8, but until the 2025 Framework mathematical literacy at grade 12 had received comparatively little attention. In the 2025 Framework, some grade 12 objectives are identified with a hashtag (#), if there are everyday applications of the objective to situations in a person's life as a community member, citizen, worker, or consumer. These hashtags have been included in the Framework for two reasons – to encourage the development of items measuring mathematical literacy and to support the identification of existing items in order to explore the feasibility of assessing and reporting on mathematical literacy.

This special study, therefore, is intended as a first step in the investigation of mathematical literacy as an assessable and reportable construct under the requirements and constraints of a NAEP Mathematics operational administration. Depending on the results of this study, future frameworks might identify mathematical literacy as a new content area at one or more grade levels. Alternatively, future frameworks may call for additional research, such as a special study focused on curriculum or assessment frameworks around the world that include mathematical literacy as a significant area.

Because this special study includes an analysis of options for reporting on mathematical literacy, it may share some common elements with Study 2 (reporting on mathematical practices). However, prior to considering reporting options, NAEP must first examine the assessability of mathematical literacy as a construct. Although it has been defined in the mathematics education literature and measured by other large-scale assessment programs, mathematical literacy is not currently a NAEP Mathematics content area or a NAEP Mathematical Practice. Therefore, the first phase of this study will focus on whether student performance in mathematical literacy is meaningfully different from performance in existing content areas and practices.

The educational measurement literature offers numerous well-understood and widely used methodologies (e.g., confirmatory factor analysis, Item Response Theory model-fit tests) for examining an assessment's dimensionality and the separability of the constructs it is meant to quantify. In addition to dimensionality tests, internal consistency statistics will provide lower-bound estimates of the reliability of students' mathematical literacy scores. This special study could also incorporate the judgment of subject-matter experts early in the assessment process by asking an independent group of mathematics content and assessment experts to identify mathematical literacy items among a larger set of items targeting various mathematics content areas and practices. If experts consistently distinguish mathematical literacy items from items that target other constructs, that would be promising (albeit incomplete) evidence of the degree to which mathematical literacy can be assessed as a unique aspect of mathematics knowledge and skill. This judgment-based study could also inform item development, highlighting the items and item features that promote identifications with mathematical literacy.

Depending on the results of the first phase of this special study, NAEP may next conduct a systematic investigation of options for *reporting* on mathematical literacy. The steps in this phase could largely mirror the design and sequence of Study 2. NAEP would first conduct a landscape scan, reviewing existing large-scale assessment programs' approaches to reporting on constructs with multiple subdomains. For example, the exploratory approach illustrated by item maps in Figures 2 and 3 may be suitable for mathematical literacy. If mathematical literacy is considered as a potential content area (rather than a practice), one reporting option would be to add a mathematical literacy icon to current NAEP Mathematics item maps. Alternatively, box and whisker plots could be presented to compare the distributions of mathematical literacy items that demand different NAEP Mathematical Practices (e.g., representing versus abstracting and generalizing versus mathematical modeling). Then, similar to Study 2, focus groups and A/B testing can be employed to verify that the *intended* interpretations of a reporting format aligns with *actual* interpretations by consumers of the report.

Regardless of the specific methodology, it is important to emphasize the exploratory nature of Study 3. It is not intended to produce procedures for scaling mathematical literacy separately from existing content areas. Even if the findings from dimensionality and reliability analyses in the first phase suggest it would be feasible, adding a subscale to the NAEP Mathematics Assessment would require extensive empirical analysis and deliberation. Rather, Study 3 represents a first step, providing foundational information about the role of mathematical literacy in the NAEP Mathematics Assessment.

Outcomes

This study will produce two key outcomes in service of the 2025 Mathematics Framework's vision for mathematical literacy:

1) NAEP will determine whether mathematical literacy is a unique dimension that can be measured accurately, reliably, and separately from each math content area and each math practice.
2) To signal the importance of mathematical literacy, NAEP will develop considerations for valid and straightforward reporting of mathematical literacy results.

APPENDIX F: NAEP MATHEMATICS PROJECT STAFF AND PANELS

Visioning Panel

[* indicates the subgroup who drafted the Framework as part of the Development Panel]

June Ahn*
Associate Professor, Education
University of California, Irvine
Irvine, CA

Kevin D. Armstrong
Middle School Principal
Metropolitan Nashville Public Schools, TN / National Association of Elementary School Principals (NAESP)
Nashville, TN / Alexandria, VA

Joan Elizabeth Auchter
Director, Professional Learning
National Association of Secondary School Principals (NASSP)
Reston, VA

Robert Q. Berry
Professor, Education / President
University of Virginia / National Council of Teachers of Mathematics (NCTM)
Charlottesville, VA / Reston, VA

David Bressoud
Professor, Mathematics / Director
Macalester College, MN / Conference Board of the Mathematical Sciences (CBMS)
St. Paul, MN / St. Paul, MN

Jinghong Cai
Research Analyst, Center for Public Education
National School Boards Association
Alexandria, VA

Diana Aurora Ceja*
Administrator
Riverside County Office of Education / TODOS: Mathematics for ALL
Riverside, CA / Tempe, AZ

Linda Ruiz Davenport*
Director of K-12 Mathematics
Boston Public Schools
Boston, MA

Sarah Ann DiMaria*
Mathematics Teacher
Austin Unified School District / Knowles Teacher Initiative
Austin, TX / Moorestown, NJ

Marielle Edgecomb*
Mathematics Teacher
Peninsula School District, ME
Prospect Harbor, ME

Amy Burns Ellis*
Associate Professor, Education
University of Georgia
Athens, GA

Linda Furuto
Professor, Education
University of Hawai'i at Mānoa
Honolulu, HI

Dewey Gottlieb
Educational Specialist / President
Hawai'i Department of Education / Association of State Supervisors of Mathematics (ASSM)
Honolulu, HI

Victoria Marguerite Hand
Associate Professor, Education
University of Colorado Boulder
Boulder, CO

Raymond Hart
Director of Research
Council of the Great City Schools
Washington, DC

Daniel Joseph Heck
Vice President
Horizon Research, Inc.
Chapel Hill, NC

Joan Herman
Director Emerita, CRESST
University of California, Los Angeles
Los Angeles, CA

Kelli Millwood Hill*
Director of Efficacy & Research
Khan Academy
Mountain View, CA

Chris Hulleman
Research Associate Professor, Education
University of Virginia
Charlottesville, VA

Jennifer Langer-Osuna*
Assistant Professor, Education
Stanford University
Stanford, CA

Katherine Elizabeth Lewis
Assistant Professor, Education
University of Washington
Seattle, WA

Kelly S. Mix *
Professor and Chair, Department of Human Development and Quantitative Methodology, Education
University of Maryland
College Park, MD

Sorsha-Maria T. Mulroe*
Mathematics Support Teacher
Howard County Public School System
Ellicott City, MD

Nora G. Ramirez*
Executive Secretary
TODOS: Mathematics for ALL
Tempe, AZ

J. Michael Shaughnessy*
Professor Emeritus, Mathematics & Statistics
Portland State University
Portland, OR

Edward Alan Silver*
Senior Associate Dean and Professor, Education and Mathematics
University of Michigan
Ann Arbor, MI

Joi A. Spencer*
Associate Dean and Associate Professor, Education
University of San Diego
San Diego, CA

Maria Teresa Tatto
Professor, Education
Arizona State University
Tempe, AZ

Zalman P. Usiskin*
Professor Emeritus, Education
University of Chicago
Chicago, IL

Nadja Young
Senior Manager, Federal Government Civilian Projects
SAS Institute
Cary, NC

Suzanne M. Wilson, Panel Chair*
Professor and Head, Department of Curriculum and Instruction, Education
University of Connecticut
Storrs, CT

Technical Advisory Committee

Derek C. Briggs
Professor, Research and Evaluation Methodology
University of Colorado, Boulder
Boulder, CO

Howard Everson
Senior Principal Research Scientist
SRI International
New York, NY

Bonnie Hain
Chief Academic Officer
CenterPoint Education Solutions
Washington, DC

Kristen L. Huff
Vice President
Curriculum Associates
North Billerica, MA

Scott Marion
Executive Director
The National Center for the Improvement of Educational Assessment (NCIEA)
Dover, NH

Jennifer Randall
Associate Professor and Director of Evaluation for the Center for Educational Assessment, Education
University of Massachusetts, Amherst
Amherst, MA

Guillermo Solano-Flores, TAC chair
Professor, Education
Stanford University
Stanford, CA

Mark Wilson
Professor, Education
University of California, Berkeley
Berkeley, CA

Advisory Committee

Hyman Bass
Professor, Education and Mathematics
University of Michigan
Ann Arbor, MI

Rochelle Gutiérrez
Professor, Curriculum and Instruction and Latina/Latino Studies
University of Illinois at Urbana-Champaign
Champaign, IL

Danny Bernard Martin
Professor, Curriculum and Instruction
University of Illinois at Chicago
Chicago, IL

Jeremy Roschelle
Executive Director, Learning Sciences
Digital Promise
San Mateo, CA

Jon Star
Professor, Education
Harvard University
Cambridge, MA

WestEd Staff

Angela Bowzer
Mathematics Content Specialist
Senior Math Content Specialist
WestEd
Columbia, MO

Ann R. Edwards
Mathematics Content Specialist
Senior Research Associate
WestEd
Redwood City, CA

Matthew Gaertner
Measurement Specialist
Director of Research, Standards, Assessment, and Accountability Services
WestEd
Austin, TX

Shandy Hauk
Mathematics Content Specialist
Senior Research Associate, WestEd /
Associate Professor, Mathematics,
San Francisco State University
Redwood City, CA / San Francisco, CA

Kellie Kim
Process Manager
Senior Research Associate
WestEd
Washington, DC

Mark Loveland
Project Co-Director
Senior Research Associate
WestEd
Redwood City, CA

Steve Schneider
Project Director
Senior Program Director, Science Technology, Engineering, and Mathematics
WestEd
Redwood City, CA

Sarah Warner
Project Coordinator
Research Associate
WestEd
Nashville, TN

Kamilah Wilson
Administrative Assistant
WestEd
Washington, DC

Council of Chief State School Officers (CCSSO) Staff

Scott Norton
Deputy Executive Director, Programs

Fen Chou
Program Director, Standards, Assessment, and Accountability

National Assessment Governing Board Staff

Michelle Blair
Project Officer
Assistant Director for Assessment Development

Sharyn Rosenberg
Assistant Director for Psychometrics

REFERENCES

Abedi, J., & Herman, J. L. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record, 112*(3), 723–746.

Achieve. (2016). *Is NAEP math out of step with the states?* Washington, DC: Author. Retrieved from https://www.achieve.org/publications/naep-math-out-step-states

Adler, J. (1999). The dilemma of transparency: Seeing and seeing through talk in the mathematics classroom. *Journal for Research in Mathematics Education*, *30*(1), 47–64.

Aguirre, J. M., Turner, E. E., Bartell, T. G., Kalinec-Craig, C., Foote, M. Q., Roth McDuffie, A., & Drake, C. (2013). Making connections in practice: How prospective elementary teachers connect to children's mathematical thinking and community funds of knowledge in mathematics instruction. *Journal of Teacher Education*, *64*(2), 178–192.

Aguirre, J.M., Anhalt, C.O., Cortez, R., Turner, E.E., & Simic-Muller, K. (2019). Engaging teachers in the powerful combination of mathematical modeling and social justice: The Flint water task. *Mathematics Teacher Educator, 7(2)* 7–26.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: Author.

Andrews, J. J., Kerr, D., Mislevy, R. J., von Davier, A., Hao, J., & Liu, L. (2017). Modeling collaborative interaction patterns in a simulation-based task. *Journal of Educational Measurement*, *54*(1), 54–69.

Baghaei, P., & Dourakhshan, A. (2016). Properties of single-response and double-response multiple-choice grammar items. *International Journal of Language Testing*, *6*(1), 33–49.

Barbosa, J. C. (2006). Mathematical modeling in classroom: A socio-critical and discursive perspective, *ZDM*, 293–301.

Bell, C. V., & Pape, S. J. (2012). Scaffolding students' opportunities to learn mathematics through social interactions. *Mathematics Education Research Journal, 24*, 423–445.

Bishop, J. (2012). "She's always been the smart one. I've always been the dumb one": Identities in the mathematics classroom. *Journal for Research in Mathematics Education, 43*(1), 34–74.

Black, L. (2004). Teacher-pupil talk in whole-class discussions and processes of social positioning within the primary school classroom. *Language and Education, 18*(5), 347–360.

Burroughs, E., & Carlson, M. A. (2019). *Fostering empathy in mathematics through mathematical modeling*. Manuscript in preparation.

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school.* Portsmouth, NH: Heinemann.

Carraher, D. W., Martinez, M. V., & Schliemann, A. D. (2008). Early algebra and mathematical generalization. *ZDM*, *40*(1), 3–22.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, *64*(8), 723–733.

Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, *18*(1), 26–31.

Carter, P. L., & Welner, K. G. (2013). *Closing the opportunity gap: What America must do to give every child an even chance*. New York: Oxford University Press.

Chapin, S. H., O'Connor, C., O'Connor, M. C., & Anderson, N. C. (2009). *Classroom discussions: Using math talk to help students learn, Grades K–6*. Sausalito, CA: Math Solutions.

Civil, M. (2016). STEM learning research through a funds of knowledge lens. *Cultural Studies of Science Education*, *11*(1), 41–59. https://doi.org/10.1007/s11422-014-9648-2

Civil, M., & Planas, N. (2004). Participation in the mathematics classroom: Does every student have a voice? *For the Learning of Mathematics, 24*(1), 7–12.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*(2), 119–142.

Cohen, E. G., & Lotan, R. A. (1997). *Working for equity in heterogeneous classrooms: Sociological theory in practice. Sociology of Education Series*. New York: Teachers College Press.

Collins, A., & Stevens, A. L. (1983). A cognitive theory of interactive teaching. In C. M. Reigeluth (Ed.). *Instructional design theories and models: An overview*. Hillsdale, NJ: Erlbaum.

Coiro, J. Knobel, M., Lankshear, C., & Leu, D. (Eds.) (2008). *The handbook of research on new literacies*. Mahwah, NJ: Erlbaum.

Crabtree, Ashleigh R. (2016). *Psychometric properties of technology-enhanced item formats: An evaluation of construct validity and technical characteristics* (Doctoral dissertation). https://doi.org/10.17077/etd.922fbj4d

Crowe, C. E., & Ma, X. (2010). Profiling student use of calculators in the learning of high school mathematics. *Evaluation & Research in Education*, *23*(3), 171–190.

Cuoco A., Goldenberg E. & Mark J. (1996) Habits of mind: an organizing principle for mathematics curriculum. Journal of Mathematical Behavior 15, 375–402.

Daro, P., Hughes, G. B., & Stancavage, F. (2015). *Study of the Alignment of the 2015 NAEP Mathematics Items at Grades 4 and 8 to the Common Core State Standards (CCSS) for Mathematics.* Washington, DC: American Institutes for Research. Retrieved from https://www.air.org/sites/default/files/downloads/report/Study-of-Alignment-NAEP-Mathematics-Items-common-core-Nov-2015.pdf

de Freitas, E., & Sinclair, N. (2016). The cognitive labor of mathematics dis/ability: Neurocognitive approaches to number sense. *International Journal of Educational* Research, *79,* 222–230.

Dunham, P., & Hennessy, S. (2008). Equity and use of educational technology in mathematics. *Research on Technology and the Teaching and Learning of Mathematics*, *1*, 345–418.

Eisenhart, M., Borko, H., Underhill, R., Brown, C., Jones, D., & Agard, P. (1993). Conceptual knowledge falls through the cracks: Complexities of learning to teach mathematics for understanding. Journal for Research in Mathematics Education, 8–40.

Eklöf, H. (2010). Skill and will: Test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy & Practice, 17*(4), 345–356. DOI:10.1080/0969594X.2010.516569

Elliott, S. N., & Bartlett, B. J. (2016). Opportunity to learn. In P. Nathan (Ed.), *Oxford handbook of education online*. New York: Oxford University Press.

Ellis, A. B. (2007). Connections between generalizing and justifying: Students' reasoning with linear relationships. *Journal for Research in Mathematics Education*, *38*(3), 194–229.

Ellis, A., Bieda, K., & Knuth, E. (2012). *Essential understandings for proof and proving in 9–12 mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Engle, R. A., Langer-Osuna, J. M., & McKinney de Royston, M. (2014). Toward a model of influence in persuasive discussions: Negotiating quality, authority, privilege, and access within a student-led argument. *Journal of the Learning Sciences*, *23*(2), 245–268.

Enyedy, N. (2003). Knowledge construction and collective practice: At the intersection of learning, talk, and social configurations in a computer-mediated mathematics classroom. *Journal of the Learning Sciences, 12*(3), 361–407.

Enyedy, N., Rubel, L., Castellón, V., Mukhopadhyay, S., Esmonde, I., & Secada, W. (2008). Revoicing in a multilingual classroom. *Mathematical Thinking and Learning*, *10*(2), 134–162.

Ernest, P. (1998). *Social constructivism as a philosophy of mathematics.* New York: State University of New York Press.

Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., et al. (2017). *Collaborative problem solving: Considerations for the National Assessment of Educational Progress*. Washington, DC: National Center for Education Statistics

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *The High School Journal, 91*(1), 29–42.

Folk, C. L., Wolfe, J., Nobre, A. C., Heideman, S., Gibson, B. S., et al. (2015). *The Handbook of Attention*. Cambridge, MA: MIT Press.

Forman, E. A., Larreamendy-Joerns, J., Stein, M. K., & Brown, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. *Learning and Instruction*, *8*(6), 527–548.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK–12 curriculum framework*. Alexandria, VA: American Statistical Association. Retrieved from https://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf

Franklin, C. et al. (in press). *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK–12 curriculum framework* [Updated]. Alexandria, VA: American Statistical Association.

Garfunkel, S., & Montgomery, M. (Eds.). (2016). *GAIMME: Guidelines for Assessment and Instruction in Mathematical Modeling Education*. Philadelphia, PA: COMAP and SIAM. Retrieved from https://www.siam.org/Publications/Reports/Detail/guidelines-for-assessment-and-instruction-in-mathematical-modeling-education

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*(2), 121–151.

Geary, D. C., Liu, F., Chen, G. P., Saults, S. J., & Hoard, M. K. (1999). Contributions of computational fluency to cross-national differences in arithmetical reasoning abilities. *Journal of Educational Psychology, 91*(4), 716–719.

Gee, J. P. (1999). *An introduction to Discourse analysis: Theory and method*. New York: Routledge.

González, N., Moll, L. C., & Amanti, C. (Eds.). (2005). *Funds of knowledge: Theorizing practices in households, communities, and classrooms.* New York: Routledge.

Goos, M. (2004). Learning mathematics in a classroom community of inquiry. *Journal for Research in Mathematics Education, 35*(4), 258–291.

Goos, M., Galbraith, P., & Renshaw, P. (2002). Socially mediated metacognition: Creating collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics, 49*, 193–223.

Gutiérrez, R. (2013). The sociopolitical turn in mathematics education. *Journal for Research in Mathematics Education*, *44*(1), 37–68.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange*, *21*(1), 6–13.

Harel, G., & Sowder, L. (1998). Students’ proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput and E. Dubinsky (Eds.), *Research on Collegiate Mathematics Education* (Vol.3), (pp. 234–283). Providence, RI: American Mathematical Society.

Harel, G., & Tall, D. (1991). The general, the abstract, and the generic in advanced mathematics. *For the Learning of Mathematics*, *11*(1), 38–42.

Herbel-Eisenmann, B. A., & Cirillo, M. (Eds.). (2009). *Promoting purposeful discourse: Teacher research in mathematics classrooms*. Reston, VA: National Council of Teachers of Mathematics.

Herman, J. L., Klein, D. C., & Abedi, J. (2000). Assessing students' opportunity to learn: Teacher and student perspectives. *Educational Measurement: Issues and Practice*, *19*(4), 16–24.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum.

Hufferd-Ackles, K., Fuson, K., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education, 35*(2), 81–116.

Hughes, G. B., Daro, P., Holtzman, D., & Middleton, K. (2013). *A Study of the Alignment Between the NAEP Mathematics Framework and the Common Core State Standards for Mathematics*. Washington, DC: American Institutes for Research. Retrieved from http://www.air.org/sites/default/files/downloads/report/NVS_combined__study_1_NAEP_alignment_with_CCSS_0.pdf

Husén, T. (Ed.). (1967). *International study of achievement in mathematics: A comparison of twelve countries.* New York: John Wiley.

Hussain, M. A., Monaghan, J., & Threlfall, J. (2013). Teacher-student development in mathematics classrooms: Interrelated zones of free movement and promoted actions. *Educational Studies of Mathematics, 82*, 285–302.

International Association for the Evaluation of Educational Achievement. (2013). *TIMSS 2011 Assessment*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Chestnut Hill, MA and International Association for the Evaluation of Educational Achievement (IEA), IEA Secretariat, Amsterdam, the Netherlands.

Jodoin, M. G. (2003). Measurement efficiency of innovative item formats in computer-based testing. *Journal of Educational Measurement*, *40*(1), 1–15.

Johnston, W. T., Stephens, M., & Ratway, B. (2018). *Review of state curricular standards in mathematics: Study report.* Washington, DC: American Institutes for Research.

Knuth, E. J., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianous, M. L. Blanton and E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 153–170). New York: Routledge.

Küchemann D. & Hoyles C. (2009) From empirical to structural reasoning: tracking changes over time. In Teaching and Learning Proof Across the Grades Hillsdale (eds D.A. Stylianou, M.L. Blanton & E.J. Knuth), pp. 171–191. Routledge, London; New York.

Lambert, R., Tan, P., Hunt, J. H., & Candella, A. (2018). Re-humanizing the mathematics education of students with disabilities: Critical perspectives on research and practice. *Investigations in Mathematics Learning*, *10*(3), 129–132.

Lau, P., Singh, P., & Hwa, T. (2009). Constructing mathematics in an interactive classroom context. *Educational Studies in Mathematics, 72*, 307–324.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.

Lerman, S. (1996). Intersubjectivity in mathematics learning: A challenge to the radical constructivist paradigm? *Journal for Research in Mathematics Education*, *27*(2), 133–150.

Lerman, S. (2000). The social turn in mathematics education research. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19–44). Westport, CT: Ablex Publishing.

Lesh, R. Post, T. and Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of* *representation in the teaching and learning of mathematics* (pp. 33–40). Hillsdale, NJ: Lawrence Erlbaum.

LOCUS: Levels of Conceptual Understanding in Statistics. (2019). Item retrieved from https://locus.statisticseducation.org/professional-development/questions/by-grade/grade-6?page=&type=prodev_const_response_question

Martin, D. (2000). *Mathematics success and failure among African-American youth: The roles of sociohistorical context, community forces, school influence, and individual agency*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Martin, D. (2009). *Mathematics teaching, learning, and liberation in the lives of Black children*. New York: Routledge.

Martineau, J., Dadey, N., & Marion, S. (2018). *Literature review on developing or revising assessment frameworks to support a transition from paper-based to digitally based assessment*. Report prepared for the National Assessment Governing Board.

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6–20.

Moll, L. C., Amanti, C., Neff, D., & González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice, 31*(2), 132–141.

Moschkovich, J. (2007). Examining mathematical discourse practices. *For the Learning of Mathematics*, *27*(1), 24–30.

Moschkovich, J. (2008). "I went by twos, he went by one": Multiple interpretations of inscriptions as resources for mathematical discussions. *Journal of the Learning Sciences*, *17*(4), 551–587.

Mullis, I. V. S., & Martin, M. O. (2016). *TIMSS 2019 Assessment Framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, and International Association for the Evaluation of Educational Achievement.

The NAEP Law. (2017, July 28). Retrieved from https://www.nagb.gov/about-naep/the-naep-law.html

Nasir, N. S., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *Journal of the Learning Sciences, 17*(2), 143–179.

Nathan, M. J., Eilam, B., & Kim, S. (2007). To disagree, we must also agree: How intersubjectivity structures and perpetuates discourse in a mathematics classroom. *The Journal of the Learning Sciences*, *16*(4), 523–563.

National Assessment Governing Board. (n.d.). *The Nation's Report Card*. Homepage: https://www.nationsreportcard.gov

National Assessment Governing Board. (2002). *NAEP item development and review policy statement*. Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/policies/Item%20Development%20and%20Review.pdf

National Assessment Governing Board. (2007). *Assessment and item specifications for the NAEP 2009 Mathematics Assessment*. Washington, DC: Author

National Assessment Governing Board. (2014a). *NAEP testing and reporting on students with disabilities and English language learners*. Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/policies/naep_testandreport_studentswithdisabilities.pdf

National Assessment Governing Board. (2014b). *Science Framework for the 2015 National Assessment of Educational Progress.* Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/science/2015- science-framework.pdf

National Assessment Governing Board. (2017). *Mathematics Framework for the 2017 National Assessment of Educational Progress*. Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/mathematics/2017-math-framework.pdf

National Assessment Governing Board. (2018a). *Developing student achievement levels for the National Assessment of Educational Progress policy statement*. Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/policies/ALS-revised-policy-statement-11-17-18.pdf

National Assessment Governing Board. (2018b). *Framework development policy statement.* Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/policies/framework-development.pdf

National Center for Education Statistics [NCES]. (n.d.). NAEP Questions Tool. Homepage: https://nces.ed.gov/NationsReportCard/nqt

NCES. (2019). Trends in International Mathematics and Science Study. Homepage: https://nces.ed.gov/timss/

National Council of Teachers of Teachers of Mathematics [NCTM]. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.

NCTM. (1991). Professional standards for teaching mathematics. Reston, VA: Author.

NCTM. (2000). Principles and standards for school mathematics. Reston, VA: Author.

NCTM. (2014). Principals to actions: Ensuring mathematical success for all. Reston, VA: Author.

NCTM. (2018). Catalyzing change in high school mathematics: Initiating critical conversations. Reston, VA: Author.

National Governors Association Center for Best Practices, Council of Chief State School Officers [CCSS-M]. (2010). *Common Core State Standards for Mathematics*. Washington DC: Authors.

National Research Council [NRC]. (2001). *Adding it up: Helping children learn mathematics*. J. Kilpatrick, J. Swafford, & B. Findell (Eds.). Washington, DC: National Academy Press.

NRC. (2009). *Mathematics learning in early childhood: Paths toward excellence and equity.* Washington DC: National Academies Press.

Organisation for Economic Co-operation and Development [OECD]. (2006). PISA released items – Mathematics. Retrieved from https://www.oecd.org/pisa/38709418.pdf

OECD. (2017). *PISA 2015 Collaborative Problem-Solving Framework***.** Retrieved from https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf

OECD. (2018). PISA for Development Mathematics Framework, in *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science*. Paris: OECD Publishing, https://doi.org/10.1787/9789264305274-5-en.

OECD. (2019). *Programme for International Student Assessment.* Homepage: http://www.oecd.org/pisa/

Partnership for Assessment of Readiness for College and Careers [PARCC] (2015). Items and mathematics practice tests. Retrieved from https://parcc-assessment.org/released-items

Pellegrino, J. W., Jones, L. R., & Mitchell, K. J. (Eds.). (1999). *Grading the Nation's Report Card: Evaluating NAEP and transforming the assessment of educational progress.* Washington, DC: National Academy Press.

Radford, L. (2007). Iconicity and contraction: A semiotic investigation of forms of algebraic generalizations of patterns in different contexts. *ZDM*, *40*(1), 83–96.

Saxe, G. B. (1988). Candy selling and math learning. *Educational Researcher, 17*(6), 14–21.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic Press.

Schmidt, W.H., Burroughs, N.A., Zoido, P., & Houang, R.T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher, 44* (7), 371–386.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... & Chen, M. (2012). Early predictors of high school mathematics achievement. Psychological Science, 23(7), 691– 697.

Silicon Valley Mathematics Initiative. (2016). *Database of Tasks & POMs.*

Sireci, S. G., & Zenisky, A. L. (2006). Innovative item formats in computer-based testing: In pursuit of improved construct representation. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 329–347). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Smarter Balanced Assessment Consortium [SBAC] (2018). Sample Items. Retrieved from http://sampleitems.smarterbalanced.org

Solano-Flores, G. (2011). Assessing the cultural validity of assessment practices: An introduction. In M. del Rosario Basterra, E. Trumbull, and G. Solano-Flores (Eds.), *Cultural validity in assessment* (pp. 19–37). New York: Routledge.

Standards and Testing Agency. (2019). *Key Stage 2, Paper 3: Reasoning.* Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/804059/STA198218e_2019_ks2_mathematics_Paper3_reasoning.pdf

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology,* Vol. 34, (pp. 379–440). San Diego, CA: Academic Press.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, *10*(4), 313–340.

Tan, P., & Kastberg, S. (2017). Calling for research collaborations and the use of dis/ability studies in mathematics education. *Journal of Urban Mathematics Education*, *10*(2), 25–38.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., et al. (2012). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA teacher education and development study in mathematics (TEDS-M). Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Teasley, S.D. and Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie and S. J. Derry (Eds). *The Computer as a Cognitive Tool.* Hillsdale, NJ: Erlbaum. 229–258.

Thompson, S. J., Johnstone, C. J., Anderson, M. E., & Miller, N. A. (2005). *Considerations for the development and review of universally designed assessments* (Technical Report 42). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments, Synthesis Report 44*. Minneapolis, MN: National Center on Educational Outcomes.

TODOS and the National Council of Supervisors of Mathematics. (2016). *Mathematics education through the lens of social justice: Acknowledgment, actions, and accountability*. Joint position statement. Retrieved from https://www.todos-math.org/socialjustice

Tripathi, P. N. (2008). Developing mathematical understanding through multiple representations. *Mathematics Teaching in the Middle School 13*(8)*,* 438–445.

Truxaw, M. P., & DeFranco, T. C. (2008). Mapping mathematics classroom discourse and its implications for models of teaching. *Journal for Research in Mathematics Education*, *39*(5), 489–525.

U.S. Department of Labor, Office of Disability Employment Policy (2010). *Skills to pay the bills: Mastering soft Skills for workplace success*. Washington, DC: U.S. Department of Labor. Retrieved from http://www.dol.gov/odep/topics/youth/softskills/softskills.pdf

van Oers, B. (2001). Educational forms of initiation in mathematical culture. *Educational Studies in Mathematics, 46*, 59–85.

Voogt, J. & Knezek, G. (Eds.) (2008). *International handbook of information technology in primary and secondary education.* New York: Springer.

Wang, J. (1998). Opportunity to learn: The impacts and policy implications. *Educational Evaluation and Policy Analysis*, *20*(3), 137–156.

Warschauer, M. (2016). Addressing the social envelope: education and the digital divide. In C. Greenhow, J. Sonnevend, and C. Agur (Eds.), *Education and social media: Toward a digital future* (pp. 29–48). Cambridge, MA: MIT Press.

Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. *Review of Research in Education*, *34*(1), 179–225.

Wickstrom, M. H., & Aytes, T. (2018). Elementary modeling: Connecting counting with sharing. *Teaching Children Mathematics, 24*(5), 300–307.

Update: 2025 NAEP Reading Framework

Board Policy for Each Framework Update

In November 2019, the Assessment Development Committee (ADC) took stock of the success in implementing the Framework Development Policy. The ADC confirmed that one role of the Committee is to assure that the framework update process is carefully followed to produce a high quality framework for each NAEP assessment. To execute this responsibility, the ADC monitors framework processes via routine project updates and provides direction to the framework panels, as needed. This guidance is intended to assure compliance with the NAEP law, Governing Board policies, Department of Education and government-wide regulations, and requirements of the contracts used to implement the framework project.

As the framework panels engage deeply in the issues specific to the subject area, the Board must exercise policy oversight by considering a wider context. This includes consideration of the role and purpose of NAEP in informing the public about student achievement, the legislative parameters for NAEP, constraints of a large-scale assessment, technical assessment standards, and issues of burden and cost-effectiveness in designing the assessment. This wider context also includes the Board's priorities as articulated in the Governing Board's Strategic Vision. Hence, for each framework process, the Board must determine:

What direction is needed from the Governing Board?

The following list of critical questions is intended to support the ADC as it monitors framework update processes to assure compliance with the Governing Board's Framework Development Policy. The goal of each update is to produce a high-quality framework. Accordingly, key outcomes are also listed.

Process

The process must be comprehensive, inclusive, and deliberative. Ongoing process questions for the Committee's monitoring efforts for the 2025 NAEP Reading Framework include:

- Does the process engage a broad spectrum of stakeholders in developing recommendations for the knowledge and skills NAEP should assess?
- Is the process informed by a broad, balanced, and inclusive set of factors, delicately balancing current curricula and instruction, research, and the nation's future needs?
- Is the process being conducted in an environment that is open, balanced, and even-handed?
- Is the Development Panel considering all viewpoints raised and debating all pertinent issues?

Given the project milestones that have already passed, the following process questions have been addressed affirmatively for the 2025 NAEP Reading Framework:

- Does the Development Panel have a proportionally higher representation of content experts and educators (compared with the Visioning Panel)?

- Does the Development Panel's content expertise collectively address all grade levels designated for the assessment?
- Did the framework update project begin with an extensive review of the current framework?

See the attached Visioning and Development Panel member listings and biographies for additional details on the composition of the Development Panel for the 2025 NAEP Reading Framework.

Outcomes

In accordance with the Board's policy, the final framework must:

- Be inclusive of content valued by the public
- Reflect high aspirations
- Focus on important, measurable indicators
- Avoid endorsing or advocating a particular instructional approach
- Be clear and accessible to educators and the general public
- Define the construct(s) to be assessed and reported upon
- Articulate item formats, sample items, and sub-content weightings to demonstrate the construct is to be measured
- Describe how much of the content domain relates to the NAEP Basic, NAEP Proficient, and NAEP Advanced levels for each grade to be tested
- Align to widely accepted professional testing standards
- Support fair and accurate measurement of student academic achievement
- Support NAEP assessment items that will be secular, neutral, and non-ideological and free from racial, cultural, gender, or regional bias

Issues Under Discussion for the Reading Framework Update

In this session, the ADC will be briefed on the progress of the Development Panel's update of the 2025 NAEP Reading Framework (see attached quarterly update).

As noted in the Issues Review (attached) used to launch the Development Panel's deliberation, there were several major questions being tackled.

Issue 1. How should the texts and reading tasks used in NAEP be updated to reflect contemporary aspirations and expectations for reading?

Issue 2. How should NAEP integrate reading and writing while maintaining NAEP Reading and NAEP Writing assessments and reporting?

Issue 3. How should NAEP account for the interplay between knowledge and reading comprehension?

Issue 4. How should NAEP take better advantage of the affordances of digitally-based assessments?

Issue 5. How should NAEP modify the content and structure of the Reading assessment and the reporting of results in order to more equitably represent students' reading achievement?

Issue 6. What new theoretical and research-based understandings about reading comprehension and its assessment need to be reflected in the framework?

In response to these issues, the Panel is considering recommending the following changes:

a) Accounting for motivation, engagement, and social emotional learning as part of student achievement in reading, updating the definition of the construct of reading
b) Expanding the construct of reading to include using, applying, navigating, and selecting texts, as well as judging relevance and trustworthiness of sources
c) Providing choice in passage selection or task
d) Commissioning texts to augment sampling of authentic texts, as needed
e) Reporting new subscales, as part of achievement results
f) Reporting text and task variables, alongside student achievement results
g) Reporting motivation, interest, knowledge, effort, self-efficacy, metacognitive skills, and social emotional learning, alongside student achievement results

Some of these changes have policy implications. For example,

- the Board has not taken a formal position with respect to social emotional learning, and it is being considered as part of the construct and the contextual variables given the state of the field;
- relatedly, some states have raised concerns about NAEP questionnaire items that could be considered intrusive, and integrating certain questionnaire items into the assessment (rather than the separate questionnaires) could lead to state participation issues;
- the NAEP Technology and Engineering Literacy Assessment has a subdomain addressing Information and Communication Technology (which also includes navigating hypertext and judging trustworthiness, for example), and the ADC has previously discussed the importance of being intentional about the ways in which NAEP assessments' content relate to each other;
- providing choice in one assessment could lead to choice as a component of other subject-area assessments;
- the current NAEP Reading Framework includes a commitment to authentic texts, and using commissioned texts could prompt a variety of interpretations; and
- new subscales for the NAEP Reading Assessment could help to make NAEP results more actionable to NAEP audiences, but both intended and unintended uses must be considered.

After receiving an update on the project's progress, ADC members will be asked for their initial feedback on these potential policy implications.

WestEd
WestEd.org

Quarterly Report

2025 NAEP READING FRAMEWORK UPDATE

WestEd

February 2020

Contract #91995918C0001

Quarterly Progress Report

Project Overview

In September 2018, the Governing Board awarded a contract to WestEd to conduct an update of the NAEP Mathematics and Reading Assessment Frameworks, Assessment and Item Specifications, and Contextual Variables. Year 1 of the project was focused on the updating of the Mathematics Framework documents, with Year 2 focused on Reading. The goal of the Reading Framework project is to update the NAEP Reading Framework documents through the work of a 32-person Visioning Panel, a 17-person Development Panel, and an 8-person Technical Advisory Committee (TAC). This will be accomplished through an initial Visioning Panel meeting, five subsequent Development Panel meetings, conducting outreach efforts to gather public comment on draft versions of the documents, and production of a final updated Reading Assessment Framework, Assessment and Item Specifications, and Contextual Variables for Reading to submit to the Governing Board by October 2020.

The Reading Framework update is to be conducted using a combination of external experts and reading specialists within WestEd. WestEd's considerable experience with NAEP comes from having led previous NAEP framework projects (the 2009 NAEP Science Framework and the 2014 NAEP Technology and Engineering Literacy (TEL) Framework). To complete this work, WestEd is partnering with the Council of Chief State School Officers (CCSSO), which assists in compiling resources for the Framework panels and in securing feedback on the updated framework, assessment and item specifications, and contextual variables. Input into the framework document update also comes from project collaborators: the Literacy Research Association (LRA), the International Literacy Association (ILA), and the National Council of Teachers of English (NCTE).

Project Team

The Project Management Team consists of Steve Schneider, Mark Loveland, Cynthia Greenleaf, Matt Gaertner, and Kellie Kim. As project director, Steve Schneider provides day-to-day leadership, guidance, and liaising with the Governing Board. Dr. Schneider has over 40 years of science, mathematics, and technology education experience and led WestEd's three previous Framework development projects. Project co-director, Mark Loveland, and Reading Content Lead, Cynthia Greenleaf, have oversight for all programmatic activities. Dr. Loveland was project coordinator for the TEL Framework development project and project co-director for the Mathematics Framework update. P. David Pearson, Professor Emeritus and former Dean of the University of California, Berkeley Graduate School of Education serves as the Reading Panel Chair. Together, he and Dr. Greenleaf lead the Visioning and Development Panel activities. Measurement Lead, Dr. Gaertner coordinates the TAC. Dr. Kim serves as Process Manager, documenting all project activities. In addition to the project leaders, the broader project team includes two additional reading subject matter experts, three project coordinators, and research assistants.

Project Plan

The project plan describes WestEd's project management and coordination of panel and TAC activities to update the NAEP Reading Assessment Framework, Assessment and Item Specifications, and Contextual Variables. The bulk of the framework update work will be carried out by the Framework Visioning and Development Panels. Comprised of 32 individuals representing various stakeholder groups, the Framework Visioning Panel formulated guidelines for developing a recommended framework, based on the state of the field. Seventeen members of the Visioning Panel constitute the Framework Development Panel. The Development Panel is charged with developing the drafts of the three project documents and engaging in the detailed deliberations to determine how to reflect the Visioning Panel guidelines in an updated framework. Dates for the Visioning Panel meeting and the five Development Panel meetings have been finalized.

Preparatory work for the Framework Panel activities has been extensive. WestEd has prepared a Project Plan, which describes the process and schedule for updating the framework documents, and a project Design Document, which serves as the blueprint for the project processes, describing outcomes and metrics, and as the touchstone for quality assurance monitoring. Additionally, a Technical Advisory Committee comprised of eight technical experts responds to technical issues raised during panel deliberations.

Progress to Date

Preparatory Activities

WestEd has worked in consultation with Governing Board staff and Governing Board members to identify a broadly-representative final list of 32 members of the Visioning and Development Panels, along with an updated list of 8 technical experts specializing in educational measurement to comprise the Reading TAC. The work of the panels and TAC has been informed by a review of the issues and a compilation of resources. The Issues Review served as a springboard for discussion by the Framework panels and addressed specific issues that are likely to be engaged in the update process. The Resource Compilation has been a "living document," with additional resources added throughout the panel activities as they have been identified.

Panel Activities

Panel activities have been successfully conducted around the Visioning Panel meeting and the first three Development Panel meetings. The October 2019 Visioning Panel meeting and pre-meeting activities focused primarily on orienting panelists to the project and to the current state of reading education and assessment, followed by the generation of guidelines for the subsequent work to be done by the

Development Panel. The guidelines provided recommendations on: 1) expanding the construct of reading; 2) expanding the definition of text; 3) extending the range of comprehension tasks that require knowledge application, including writing from sources; 4) augmenting and expanding the cognitive targets and the approach to reporting performance on them; 5) expanding how meaning vocabulary is defined and measured; and 6) include, measure, and report on the role of engagement in reading performance.

The first Development Panel meeting, held in November 2019, used the Visioning Panel guidelines and the Issues Review to identify broad areas of the current Reading Framework that would serve as the starting point for the update process. Working in small groups in the first meeting and in between meetings, the Development Panel conducted a thorough examination of the current Framework and provided targeted recommendations for the update of the Framework and the Contextual Variables. The second Development Panel meeting, conducted in December 2020, focused on a multi-faceted approach to producing an updated model behind the 2025 NAEP Reading Framework, which identifies three key components: reader, text, and activity, along with an expanded definition of "reading." The third Development Panel meeting, convened in January 2020, took that model, came to consensus on a revised table of contents for the Framework, and pivoted the small groups' attention to either updating existing chapters or writing new chapters.

TAC Activities

The TAC has met on two occasions, in December 2019 to provide feedback on the Visioning Panel Guidelines and again in January 2020 to respond to questions from the first two Development Panel meetings. The first TAC meeting addressed issues related to specific feedback on proposed changes to the reading construct and the cognitive targets. TAC feedback on these issues were reported at the second Development Panel meeting.

The second TAC meeting focused on issues related to questions around the proposed Reading Framework chapter structure, new questions around cognitive targets, and student choice on the NAEP Reading Assessment. Responses were again reported back to the Development Panel at the next meeting.

Next Steps

Panel Activities

The fourth and penultimate face-to-face Developmental Panel meeting will be conducted on March 17-18, 2020 in Washington, DC. The focus of this meeting will be to reach consensus on critical decisions needed to complete draft versions of each Framework chapter and the ALDs, and then work to complete and review those pieces of the Framework. The Panel will also discuss the process, timelines, and assignments for engaging in outreach activities leading up to and during the public comment period, from June 1 to July 15. The Development Panel will meet virtually on a webinar in April to discuss final revisions of the Framework draft. The Visioning Panel will meet virtually via webinar in late May for an initial review of the Framework draft. Following the public comment period, the Development Panel will meet again virtually on a webinar in August to consider suggested revisions to the Framework draft. A final in-person meeting of the Development Panel is scheduled for September 2020 to resolve any remaining issues needed to finalize the Framework documents for submission to the Governing Board in mid-October.

Drafts of the Updated Framework

WestEd has developed a timeline and process for generating final drafts of the Framework documents for public review and comment, starting on June 1, 2020. Draft 1 will consist of individual chapters to be crafted by assigned panelists in the month following the third Development Panel meeting. An initial Draft 2 will combine the chapter drafts into a single Framework document prior to the fourth Development Panel meeting, followed by a series of internal (project staff) and external (Governing Board staff, TAC) reviews. Project staff will incorporate the feedback from these reviews in preparation for a Governing Board staff review. Based on feedback from this review, Draft 3 of the updated Framework and recommended Contextual Variables will be prepared in late May 2020 for public comment from June 1 to July 15. Draft 4 of the Framework will include revisions in response to feedback received during the public comment period including the Board's guidance on policy questions. This draft will be the focus of the final Development Panel meeting in September 2020. The final version of the 2025 NAEP Reading Framework will be submitted to the Governing Board in mid-October 2020.

Outreach

Planning for outreach activities is underway. Outreach activities are ongoing through mid-July 2020, serving multiple purposes: raise awareness of the Reading Framework update, engage with stakeholders, and gather external feedback and public comment on the draft Framework documents. Outreach aims to solicit substantive feedback in significant numbers from each of the stakeholder constituencies: teachers, curriculum specialists, content experts, assessment specialists, state administrators, local school administrators, policymakers, business representatives, parents, users of assessment data, researchers and technical experts, and members of the public.

Members of the Visioning and Development Panels solicit feedback from their member organizations through in-person and virtual meetings, while WestEd actively solicits feedback from additional stakeholder organizations through a variety of meeting formats and outreach activities. In all instances, groups follow procedures for securing input and ensuring representation of diverse views. WestEd staff will tabulate feedback and prepare summary documents for the Governing Board and the Development Panel. CCSSO will lead a series of outreach efforts to solicit feedback on draft versions of the Framework documents through its extensive membership network.

Collaborating representatives (e.g., organizational representatives on the Visioning Panel) will be given ample resources and support in order to host feedback forums with stakeholder constituents. Whenever possible, project staff will attend and support feedback forum hosts.

Final versions of the Framework documents will be developed for submission to the Governing Board on October 15, 2020. Along with the Framework documents, an annotated summary document will be developed describing the most significant changes to the Framework draft since the public comment draft (Draft 3). Any changes made to the Reading Framework documents will be carefully documented for transparency so the Governing Board can see all decision points with rationales.

Milestones

The major milestones of the project are summarized below.

Milestone	Dates
Project Kickoff Meetings	June – July 2019
Project Plan Development	June – September 2019
Design Document Development	July – September 2019
Identification of Visioning and Development Panelists and TAC Members	August – September 2019
Issues Paper and Resource Compilation Development	September – October 2019
Visioning Panel Meeting	October 2019
Development Panel Meetings	November 2019 – September 2020
Convene TAC	2-3 weeks after each panel meeting and prior to submission of draft framework documents
Draft Versions of Framework Documents	November 2019 – May 2020
Gather Public Comment	June 2020 – July 2020
Develop Final Versions of Framework Documents	July 2020 – October 2020
Submit Final Process Report	November 2020

Issues Review

2025 NAEP READING FRAMEWORK UPDATE

Cynthia L. Greenleaf, WestEd

Mira-Lisa Katz, WestEd, Prof. Emerita, Sonoma State University

Georgia Earnest García, Prof. Emerita, University of Illinois, Urbana-Champaign

Contract #91995918C0001

Introduction

This paper, which is intended to serve as a springboard for discussion in the 2025 NAEP Reading Update Project, outlines the issues that are likely to surface in the update of the National Assessment of Educational Progress (NAEP) Reading Framework. While this review is intentionally limited in scope, it is supplemented by an annotated Resource Compilation containing much more information pertinent to the task. Together, this Issues Review and the Resource Compilation provide the central library for the panel in the work of updating the NAEP Reading Framework.

Background and History

The purpose of NAEP is to provide fair and accurate measurement of student academic achievement. Also known as "The Nation's Report Card," NAEP collects and reports information on student performance based on samples of students in grades 4, 8, and 12. At grades 4 and 8, the NAEP Reading Assessment provides results for the nation, states, and 27 large districts that volunteer to participate in the Trial Urban District Assessment (TUDA). NAEP is critically important in the nation's evaluation of the condition and progress of education. The NAEP Reading Assessment has served as a measure of trends in academic achievement of U.S. elementary and secondary students. Each NAEP Assessment is guided by a framework and associated documents that specify the knowledge and skills to be tested at the 4th, 8th, and 12th grade, the format of the assessment, the definition of NAEP achievement levels, and the contextual variables for examining and presenting the results. As part of overseeing NAEP, the National Assessment Governing Board (the Board) oversees development of these frameworks. Each framework update takes into account factors such as state standards and assessments, international standards and assessments, exemplary research, and widely accepted professional standards (e.g., from AERA, APA, ILA, LRA, NCTE).

> THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP) is authorized by Congress and funded by the federal government. It is the only nationally representative and continuing assessment of what America's students know and can do. For more than 40 years, NAEP has been charged with collecting and reporting information on student achievement in mathematics, reading, science, U.S. history, writing, and other subjects. Originally, assessments were given to students at ages 9, 13, and 17. With the establishment of the National Assessment Governing Board in 1988, NAEP began assessing students at grades 4, 8, and 12. State-level reporting began with the 1990 reading assessment for grades 4 and 8. The NAEP Trial Urban District Assessment (TUDA) began in 2003, reporting results for 27 districts as of 2017.
>
> NAEP reports provide descriptive information about student performance in various subjects, including basic and higher order skills, and comparisons of performance by race/ethnicity, gender, type of community, and geographic region. They also show relationships between achievement and certain background variables, such as time spent on homework or educational level of parents.

The comprehensive process for the 2025 NAEP Reading Update project involves a Visioning Panel dedicated to crafting a set of guidelines for areas of focus in the update. These guidelines direct the Development Panel in its work to produce updates to assessment objectives, specifications, achievement level descriptions, and contextual variables. Each panel includes researchers, educators, business leaders, and policymakers.

The National Assessment Governing Board was established in 1988 to oversee the ongoing process of NAEP framework development, test specification, administration, and reporting. The first NAEP Reading Assessment framework was completed in 1990 and remained in place from 1992 to 2007. In 2007, the Board revisited the core of the framework to examine whether an update was needed. The new framework, developed through a deliberative process of panel and public engagement, was completed in 2004. Since the No Child Left Behind Act of 2001 and continuing with the Every Student Succeeds Act of 2015, NAEP has assessed reading in grades 4 and 8 every two years and in grade 12 every four years.

Distinct Features of NAEP Assessments

An assessment framework differs from a set of curricular standards. An assessment framework offers a blueprint specifying what to measure and how to measure accomplishment in a domain such as reading or mathematics. It makes clear statements about what should be assessed on NAEP, representing what students should know and be able to do at different stages of their development and about the types of texts and tasks (questions or items) best suited to measuring this knowing and doing. By contrast, standards documents offer more explicit and specific statements about the scope and sequence for instruction—what content is covered and when.

NAEP Reading Assessments are unique in that they are not reported by student, by school, or by district, with the exception of the 27 Trial Urban District Assessment (TUDA) districts. By law and by design, NAEP does not produce results for individual students or schools. Further, not all students in a district or school take the NAEP assessment, and no single student takes all of the assessment. Rather, a matrix sampling strategy ensures that enough students take each component of the test to provide a robust, composite portrait of reading attainment for the nation, for participating states, and for various demographic groups.

The Current NAEP Reading Framework and Assessment

As the background and history above suggests, NAEP frameworks and assessments are reviewed and updated periodically in accordance with the National Governing Board's Policy on Framework Development. This process ensures that NAEP keeps pace with what students are expected to know and be able to do and continues to play a critical role in the nation's evaluation of the condition and progress of education.

The current NAEP Reading Framework has been in place since the 2009 NAEP Reading Assessment, and empirical analyses have supported continued reporting of student achievement trends extending back to 1992 for grades 4, 8, and 12. Compared to the previous framework, the 2009 Reading Framework increased the emphasis on informational texts, redefined reading cognitive processes (behaviors and skills), introduced a new systematic assessment of vocabulary knowledge and added poetry to grade 4.

Because the nature of texts affects comprehension, and different text types must be read and interpreted using different skills, the NAEP Reading Assessment includes two distinct types of texts: literary and informational. Since 2009, the NAEP Reading Assessment also addresses vocabulary in a systematic way, to assess the interpretation of words in the context of a passage. The vocabulary questions function both as a measure of passage comprehension and as a test of knowledge of the meanings of specific words.

The cognitive targets developed for the current (2009) Reading Framework identify the mental processes or kinds of thinking that underlie reading comprehension: locate and recall, integrate and interpret, and critique and evaluate. These targets, along with vocabulary, have shaped the test specifications for the types of tasks students have been asked to carry out. The framework has also established achievement levels at each of the tested grades, specifying what knowledge and skills are needed to attain the NAEP Basic, NAEP Proficient, and NAEP Advanced achievement levels.

NAEP Reading Framework and Assessments Respond to a Changing Education Context

In response to the increasing role of digital technology in students' learning, the NAEP Reading Assessment has been administered as a digital-based assessment at grades 4 and 8 beginning in 2017. The 2017 assessments at grades 4 and 8 were largely comprised of previous paper-based assessment questions, adapted to fit a tablet screen and to address the same content and measurement targets. New types of items aligned to the framework were also developed to take advantage of the digital delivery system; additionally, studies were conducted to ensure that the digital delivery system was comparable to the print based assessment.

The digitally administered reading assessments at grades 4 and 8 were designed to continue reporting trends in student performance dating back to 1992. The first digital-based NAEP Reading assessment at grade 12 was conducted in 2019. Going forward all NAEP Reading Assessments will be digitally based.

> The Board has kept the NAEP Reading framework steady to support content stability and trend reporting during a time of sweeping changes in assessments across states. The 2017 assessment content was developed using the same frameworks used to develop the 2015 paper-based assessments and prior assessments since 2009.

Given the recent national focus on ensuring students' college and career readiness, the most recent NAEP Reading Framework update included a new purpose for NAEP 12th grade testing: reporting on how well prepared 12th grade students are for postsecondary training and education. This significant new purpose led to a shift for the 12th grade reading framework and some revisions in the reading processes assessed.

Moreover, changes are being considered for the administration of upcoming NAEP Assessments. NAEP Assessments are typically administered to each student in two 30-minute blocks focused on a single subject, such as Reading. Starting with the next schedule of assessments for NAEP 2021-2029, most students selected to participate in NAEP may take two-blocks of one subject, followed by a break, and then one-block of another subject. The schedule of assessments for 2021-2029 also indicates that the next NAEP Writing assessment will occur in 2029 and may reflect a new framework based on a framework review that will be conducted by the Governing Board's Assessment Development Committee.

Changes in the Context of Reading Research, Education and Assessment Relevant to the NAEP Reading Framework Update

As an independent, national monitor of student reading achievement, the NAEP Reading Framework must be both independent of particular curricula as well as inclusive of student learning across a range of curricula used in different states and school districts. The framework must also reflect best research and emerging themes in the field. Since the most recent revision of the NAEP Reading Framework in 2004, there have been shifts in both expectations for what students should know and be able to do and developments in assessment from consortia and states.

Most prominent in new standards is the call for readers to engage with complex text—not simply in terms of typical "text difficulty" but in terms of presentation of ideas that call for close, attentive reading and depth of understanding (Valencia, Wixson & Pearson, 2014). Writing from sources figures prominently in new standards, suggesting a role for considering how reading/writing relationships are handled in the framework (Lee, Hawley, Browder, Flowers & Wakeman, 2015; McDonald, Salomone, Gutierrez & Japtok, 2016; Mo & Troia, 2017; Peterson, 2017). New standards also uniformly emphasize the multimodal nature of reading, including using a variety of text types to conduct research, critique sources, and to communicate understanding through writing (Breakstone, McGrew, Smith, Ortega & Wineburg, 2018; Leu, Kinzer, Coiro, Castek & Henry, 2017; McGrew, Breakstone, Ortega, Smith & Wineburg, 2018).

As standards have been updated, a number of new reading assessments have been developed to assess them. PARCC (Partnership for Assessment of Readiness for College and Careers), Smarter Balanced, PISA (Programme for International Student Assessment), PIRLS (Progress in International Reading Literacy Study) and GISA (Global, Integrated Scenario-Based Assessment) (the last was developed under the Department of Education's Reading for Understanding Initiative) are examples of tests that made efforts to instantiate new standards. Unique features of this generation of new assessments include synthesis across multiple texts, technology enhanced items, items with multiple correct answers, and multimodal features.

A great deal of research attention over the past decade has focused on the nature of disciplinary texts and tasks that represent learning and understanding in disciplinary content areas (Goldman et al., 2016), along with the role of academic vocabulary in such literacy and learning (LaRusso, et al, 2016). Simultaneously, technological innovations have brought about changes in the format of texts as well as approaches to reading (Oranje, et al., 2015). Researchers are identifying the ways that online reading capability is both similar to, and distinct from, reading text printing on paper (Coiro, 2011; Coiro, Lankshear, Knobel & Leu, 2014; Singer & Alexander, 2017).

Additionally, over the past two decades, the population of students in U.S. schools has become increasingly diverse (Bryant, Triplett, Watson & Lewis, 2017). Students' reading proficiencies affect their economic and civic participation in life the nation (Business Roundtable, 2017; NCEE, 2013). At the same time, texts inevitably are cultural and political in nature, drawing on frames of reference that may not be universally shared (Lafontaine, Baye, Vieluf & Monseur, 2015; Wexler, 2018). Recent vocabulary studies demonstrate that readers draw on multiple dimensions of vocabulary knowledge, extending understandings of the role of vocabulary in meaning making (e.g. Larusso, et al., 2016). And new understandings of translanguaging have provided insight into how meaning making engages multiple linguistic and cultural processes for bilingual and biliterate readers (Pacheco & Miller, 2016; Pacheco & Smith, 2015).

Finally, affective and non-cognitive dimensions of reading and learning influence student performance on assessment tasks. Student interest and motivation are known to affect reading performance (Guthrie, Klauda & Ho, 2013), along with students' purposes for reading (Kendeou, Van den Broek, Helder & Karlsson, 2014; Larusso, et al., 2016). Recent work on socioemotional factors such as self-efficacy, growth mindsets, metacognition, and self-regulation impacting performance demonstrate that these factors may also be relevant and important to measure (Dweck & Molden, 2005; Farrington, et al., 2012; Hall, 2016; Taylor, Oberle, Durlak & Weissberg, 2017).

Given these advances in the field, in updating the NAEP Reading Framework for 2025 and beyond, it will be important to consider how the NAEP Reading Framework and test specifications present text types, topics, and tasks in light of changed expectations for student reading, new research, technological advances, and differences in students' backgrounds, experiences, motivations, and interests.

Updating the NAEP Reading Framework for the 2025 NAEP Reading Assessment

In preparation for updating the NAEP Reading Framework, the Governing Board commissioned white papers from content experts to inform deliberations and to shape the Governing Board's charge to the panels that will be convened as part of the framework update process.

Expert Reviews of the Current NAEP Reading Framework. Experts examined the current NAEP Reading Framework and provided reviews in response to the questions, "Does the NAEP Reading Framework need to be revised? If so, why and how?" (Afflerbach, Allen, Alexander, Duke, Hoffman, McKeown, Wealdon).

The Board's Assessment Development Committee led a review of the current NAEP Reading Framework in Spring 2018. They solicited expert commentary to determine next steps for the NAEP Reading Framework. They recruited five experts in reading to review the 2017 NAEP Reading Framework and provide recommendations regarding revisions to reflect current research and knowledge in reading comprehension. These experts presented their recommendations in a panel discussion hosted by the Assessment Development Committee on March 2, 2018. The Board also invited papers from experts who drafted the current NAEP Reading Framework and then worked with item development for the NAEP Reading Assessment, to gather additional insights based on this more in-depth engagement with the assessment. The key documents are included, in full, in the Resource Compilation and have informed the issues put forward in this Issues Review.

In addition, in September of 2019 the Governing Board convened state, district, academia, policy, and assessment experts to obtain guidance on how states' integrated approaches to assessing reading and writing might inform NAEP frameworks and assessments. Experts shared trends in district, state, and international approaches to integrated assessment and reporting of reading and writing achievement and the advantages and disadvantages of different approaches. They were asked to offer guidance on what approaches the Governing Board might consider within existing constraints when revising the NAEP Reading and Writing Frameworks to blend and/or coordinate the two assessments. The expert panel deliberations and guidance have informed the issues articulated in this *Issues Review*.

Expert Panel Meeting on English Language Arts Assessment. Experts were asked to offer guidance in response to the questions, "What approach to integrating the reading and writing assessments is most appropriate for NAEP, given NAEP's goals and related program legislation? What key issues should the Governing Board consider with the goal to integrate the NAEP Reading and Writing Frameworks while maintaining separate Reading and Writing assessments and scores?"

The Charge to the Visioning Panel

The Governing Board's newly revised Framework Policy calls for a Visioning Panel, which establishes broad guidelines for the modification of a framework that it has been asked to review. A subset of the members of that panel will participate in a Development Panel, which will implement the changes. In addition to developing a written set of guidelines for revising the NAEP Reading Framework, the Visioning Panel is also asked to provide guidelines for revision of the Specifications document that describe how the NAEP assessment items are to be developed and the reading-specific contextual questions that are asked of students, teachers, and school administrators.

The National Assessment Governing Board Charge to the Visioning Panel For the2025 National Assessment of Educational Progress (NAEP) Reading Framework

The Visioning and Development Panels will recommend to the Board necessary changes in the NAEP Reading Framework at grades 4, 8, and 12 that maximize the value of NAEP to the nation; and the Panels are also tasked with considering opportunities to extend the depth of measurement and reporting given the affordances of digital based assessment. The update process shall result in three documents: a recommended framework, assessment and item specifications, and recommendations for contextual variables that relate to student achievement in reading.

The Issues

Given this history and the current policy and assessment landscape operating within and across states in the U.S., several issues are critical for the Visioning Panel to address.

Issue 1. How should the texts and reading tasks used in NAEP be updated to reflect contemporary aspirations and expectations for reading?

Issue 2. How should NAEP integrate reading and writing while maintaining NAEP Reading and NAEP Writing assessments and reporting?

Issue 3. How should NAEP account for the interplay between knowledge and reading comprehension?

Issue 4. How should NAEP take better advantage of the affordances of digitally-based assessments?

Issue 5. How should NAEP modify the content and structure of the Reading assessment and the reporting of results in order to more equitably represent students' reading achievement?

Issue 6. What new theoretical and research-based understandings about reading comprehension and its assessment need to be reflected in the framework?

We regard this list as provisional and expect the Visioning Panel to add others as relevant and appropriate.

Issue 1. How should the texts and reading tasks used in NAEP be updated to reflect contemporary aspirations and expectations for reading?

A key issue for the Visioning Panel will be to consider where and to what extent the current NAEP reading framework assesses what is expected of students in current instructional standards and in the broader contemporary context. Students are increasingly expected to read and integrate insights and information across multiple texts presented in varied modalities, both in print and online, in order to construct explanations and arguments and build and communicate knowledge. The proliferation of information sources requires students to exercise critical judgment about source relevance, trust-worthiness and perspective. Similarly, in reading literature, students are expected to analyze and appreciate how authors use literary devices and elements of craft to achieve literary goals. National and state standards and assessments, international frameworks and assessments, and college and career standards press for student engagement with complex texts and tasks across academic disciplines. How should the texts, tasks, assessment objectives (including cognitive targets), and specifications in NAEP reading assessments be updated to keep pace with these developments?

Issue 2. How should NAEP integrate reading and writing while maintaining NAEP Reading and NAEP Writing assessments and reporting?

Since the last update of the NAEP Reading Framework, new standards in English Language Arts and new college and career readiness standards have moved toward the integration of reading and writing, both in curriculum and assessment. Writing with sources figures prominently in new standards, suggesting a role for considering how reading/writing relationships are handled in the NAEP Reading Framework. Currently separate Reading and Writing assessments (and score reporting) are legislatively mandated for NAEP; by contrast, states have adopted standards that support increasingly integrated approaches to the teaching and assessment of reading and writing. Further, states are required to participate in the NAEP Reading Assessment at grades 4 and 8 through the Every Student Succeeds Act, while state participation in the NAEP Writing Assessment is voluntary. How should this state-level context inform NAEP frameworks and assessments? This also relates to the definition of reading in the NAEP Reading Framework. For example, experts across English language arts were recently convened by the Governing Board, and they suggested that writing to sources may be an appropriate aspect of a reading construct that includes application of what has been read. When the NAEP Writing Framework is next updated, these experts noted that writing with sources could be addressed in the context of a writing construct that also includes writing without sources. The Visioning Panel is asked to consider whether this approach to integrating the NAEP Reading and Writing assessments is appropriate, while addressing NAEP's legislative mandates and constraints.

Issue 3. How should NAEP account for the interplay between knowledge and reading comprehension?

Understanding and accounting for the role of knowledge in reading comprehension has long plagued the field. In the past, knowledge was understood to fuel comprehension. More recently, knowledge is recognized as an outcome of comprehension as well. Currently, NAEP does not address the knowledge-comprehension relationship directly. Instead it "accounts for" knowledge by sampling passages across a wide range of topics in both literary and informational genres. As several experts have noted, the current approach is inadequate to ensure equitable assessment of the variations in knowledge, experience, and abilities that students bring to the task. Importantly, the background knowledge demands of reading include not only familiarity with the content or topics of texts, but also students' prior experience with particular types of texts, genres, and reading tasks. The Visioning Panel is asked to consider new approaches to level the background knowledge playing field or at the very least, to account for differential knowledge among students taking the assessment. Such approaches might include providing necessary background knowledge prior to a reading passage (for example through a video or text preface), building a knowledge onramp across multiple texts, providing feedback after each item to ensure that all students approach subsequent tasks with comparable knowledge resources, and/or measuring knowledge inputs and outcomes as part of the assessment.

Issue 4. How should NAEP take better advantage of the affordances of digitally-based assessments?

Several expert panelists noted that additional skills are required in order to read successfully in digital environments. Moreover, digitally based assessments offer new possibilities for the range of texts and the types of tasks used in reading assessment. Scenario-based assessments present purposeful tasks for student engagement in reading across multiple texts. New developments also include building avatar-enriched social contexts for reading as well as novel and more dynamic response formats. The Visioning Panel will be provided with a presentation by NCES of digitally-based assessments that have already been developed for the most recent NAEP reading administrations These examples illustrate the affordances of computer administration not only to present digital texts and graphic informational displays, but also to provide a rich context for purposeful reading and meaning making, and to engage and maintain students' motivation. The Visioning Panel is asked to consider how to update the framework to reflect the NAEP assessment as it is currently operationalized as well as how to exploit new opportunities offered by digital innovation when revising the assessment objectives, specifications document, and context questions.

Issue 5. How should NAEP modify the content and structure of the Reading assessment and the reporting of results in order to more equitably represent students' reading achievement?

Equity is a concern with any assessment, and arguably greater in the case of NAEP, because it is administered broadly across many different populations. Visioning Panel members are asked to bring their considerable backgrounds, experiences, and wisdom to the challenge of developing guidelines for how the framework, specifications document, and surveys for students, teachers, and schools can be crafted to make the NAEP reading assessment as fair as possible to all populations. Specifically, panelists are asked to consider how NAEP can better acknowledge students' primary languages, cultural resources, and learning needs consistent with NAEP's definition of reading. Additionally, how can NAEP better measure and report students' opportunities to learn, as well as their motivation and engagement with assessment texts and tasks they encounter on the NAEP assessment? Efforts such as these may lead to a more equitable assessment for all children. Among the ideas offered in the resource compilation are those associated with cultural validity in assessment (Solano-Flores, 2011) and the assessment of English language learners (Pitoniak, Young, Martiniello, King, Buteux & Ginsburgh, 2009).

Issue 6. What new theoretical and research-based understandings about reading comprehension and its assessment need to be reflected in the framework?

The past decade has brought considerable change to our understanding of the nature, teaching, and assessment of reading comprehension. These changes include new understandings of the roles that key factors play in shaping and explaining reading comprehension; these include text, purpose and task, knowledge, and vocabulary. New forms of texts, multimodality, and multi-text comprehension create new targets for instruction and therefore assessment. Learning in academic disciplines now understood to require reading and learning with discipline-specific texts of varied genres to build valued types of knowledge specific to these fields. Thus, students read across and integrate information from multiple sources as they comprehend, critique, and construct arguments and explanations about the ideas and issues they encounter while reading. They engage in reasoning processes that reflect discipline-specific ways of thinking and building knowledge. Finally, new developments in the understanding of the nature of vocabulary provide an opportunity for NAEP to reshape the ways it assesses vocabulary in the Reading assessment. As several expert reviewers have noted, the current framework is not informed by these recent developments. In the process of grappling with these new developments in theory and research, the Visioning Panel will be required to address two fundamental but vexing questions: What is reading? What is text? The Visioning Panel is asked to develop guidelines that reflect our current understanding of reading comprehension and in the process address these fundamental questions.

Conclusion

While no change in testing purpose is proposed for this update, the update needs to anticipate the nation's future needs (e.g., for the workplace and economic competitiveness, for civic participation, and for supporting individual aspirations) and the associated educational aspirations that are inscribed in desired levels of achievement we set for our students and our nation. This includes anticipating future content shifts by noting how states are adopting and adapting their standards.

Most of the questions in this Issues Review have no easy answers, but the collective knowledge and experience of panel members provide the nation in general, and the educational system in particular, an opportunity to guide the NAEP Framework revision process in ways that will allow NAEP to address the needs of future generations to become skillful, thoughtful and critical readers. Over the years, NAEP assessment frameworks have provided a valuable resource to state and district educators in developing their content standards. We should expect no less of this update for the 2025 NAEP Reading Assessment.

Panelists face the challenging task of making choices that will shape the rich array of texts, tasks, processes, and abilities that define reading comprehension for the nation. The choices will be influenced by the experience, aspirations, and knowledge of what reading education should be. The panel has an opportunity to make important recommendations about what is reported in the Nations Report Card and to recommend that the Governing Board authorize special studies for questions that can only be resolved with additional empirical evidence. The best way to achieve these goals is to share ideas and aspirations with one another openly, to challenge and discuss these thoughts with one another, and—in the end—to collaboratively weave a tapestry of consensus for recommendations to the Governing Board for the next NAEP Reading Framework.

References

Afflerbach, P. (2018). Recommendations for revision of the Reading Framework for the National Assessment of Educational Progress. In *2018 Expert Panel Discussion: NAEP Assessment of Reading*. Washington, D.C.: National Assessment Governing Board.

Alexander, F. (2018). The Reading Framework for the National Assessment of Educational Progress: Preparing for another decade of reading assessment. In *2018 Expert Panel Discussion: NAEP Assessment of Reading*. Washington, D.C.: National Assessment Governing Board.

Allen, D. (2018). The Reading Framework for the National Assessment of Educational Progress: Recommendations. In *2018 Expert Panel Discussion: NAEP Assessment of Reading*. Washington, D.C.: National Assessment Governing Board.

Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan 99*(6), 27-32.

Bryant, A. C., Triplett, N. P., Watson, M. J., & Lewis, C. W. (2017). The browning of American public schools: Evidence of increasing racial diversity and the implications for policy, practice, and student outcomes. *The Urban Review*, 49(2), 263-278.

Business Roundtable. (2017). *Work in progress: How CEOs are helping close America's skills gaps*. Retrieved May 2018, from https://www.businessroundtable.org/skills

Coiro, J. (2011). Predicting reading comprehension on the internet: Contributions of offline reading skills, online reading skills, and prior knowledge. *Journal of Literacy Research*, 43(4), 352–392.

Coiro, J., Knobel, M., Lankshear, C., & Leu, D. J. (2014). *Handbook of research on new literacies*. Routledge.

Duke, N.K. (2018). Recommendations for revision of the Reading Framework for the National Assessment of Educational Progress. In *2018 Expert Panel Discussion: NAEP Assessment of Reading*. Washington, D.C.: National Assessment Governing Board.

Dweck, C., & Molden, D. (2005). Self-theories: Their impact on competence motivation and acquisition. In A. Elliot, and C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 122-140). New York: Guilford Press.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance*. Chicago, IL: Consortium on Chicago School Research.

Goldman, S., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., Lee, C.D., Shanahan, C., & Project READI. (2016). Disciplinary literacies and learning to read for understanding: A conceptual framework for disciplinary literacy. *Educational Psychologist*, 51(2), 219-246. http://dx.doi.org/10.1080/00461520.2016.1168741.

Guthrie, J. T., Klauda, S. L., & Ho, A. N. (2013). Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly*, *48*(1), 9-26

Hall, L. A. (2016). The role of identity in reading comprehension development. *Reading & Writing Quarterly*, *32*(1), 56-80.

Heid, M. K. (2018). Updating the NAEP Mathematics Framework. In *2018 Expert Panel Discussion: NAEP Assessment of Mathematics*. Washington, D.C.: National Assessment Governing Board.

Hoffman, J. (2018). The Reading Framework for the National Assessment of Educational Progress: Challenges and opportunities moving forward. In *2018 Expert Panel Discussion: NAEP Assessment of Reading*. Washington, D.C.: National Assessment Governing Board.

Kendeou, P., Van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, *29*(1), 10-16.

Lafontaine, D., Baye, A., Vieluf, S., & Monseur, C. (2015). Equity in opportunity-to-learn and achievement in reading: A secondary analysis of PISA 2009 data. *Studies in Educational Evaluation*, 47, 1-11.

LaRusso, M., Kim, H. Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., et al. (2016). Contributions of academic language, perspective taking, and complex reasoning to deep reading comprehension. *Journal of Research on Educational Effectiveness*, 9(2), 201-222.

LaRusso, M., Kim, H. Y., Selman, R., Uccelli, P., Dawson, T., Jones, S., et al. (2016). Contributions of academic language, perspective taking, and complex reasoning to deep reading comprehension. *Journal of Research on Educational Effectiveness*, 9(2), 201-222.

Lee, A., Hawley, K., Browder, D. M., Flowers, C., & Wakeman, S. (2016). Teaching writing in response to text to students with developmental disabilities who participate in alternate assessments. *Education and Training in Autism and Developmental Disabilities*, 51(3), 238-251.

Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2017). New literacies: A dual-level theory of the changing nature of literacy, instruction, and assessment. *Journal of Education*, 197(2), 1-18.

McDonald, S., Salomone, W., Gutierrez, S., & Japtok, M. (2016). *The Writer's Response: A Reading-based Approach to Writing*. Boston, MA: Cengage Learning.

Mcgrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of Civic Online Reasoning. *Theory & Research in Social Education*, 46(2), 165-193.

McKeown, M. (2018). Considerations for revisiting the NAEP Reading Framework. In *2018 Expert Panel Discussion: NAEP Assessment of Reading*. Washington, D.C.: National Assessment Governing Board.

Mo, Y. & Troia, G. A. (2017). Similarities and differences in constructs represented by U.S. States' middle school writing tests and the 2007 National Assessment of Educational Progress writing assessment. *Assessing Writing*, 33, 48-67.

National Center for Education and the Economy [NCEE]. (2013). *What does it really mean to be college and workforce ready? The English literacy required of first year community college students*. Washington, DC: NCEE. Retrieved June 20, 2018 from http://ncee.org/wp-content/uploads/2013/05/NCEE _EnglishReport_May2013.pdf.

Oranje, A., Moran, R., Campbell, J., Dresher, A., Persky, H., Dion, G., and Scalise, K. (2015). *NAEP's Transition to Digitally Based Assessment*. As cited in Behuniak (2015).

Pacheco, M. B., & Miller, M. E. (2016). Making meaning through translanguaging in the literacy classroom. *The Reading Teacher*, *69*(5), 533-537.

Pacheco, M. B., & Smith, B. E. (2015). Across languages, modes, and identities: Bilingual adolescents' multimodal codemeshing in the literacy classroom. *Bilingual Research Journal*, *38*(3), 292-312.

Peterson, D. S. (2017). Engaging elementary students in higher order talk and writing about text. *Journal of Early Childhood Literacy*, http://journals.sagepub.com/doi/full/10.1177/1468798417690918

Pitoniak, M. J., Young, J. W., Martiniello, M., King, T. C., Buteux, A., & Ginsburgh, M. (2009). Guidelines for the assessment of English-language learners. Educational Testing Service (ETS). Retrieved from https://www.ets.org/s/about/pdf/ell_guidelines.pdf.

Singer, L. M., & Alexander, P. A. (2017). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. *The Journal of Experimental Education*, 85(1), 155-172.

Solano-Flores, G. (2011). Assessing the cultural validity of assessment practices: An introduction. In M. del Rosario Basterra, E. Trumbull, and G. Solano-Flores (Eds.), *Cultural validity in assessment* (pp. 19-37). New York: Routledge.

Taylor, R. D., Oberle, E., Durlak, J. A. & Weissberg, R. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, *88*(4), 1156–1171.

Valencia, S. W., Wixson, K. K., & Pearson, P. D. (2014). Putting text complexity in context: Refocusing on comprehension of complex text. *Elementary School Journal*, 115(2), 270-289.

Wealdon, W. (2018). The Reading Framework for the National Assessment of Educational Progress Feedback. In *2018 Expert Panel Discussion: NAEP Assessment of Reading*. Washington, D.C.: National Assessment Governing Board.

Wexler, N. (April 13, 2018). Why American students haven't gotten better at reading in 20 years. *The Atlantic.* Downloaded from https://www.theatlantic.com/education/archive/2018/04/-american-students-reading/557915/

Reading Visioning Panelists by Stakeholder Group and Issue Attachment B

Yellow highlight indicates a Development Panelist

	Developmental Trends	*English Learners*	*Equity & Special Populations*	*Special Education*	*Socioemotional Factors*	*Technology*	*Generalist*
Academic	Kathleen Hinchman (Syracuse University)	Mariana Pacheco (University of Wisconsin, Madison)	Christy Howard (East Carolina University)			Julie Coiro (University of Rhode Island) Byeong-Young Cho (University of Pittsburgh)	
Assessment Specialist	Peter Afflerbach (University of Maryland) Panayiota Kendeou (University of Minnesota) Victoria Young (Texas Education Agency)	Paola Uccelli (Harvard University)					Bonnie Hain (CenterPoint Education Solutions) Jim Patterson (College Board)
Business							Minerva Anaya-St John (A-SJ Properties, Inc.) Robert Rothman (National Center on Education and the Economy)
Content Specialist	Gina Cervetti (University of Michigan)		Carol Lee (Northwestern University)				Nancy Brynelson (California State University) Sue Pimentel (Achieve the Core)
Educational Organization			Karen Malone (Bureau of Indian Education, Navajo District) Allison Skerrett (University of Texas at Austin; LRA)	Carol Connor (University of California, Irvine; SSSR)		Elena Forzani (Boston University; ILA) Emily Kirkpatrick (NCTE)	Cindy Parker (CCSSO)
Research					John Guthrie (University of Maryland)		
State/District ELA/Literacy Director/Superintendent							Robin Hall (CGCS) Jinghong Cai (NSBA)
Teacher/Principal	Sarah Aguirre (Northside ISD, TX)				Carolyn Aguirre (Cesar Chavez MS, CA) Alicia Ross (Blue Ridge HS, PA)		Eric Turman (Reading HS, PA) Paul Wenger (NAESP) Josephine Franklin (NASSP)

Reading Visioning Panelists by Stakeholder Group and Gender
Yellow highlight indicates a Development Panelist

Attachment B

	Female	*Male*
Academic	Julie Coiro Kathleen Hinchman Christy Howard Mariana Pacheco	Byeong-Young Cho
Assessment Specialist	Bonnie Hain Panayiota Kendeou Paola Uccelli Victoria Young	Peter Afflerbach Jim Patterson
Business	Minerva Anaya-St John	Robert Rothman
Content Specialist	Nancy Brynelson Gina Cervetti Carol Lee Sue Pimentel	
Educational Organization	Carol Connor Elena Forzani Emily Kirkpatrick Karen Malone Cindy Parker Allison Skerrett	
Research		John Guthrie
State/District ELA/Literacy Director/Superintendent	Jinghong Cai Robin Hall	
Teacher/Principal	Carolyn Aguirre Sarah Aguirre Josephine Franklin Alicia Ross	Eric Turman Paul Wenger

Reading Visioning Panelists by Stakeholder Group and Geographic Region

Attachment B

Yellow highlight indicates a Development Panelist

	West	*Southwest*	*Midwest*	*Northeast*	*Southeast*	*Noncontiguous*
Academic			Mariana Pacheco	Julie Coiro Kathleen Hinchman Byeong-Young Cho	Christy Howard	
Assessment Specialist		Victoria Young	Panayiota Kendeou Jim Patterson	Paola Uccelli Peter Afflerbach	Bonnie Hain	
Business		Minerva Anaya-St John		Robert Rothman		
Content Specialist	Nancy Brynelson		Gina Cervetti Carol Lee		Sue Pimentel	
Educational Organization	Karen Malone	Carol Connor Allison Skerrett	Emily Kirkpatrick	Elena Forzani	Cindy Parker	
Research					John Guthrie	
State/District ELA/Literacy Director/Superintendent				Jinghong Cai	Robin Hall	
Teacher/Principal	Carolyn Aguirre	Sarah Aguirre	Paul Wenger	Alicia Ross Eric Turman	Josephine Franklin	

Reading Visioning Panelists by Stakeholder Group and Main Grade Band
Yellow highlight indicates a Development Panelist

	Elementary	*Middle*	*High*	*PreK-12*	*Post-Secondary*
Academic	Mariana Pacheco	Byeong-Young Cho		Julie Coiro Kathleen Hinchman Christy Howard	
Assessment Specialist	Panayiota Kendeou	Paola Uccelli	Jim Patterson	Victoria Young Peter Afflerbach Bonnie Hain	
Business				Minerva Anaya-St John Robert Rothman	
Content Specialist	Gina Cervetti			Nancy Brynelson Carol Lee Sue Pimentel	
Educational Organization	Carol Connor	Elena Forzani	Karen Malone	Allison Skerrett Emily Kirkpatrick Cindy Parker	
Research		John Guthrie			
State/District ELA/Literacy Director/Superintendent	Robin Hall			Jinghong Cai	
Teacher/Principal	Sarah Aguirre Paul Wenger	Carolyn Aguirre	Alicia Ross Eric Turman	Josephine Franklin	

Bonnie Hain and Nancy Brynelson also have post-secondary experience

Reading Visioning Panelists by Stakeholder Group and Race/Ethnicity
Yellow highlight indicates a Development Panelist

Attachment B

	Asian	*Black or African American*	*Hispanic or Latino*	*Native American*	*White*
Academic	Byeong-Young Cho	Christy Howard	Mariana Pacheco		Julie Coiro Kathleen Hinchman
Assessment Specialist			Paola Uccelli		Peter Afflerbach Bonnie Hain Panayiota Kendeou Jim Patterson Victoria Young
Business					Minerva Anaya-St John Robert Rothman
Content Specialist		Carol Lee			Nancy Brynelson Gina Cervetti Sue Pimentel
Educational Organization		Allison Skerrett		Karen Malone	Carol Connor Elena Forzani Emily Kirkpatrick Cindy Parker
Research					John Guthrie
State/District ELA/Literacy Director/Superintendent	Jinghong Cai	Robin Hall			
Teacher/Principal		Josephine Franklin Eric Turman			Carolyn Aguirre Sarah Aguirre Alicia Ross Paul Wenger

Reading Visioning Panelists by Stakeholder Group and Teaching Experience

Attachment B

Yellow highlight indicates a Development Panelist

	Yes	*No*
Academic	Julie Coiro Kathleen Hinchman Christy Howard Mariana Pacheco	Byeong-Young Cho
Assessment Specialist	Peter Afflerbach Bonnie Hain Jim Patterson	Panayiota Kendeou Paola Uccelli Victoria Young
Business		Minerva Anaya-St John Robert Rothman
Content Specialist	Nancy Brynelson Carol Lee Sue Pimentel	Gina Cervetti
Educational Organization	Carol Connor Elena Forzani Karen Malone Cindy Parker Allison Skerrett	Emily Kirkpatrick
Research		John Guthrie
State/District ELA/Literacy Director/Superintendent	Jinghong Cai Robin Hall	
Teacher/Principal	Carolyn Aguirre Sarah Aguirre Josephine Franklin Alicia Ross Eric Turman Paul Wenger	

Reading Development Panelists by Stakeholder Group and Age Range

Attachment B

Yellow highlight indicates a Development Panelist

	20-29	*30-39*	*40-49*	*50-59*	*60-69*	*70-79*
Academic			Byeong-Young Cho Christy Howard Mariana Pacheco	Julie Coiro	Kathleen Hinchman	
Assessment Specialist			Panayiota Kendeou Paola Uccelli	Bonnie Hain Jim Patterson	Peter Afflerbach Victoria Young	
Business					Minerva Anaya-St John Robert Rothman	
Content Specialist			Gina Cervetti		Nancy Brynelson Sue Pimentel	Carol Lee
Educational Organization		Elena Forzani	Emily Kirkpatrick Allison Skerrett	Karen Malone Cindy Parker	Carol Connor	
Research						John Guthrie
State/District ELA/Literacy Director/Superintendent				Jinghong Cai	Robin Hall	
Teacher/Principal			Sarah Aguirre Eric Turman	Carolyn Aguirre Alicia Ross Paul Wenger	Josephine Franklin	

2025 NAEP Reading Framework Update Visioning Panel List

Asterisk indicates also a member of the Development Panel

David

Dr. P. David Pearson*
Emeritus Faculty Member; Visioning and Development Panel Chair
University of California, Berkeley

P. David Pearson is the Evelyn Lois Corey Emeritus Chair in Instructional Science within the Graduate School of Education at the University of California, Berkeley, where he served as Dean from 2001-2010. His current research focuses on literacy history and policy, including assessment work on statewide assessment in Minnesota and Illinois, the New Standards movement in the 1990s, Smarter Balanced in 2010-2015, and NAEP (continuously since 1973).

Prior to coming to Berkeley in 2001, he served as the John A. Hannah Distinguished Professor of Education in the College of Education at Michigan State and as Co-Director of the Center for the Improvement of Early Reading Achievement. Even earlier, he was Dean of the College of Education, Co-Director of the Center for the Study of Reading, and Professor of Curriculum and Instruction at the University of Illinois. His initial professorial appointment was at the University of Minnesota in Minneapolis from 1969-1978.

He has been active in a range of leadership roles in professional organizations, most notably the International Literacy Association, the National Council of Teachers of English, the American Educational Research Association, the Literacy Research Association, and the National Academy of Education.

He has written and co-edited several books about research and practice, most notably the Handbook of Reading Research. He has served on the boards of many educational research journals. His 300+ books, articles and chapters, written with over 200 co-authors, appear in a range of outlets for a wide range of audiences—teachers, scholars, and policy makers.

2025 NAEP Reading Framework Update Visioning Panel List

Peter

Dr. Peter Afflerbach*
Professor of Education
University of Maryland

Dr. Peter Afflerbach is Professor of Education at the University of Maryland. Dr. Afflerbach's research interests focus on individual differences in reading, the differences and similarities of reading comprehension strategies for print and digital reading, reading assessment, and the verbal reporting methodology. Dr. Afflerbach has served on the National Academy of Education and National Academy of Science committees related to literacy, and the migration of large-scale tests from traditional to digital formats. He is currently concluding a synthesis of the reading comprehension instruction research conducted under the Reading for Understanding funding initiative. Dr. Afflerbach is Chair of the Literacy Assessment Task Force of the International Literacy Association. He was elected to the International Literacy Association's Reading Hall of Fame in 2009. Dr. Afflerbach is the editor of the Handbook of Individual Differences in Reading: Reader, Text, and Context (2016), and co-editor of the Handbook of Reading Research, 4th Edition (2010) and 5th Edition (in press). He has published in numerous theoretical and practical journals, including Reading Research Quarterly, Cognition and Instruction, Elementary School Journal, Journal of Adolescent and Adult Literacy, Language Arts, Theory into Practice, and The Reading Teacher.

Carolyn

Ms. Carolyn Aguirre
Middle School Science Teacher and Department Head
New Haven Unified School District

I moved to the Bay Area in 1993 to teach and attend Cal State Hayward, where I earned my teaching credential and my master's degree in Curriculum Development. I have been teaching in the New Haven Unified School District since 2000, first at Barnard White Middle School, and then at Cesar Chavez Middle School. Before that, I worked in several other school districts in the San Francisco Bay Area. In my 26 years teaching in California, I have taught all three grade levels of middle school science, as well as 8th grade Math and Algebra. Before moving to the Bay Area, I was a Peace Corps volunteer, serving as a high school Science and Math teacher in the Kingdom of Swaziland in Southern Africa. I entered the Peace Corps in 1989 after graduating with a degree in Biology and Spanish from Occidental College in Los Angeles.

2025 NAEP Reading Framework Update Visioning Panel List

Sarah

Ms. Sarah Aguirre*
English Language Arts Teacher
Hobby Middle School, Northside ISD

Sarah is an ELA teacher at Northside ISD in San Antonio, TX. Previously, Sarah was a Field Education Specialist at the University of Texas at San Antonio. There, she conducted research and curriculum writing on several grants. Additionally, she was a reading specialist and literacy coach at a high-needs elementary campus through a grant-funded project. Her experience as an educational coordinator for the UTSA and USAID Read Malawi project in Africa inspired her love for international students. Sarah was the team leader of the Newcomer program at Colonies North Elementary in Northside ISD for 5 years where she taught children with refugee status, many of whom had interrupted or no formal education. She is on the board of Refugee Services at Catholic Charities of San Antonio, a 2016 finalist for the HEB Excellence in Education Award, 2017 Region 20 ESL teacher of the year, and has published an article for The Reading Teacher.

Minnie

Mrs. Minerva Anaya-St John
Lieutenant Colonel (Ret)
United States Airforce

Minerva Anaya-St John was born in Pharr Texas, Oct 17, 1955. During her early years she joined her family working in the fields as a migrant worker. She graduated St. Edwards University in 1977 with Criminal Justice and History degrees. She then joined the Air Force as a second lieutenant. While in the Air Force she commanded/flew AWACS missions in Asia, the Middle East and South America. She also served on the Vice Presidents' Task Force on Drugs, was the first woman to serve in the Pentagon as the executive officer for the Director of Operation for the Air Force and was the Chief of Air Operations at US Central Command. After she left the Air Force, she founded a development and construction company whose projects ranged from first-time home buyer residential to multi-family and commercial construction. Minerva remains in the construction and real estate business to this day.

Nancy

Ms. Nancy Brynelson*
Co-Director, Retired
Center for the Advancement of Reading and Writing, California State University, Chancellor's Office

Nancy Brynelson recently retired as the co-director of the CSU Center for the Advancement of Reading and Writing. Before arriving at the CSU, she served as a bilingual teacher, elementary school principal, school district administrator, and language arts consultant for the California Department of Education. Currently, she oversees the CSU's Expository Reading and Writing Curriculum and several related federal grants. She also co-wrote the 2015 English Language Arts/English Language Development Framework for California Public Schools: Kindergarten Through Grade Twelve. A 2010 inductee into the California Reading Association (CRA) Reading Hall of Fame, she is also the recipient of the CRA 2014 Marcus Foster Memorial Award and the California Association of Teachers of English 2017 Award of Merit.

2025 NAEP Reading Framework Update Visioning Panel List

Jinghong

Dr. Jinghong Cai
Research analyst
National School Boards Association, Center for Public Education

Jinghong Cai, Ph.D. in Curriculum and Instruction. Cai is the research analyst for the Center for Public Education, National School Boards Association. She is a quantitative researcher, and her research focuses on math in early childhood education and policy issues related to students' academic achievement.

Gina

Dr. Gina Cervetti*
Associate Professor
University of Michigan

Gina Cervetti specializes in literacy development and instruction. Her work involves three central concerns: the potential benefits of content-area learning for literacy development, the role of world knowledge in literacy development, and the nature of vocabulary/language instruction that supports reading comprehension. She has been particularly interested in science as a context for elementary students' reading, writing, and language development. She has examined how the collaborative, experiential, and knowledge-enhancing qualities of inquiry-based science instruction can fuel students' engagement and growth in literacy. She has served as a principal investigator on several grants investigating integrated science-literacy instruction with a focus on how science might serve as an especially rich pedagogical context for emerging bilingual students. Cervetti is currently investigating how knowledge-enriching reading and instruction might support students' acquisition of vocabulary knowledge and their ability to engage in complex forms of reasoning within and across texts. She is also involved in investigations of the language demands of school texts and ways to support students' acquisition of word knowledge and conceptual knowledge in support of comprehension. Following her doctoral work in educational psychology at Michigan State University, Cervetti worked for several years as a postdoctoral scholar and researcher at the University of California, Berkeley, on the Seeds of Science/Roots of Reading program. Cervetti joined the University of Michigan in 2011, following three years as an assistant professor at the University of Colorado, Boulder.

2025 NAEP Reading Framework Update Visioning Panel List

Byeong-Young

Dr. Byeong-Young Cho*
Associate Professor
University of Pittsburgh

I am an associate professor of literacy education in the Department of Instruction and Learning at the University of Pittsburgh's School of Education and a research scientist at Pitt's Learning Research and Development Center. My research focuses on understanding cognitive, metacognitive, and epistemic dimensions of student reading and learning in a complex task environment. My recent work examines classroom practices that support student learning and engagement through accessing, processing, and using multiple texts in disciplinary and digital literacies instruction. I have been co-leading various research projects, such as those that investigate middle school learners' historical reading through multisource text inquiries, evidence-centered assessment of digital reading skills, and metacognitively oriented digital literacy intervention for high school learners. I have published my work in scholarly journals such as Cognition and Instruction, Reading Research Quarterly, and American Educational Research Journal, to name a few. I have presented my work regularly at the national and international conferences of leading professional organizations, including the American Educational Research Association and the Literacy Research Association.

Julie

Dr. Julie Coiro*
Associate Professor
University of Rhode Island

Julie Coiro is associate professor in the School of Education at the University of Rhode Island, in the United States, where she teaches courses in reading and digital literacy and co-directs the Ph.D. in Education program and the Graduate Certificate in Digital Literacy. Julie conducts research and speaks nationally and internationally about digital literacies, online reading comprehension strategy instruction, collaborative knowledge building during inquiry, and effective practices for technology integration and professional development. Julie has served as Co-PI on a USDE federally funded research project to develop a series of valid and reliable assessments of online reading comprehension, and a project funded by NAEP-SAIL with colleagues in the US and Finland to explore how students work together to conduct online inquiry and build consensus across multiple online sources. Her work appears in journals such as Reading Research Quarterly, The Reading Teacher, Educational Leadership, and The Journal of Education. She also co-edited the Handbook of Research on New Literacies (2008) and co-authored Teaching with the Internet K-12(2004). Julie's newest co-authored book is titled From Curiosity to Deep Learning: Personal Digital Inquiry in Grades K-5 with Stenhouse (2019).

2025 NAEP Reading Framework Update Visioning Panel List

Carol

Dr. Carol Connor*
Chancellor's Professor in Education
University of California, Irvine

Carol McDonald Connor, Ph.D., CCC-SLP, is a Chancellor's Professor in Education at University of California, Irvine. Her research investigates individual child differences and the links between children's language and literacy development with the goal of illuminating reasons for the perplexing difficulties children who are atypical and diverse learners, including children with dyslexia, have developing basic and advanced literacy skills. Most recently, her research interests have focused on how to individualize (personalize) students' learning opportunities in the classroom –using technology– from preschool through fifth grade and developing and evaluating new technologies to improve teacher efficacy and students' literacy, math, and science outcomes. Awarded the PECASE in 2008, she is also a fellow of AERA and APA. Currently, she is the principal investigator for studies funded by the US Department of Education, Institute for Education Sciences and the National Institute for Child Health and Human Development, including the Early Learning Research Network and the FCRR Learning Disabilities Research Center. She is also past Editor of the Journal for Research in Educational Effectiveness and past Associate Editor for Child Development and currently an Associate Editor for AERA Open.

Elena

Dr. Elena Forzani*
Assistant Professor in Literacy Education
Wheelock College of Education & Human Development

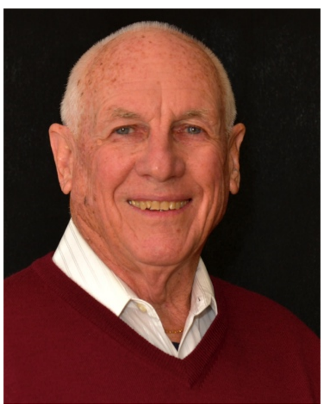
Elena Forzani is an Assistant Professor in Literacy Education at the Wheelock College of Education & Human Development, Boston University, where she teaches undergraduate and graduate courses in literacy assessment and instruction. Her research focuses on understanding how students across the elementary and secondary grades comprehend and use online information, with special attention to the evaluation of online, disciplinary texts. Prior to joining Wheelock, Dr. Forzani was the Assistant Research Director for PIRLS, an international reading assessment housed at Boston College. She was also a fellow at the New Literacies Research Lab at the University of Connecticut, where she worked on the ORCA (Online Research and Comprehension Assessment) Project. Dr. Forzani previously taught high school English and Reading in New Haven, Connecticut, as well as first grade in Louisiana. She earned her Ph.D. in Educational Psychology from the University of Connecticut.

2025 NAEP Reading Framework Update Visioning Panel List

Josephine

Ms. Josephine Franklin
Associate Director
National Association of Secondary School Principals

Josephine Franklin is the Associate Director for Professional Learning at the National Association of Secondary School Principals. As such, she manages principal recognition programs that acknowledge middle level and high school principals and assistant principals from across the country for their leadership and making a positive, significant difference in schools and communities. Also, she manages a communications grant to disseminate information around The Wallace Foundation principal pipeline initiative; and manages NASSP professional learning workshops and the development of Leading Success, an online toolkit. Prior to working at NASSP, she served in a variety of positions with Educational Research Service including management of information services and resource development. Ms. Franklin began her career teaching in the Orange City School District in New Jersey. She has earned a B.A. from Newark State College, M.A. from Kean University in Early Childhood Education and M. Ed from American University in Educational Administration.



John

Dr. John Guthrie*
Jean Mullin Professor
University of Maryland

John Guthrie, Ph.D., is the Jean Mullan Professor of Literacy Emeritus in Department of Human Development and Quantitative Methodology at the University of Maryland at College Park. He received his Ph.D. in Educational Psychology from the University of Illinois in 1968. After being a faculty member at The Johns Hopkins University, he became Research Director at the International Reading Association from 1974-1984. At the University of Maryland, from 1992 to 1997, he was co-director of the National Reading Research Center, funded by the U.S. Department of Education. From 2007-2012, he was Principal Investigator of a 5-year grant from the National Institute of Child Health and Human Development (NICHD) to investigate adolescent reading, focusing on Grade 7 students in a district-wide study. Dr. Guthrie has contributed to such volumes as Handbook of Reading Research (2000), Comprehension Instruction: Research Based Best Practices (2002), What Research Has to Say About Reading Instruction (2002) and Reading Comprehension: The RAND Report for Education (2003). He is a frequent contributor to the peer-reviewed journals of Reading Research Quarterly and the Journal of Educational Psychology and serves on the editorial board for them. Dr. Guthrie is the recipient of the Oscar Causey Award for Outstanding Reading Research and is a member of the International Reading Association Hall of Fame. In 2004, he received the University of Maryland Regent's Faculty Award for research/scholarship/creative activity. In 2011, he was elected to the National Academy of Education addresses research to national policy. In 2012, he was appointed to the Literacy Research Panel of the International Reading Association that investigates literacy policy. In 2017, he was awarded the William S. Gray Citation of Merit. Awarded for Outstanding Lifetime Contributions to Literacy by the International Literacy Association.

2025 NAEP Reading Framework Update Visioning Panel List

Bonnie

Dr. Bonnie Hain*
Chief of Academics and Districts Services
CenterPoint Education Solutions

As the Chief of Academics and District Services, Dr. Bonnie Hain oversees design and development of CenterPoint's product and services to ensure they are of the highest quality and meet the needs of educators. She also works directly with districts and schools across the country to deliver high-quality professional learning on standards implementation, instruction, and assessment literacy. Bonnie has over 25 years of experience in the field of education as a teacher, administrator, researcher, and a Reading and Language Arts assessment developer. She has led assessment design and development projects for districts across the United States, for the Maryland State Department of Education, and for the Partnership for Assessment of College and Careers (PARCC). Bonnie earned her bachelor's degree in Spanish/English education from The State University of New York at Albany, a master's degree from Virginia Tech, and her Ph.D. in English from Stony Brook University. A mother of three grown children and a grandmother of two, Bonnie resides currently with her family near Baltimore, Maryland.

Robin

Dr. Robin Hall
ELA and Literacy Director
Council of the Great City Schools

Dr. Robin Hall is the Director of Language Arts and Literacy for the Council of the Great City Schools. As a member of the Council's academic department, she supports the work of urban educators to improve student achievement for all students by sharing high-leverage information through publications, videos, and webinars, joining strategic support team site visits, and participating in job-alike conferences to facilitate networking and collaboration among member districts. Major efforts this year include providing technical assistance and written guidance for developing and implementing high-quality curriculum documents and professional development to support school staff in elevating teaching and learning to align to college-and career-readiness standards. Dr. Hall also served in various capacities over the course of thirty years in Atlanta Public Schools. She received her B.A. Degree in English from Vassar College and received her M.A. and D.A.H. Degrees from Clark Atlanta University. She is married with two daughters, a granddaughter, and two grandsons.

2025 NAEP Reading Framework Update Visioning Panel List

Kathy

Dr. Kathleen Hinchman*
Professor
Syracuse University

Once a middle school teacher, Kathleen A. Hinchman now teaches undergraduate and graduate classes in childhood and adolescent literacy. Her research is primarily qualitative or design-based and explores youths' and teachers' perspectives toward literacy instruction. She has published in multiple journals and co-authored or edited such texts as Best Practices in Adolescent Literacy Development, Adolescent Literacies: A Handbook of Practice-Based Research, and Teaching Adolescents Who Struggle with Reading. She is currently co-editor of the Journal of Adolescent & Adult Literacy. She has also served as President of the Central New York Reading Council, the New York State Reading Association, and President of the Literacy Research Association (formerly the National Reading Conference). She has also served on multiple New York State English Language Arts standards and assessment committees and as a participant on a Common Core State Standard validation study.

Christy

Dr. Christy Howard
Assistant Professor
East Carolina University

Christy Howard is an Assistant Professor in Literacy Studies at East Carolina University. Prior to coming to ECU, she served as a middle school English Language Arts classroom teacher, an English Language Arts curriculum specialist and an instructional support coach. These roles prepared her for her work at East Carolina University in preparing preservice and in-service teachers to meet the literacy needs of all students. Her research, teaching and service focus on content area literacy instruction, culturally responsive pedagogy and teacher preparation.

2025 NAEP Reading Framework Update Visioning Panel List

Panayiota

Dr. Panayiota Kendeou
Professor
University of Minnesota, Guy Bond Endowed Chair in Reading

Dr. Kendeou investigates the development of higher-order language and cognitive skills that support reading comprehension. In her research she develops theoretical models that explain how students acquire and revise knowledge during reading, and uses those models to design and test innovative, educational technology that transforms reading instruction and assessment (e.g., the federally funded projects TELCI/ELCII; iSTART-Early). Dr. Kendeou is Associate Editor of the Journal of Educational Psychology (and the Incoming Editor in 2020); she also serves on the editorial boards of Scientific Studies of Reading, Contemporary Educational Psychology, Learning and Instruction, Discourse Processes, and Reading Research Quarterly. She has 95+ publications, has served on several advisory boards (e.g., PIAAC, PIRLS), and she is the recipient of several early career awards. She is a member of the American Educational Research Association (AERA), the Society for Text and Discourse (ST&D), the Society for the Scientific Study of Reading (SSSR), the American Psychological Association (APA), and the Psychonomic Society.

Emily

Ms. Emily Kirkpatrick
Executive Director
National Council of Teachers of English

Emily Kirkpatrick is an experienced senior executive with deep expertise in organizational strategy, programmatic innovation, external communications, and fundraising in the education and nonprofit space. Ms. Kirkpatrick assumed her position as the Executive Director of the National Council of Teachers of English (NCTE)—the oldest and one of the largest literacy and education organizations in the United States—in November 2015, overseeing the professional home to English language arts teachers from PreK through university and amplifying the voices of educators through connection, collaborations, and a shared mission to improve the teaching and learning of English. Prior to NCTE, Ms. Kirkpatrick served in multiple leadership roles at the National Center for Families Learning (NCFL), an organization dedicated to eradicating poverty through education solutions for families. During her tenure at NCFL, Ms. Kirkpatrick created the award-winning digital learning platform Wonderopolis®, which reached millions of children across the globe and which was recognized by TIME Magazine as one of the top 50 websites of 2011. A transformative leader, Ms. Kirkpatrick has dedicated her career to public service and civic engagement, seeking to increase national literacy and social mobility, amplify educator voices, and advance the inclusion and empowerment of women. A native and longtime resident of Kentucky, Ms. Kirkpatrick has also served in planning and public relations roles at the Kentucky Office of the Secretary of Education, Arts and Humanities, and advanced the inclusion of women in public service positions while at the Kentucky Commission on Women. She earned her MBA with honors from Bellarmine University and her BA from Centre College in Kentucky.

2025 NAEP Reading Framework Update Visioning Panel List

Carol

Dr. Carol Lee*
Professor
Northwestern University

Carol D. Lee is the Edwina S. Tarry Professor of Education in the School of Education and Social Policy and in African-American Studies at Northwestern University in Evanston, Illinois, U.S.A. She received her Ph.D. from the University of Chicago. She is a past president of the American Educational Research Association (AERA), AERA's past representative to the World Educational Research Association, past vice-president of Division G (Social Contexts of Education) of the American Educational Research Association, past president of the National Conference on Research in Language and Literacy, and past co-chair of the Research Assembly of the National Council of Teachers of English. She is a member of the National Academy of Education in the United States, a fellow of the American Educational Research Association, a fellow of the National Conference on Research in Language and Literacy, and a former fellow at the Center for Advanced Studies in the Behavioral Sciences. She is a recipient of the Distinguished Service Award from the National Council of Teachers of English, Scholars of Color Distinguished Scholar Award from the American Educational Research Association, the Walder Award for Research Excellence at Northwestern University, the Distinguished Alumni Award from the College of Liberal Arts at the University of Illinois-Urbana, The President's Pacesetters Award from the American Association of Blacks in Higher Education, the Lifetime Achievement Award from the American Association of Colleges of Teacher Education and an honorary doctorate from the University of Pretoria, South Africa. She has led three international delegations in education on behalf of the People to People's Ambassador Program to South Africa and the People's Republic of China. She is the author or co-editor of three books, 4 monographs, and has published over 62 journal articles and book or handbook chapters in the field of education.

Karen

Ms. Karen Malone
Curriculum, Instruction, and Assessment Education Specialist
Window Rock, AZ

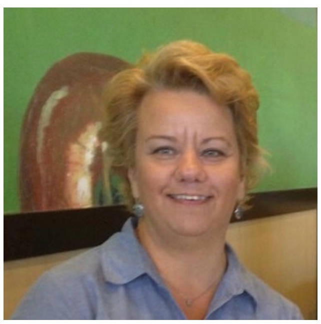
Karen Malone has worked in education for 24 years, gaining experience in instruction, curriculum, coaching, and principal leadership. She attained her Master of Education in Educational Administration from Grand Canyon University. As a seasoned teacher she is passionate about improving Native education and preparing Native students to be college and career ready. In addition to being the Curriculum, Instruction and Assessment Education Specialist, she is also involved in the Strategic Plan Implementation of the Bureau of Indian Education as a unit lead and she serves as a school board member for an indigenous school in New Mexico. Her work most recently has been in facilitating a financial literacy pilot program in Bureau operated schools across Arizona and New Mexico that serve 100% Native students. Outside of the office, Karen enjoys family, camping and traveling. As a lifelong resident of the Southwest, she is captivated by the beautiful sunsets and the endless miles of open country.

2025 NAEP Reading Framework Update Visioning Panel List

Mariana

Dr. Mariana Pacheco*
Associate Professor
University of Wisconsin, Madison, School of Education, Department of Curriculum and Instruction

Mariana Pacheco received her Ph.D. from the UCLA in 2005 (Division of Urban Schooling). She is a former elementary bilingual (English-Spanish) teacher in Southern California. Mariana Pacheco's research focuses on meaningful opportunities for bi/multilingual and English Learner students to use their full cultural, linguistic, and intellectual resources for learning and self-determination. She employs ethnographic and anthropological methods to understand sociopolitical and sociocultural processes related to language, teaching, learning, and curriculum. Her work contributes to theorizations and empirical knowledge of policies, programs, and practices that amplify what 'counts' as knowledge and that enhance bi/multilingual students' academic potential through asset-based and strength-based educational practices, particularly for Chican@/Latin@, (im)migrant, and modest-income backgrounds.



Cindy

Mrs. Cindy Parker
Middle and High School ELA Teacher

Cindy Parker has been an educator for 30 years, serving as a middle and high school ELA teacher, and retired from the Kentucky Department of Education, where she held various roles, including literacy coordinator, grant coordinator, and director of the Division of Next Generation Professionals. She has a BA in English from the University of Kentucky, MA from Eastern KY University, and earned National Board Certification in Adolescent/Young Adult English language arts. She is a past president of the Kentucky Reading Association, International Literacy Association (ILA) member, served on the ILA Common Core State Standards Committee, and a committee that revised the ILA Standards for Literacy Professionals. She is a Kentucky State Literacy team member, an adjunct instructor at the University of Kentucky in the College of Education, works for the Central Kentucky Educational Cooperative as the Special Projects Coordinator, and is the advisor for the Council of Chief State School Officers ELA Collaborative.

2025 NAEP Reading Framework Update Visioning Panel List

Jim

Dr. James Patterson
College Board

Jim Patterson PhD has spent twenty-five years in the fields of teaching, assessment, and standards. After two and a half years of secondary-level English and journalism teaching, he began work at ACT, Inc., in 1996 in ELA test development. From 1998 to 2013, he served first as the content lead for the ACT, PLAN, and EXPLORE Reading tests, spanning grades 8–12, and then in the same capacity for both the English and Reading tests. He also helped design the ELA portions of ACT Aspire (for grades 3–10). In 2013, Jim became senior director (later, executive director) for the ELA/literacy portions of the SAT Suite of Assessments at the College Board, helping redesign and then develop those portions of the SAT Suite (the SAT, PSAT/NMSQT and PSAT 10, and PSAT 8/9, covering grades 8–12) and also leading the design and initial development of the next-generation ACCUPLACER reading and writing college placement tests. Beginning in August 2019, he shifted roles within the College Board to focus on program connections and content strategy for the SAT Suite tests. From 2009 to 2010, Jim served as one of three lead writers for the Common Core State Standards for ELA/Literacy. His main contributors were developing the standards' text complexity materials, drafting the Language standards, editing the standards' evidence appendix, and writing the introductory material for the standards document. Jim earned a Bachelor of Journalism degree, magna cum laude, from the University of Missouri-Columbia in 1992; a Master of Arts in teaching degree in secondary English education from the University of Iowa in 1994; and a PhD in educational policy and leadership studies from Iowa in 2012.

Sue

Ms. Susan Pimentel
Founding Partner
Student Achievement Partners

Susan is a founding partner of Student Achievement Partners, a nonprofit devoted to accelerating student achievement by supporting effective and innovative implementation of college-and career-readiness (CCR) standards. She is also co-founder of StandardsWork, a nonprofit leading the Knowledge Matters campaign. After leading the development of the Common Core State Standards for English Language Arts/Literacy, Susan served as a member of the Understanding Language Project of Stanford University. In that capacity, she co-authored, Realizing Opportunities for English Learners in the Common Core English Language Arts and Disciplinary Literacy Standards. A recent publication, co-authored with Ross Wiener of the Aspen Institute, Practice What You Teach: Connecting Curriculum and Professional Learning in Schools highlights the work jurisdictions are doing to integrate high-quality instructional materials with professional learning. A 2018 commentary published in EdWeek, Why Doesn't Every Teacher Know the Research on Reading Instruction, shares three evidence-based practices that can boost reading proficiency. Ms. Pimentel served two terms on the National Assessment Governing Board, an independent, bipartisan board that sets policy for the national assessment. She became vice-chair of the body in November 2012. She holds a bachelor's degree in early childhood education and a law degree from Cornell University.

2025 NAEP Reading Framework Update Visioning Panel List

Alicia

Ms. Alicia Ross*
Teacher and Educational Consultant
Blue Ridge Middle/High School

I am a high school Social Studies teacher at Blue Ridge High School in New Milford, PA. We are a small rural school district in the northeastern corner of the state. I just completed my twentieth year in education. I teach AP US Government and Politics, AP Macroeconomics, General Economics/Global Studies course, and Law/Sociology. I currently serve as the teacher-leader at my school for our Reading Apprenticeship Professional Learning Community. Due to my intense interest in serving my students and addressing their literacy needs, I just completed my second master's degree. This second degree is in Reading Instruction from Wilkes University. I am a consultant for Reading Apprenticeship and for the College Board's AP US Government and Politics workshops and summer institutes. I currently live in Throop, PA and have one grown son who practices law in New York. I am avid reader, runner, and pickleball player!

Robert

Mr. Robert Rothman*
Senior Editor
National Center on Education and the Economy

Robert Rothman is a senior editor at the National Center on Education and the Economy and a writer and editor for numerous education organizations. Previously, he was a senior fellow at the Alliance for Excellent Education, a Washington, D.C.–based policy and advocacy organization, and he was a senior editor at the Annenberg Institute for School Reform, where he edited the Institute's quarterly magazine, Voices in Urban Education. He was also a study director at the National Research Council, where he led a committee on testing and assessment in the federal Title I program, which produced the report Testing, Teaching, and Learning (edited with Richard F. Elmore) and a committee on teacher testing. A nationally known education writer and editor, Mr. Rothman has written numerous reports and articles on a wide range of education issues. He is the author of Something in Common: The Common Core Standards and the Next Chapter of in American Education (2011) and Measuring Up: Standards, Assessments and School Reform (1995), and the editor of City Schools (2007). Mr. Rothman holds a degree in political science from Yale University.

2025 NAEP Reading Framework Update Visioning Panel List

Allison

Dr. Allison Skerrett*
Professor, Department of Curriculum and Instruction
Director of Teacher Education, College of Education, The University of Texas at Austin

Dr. Skerrett is a professor of language and literacy in the Department of Curriculum and Instruction at The University of Texas at Austin. Professor Skerrett is also Director of Teacher Education for The University's College of Education. Dr. Skerrett's expertise includes secondary reading and English education; adolescents' literacy practices, including those of transnational/migrant youth; secondary English teacher preparation; urban education and sociocultural influences on teaching and learning.

Eric

Mr. Eric Turman
Principal
Reading High School

Reading High School was always categorized as one of the lowest performing schools in the state of Pennsylvania. Eric was determined that Reading would no longer be part of any state report that categorizes the school as low performing. Reading High School has increased its graduation percentage from 53 to almost 70 percent over the past several years and the dropout rate has decreased from 13 to 5 percent. During the 2016 and 2019 school years, Reading High School received the Bronze and Silver medal from US & News Report as one of the top high schools in the country. This is a tremendous honor considering 95% of the students who attend Reading High School are categorized as underserved and almost 85% of the student body is Latino. Under Eric's leadership he has built a community of success where every teacher, administrator, staff and parent in the Reading School District has played a role in the success of the students. Eric is a strong leader with a shared vision that has carried him and his team to have a tremendous impact on the children attending Reading Senior High.

2025 NAEP Reading Framework Update Visioning Panel List

Paola

Dr. Paola Uccelli*
Professor
Harvard Graduate School of Education

Paola Uccelli is a professor at the Harvard Graduate School of Education. With a background in linguistics, she studies socio-cultural and individual differences in language and literacy development throughout the school years. Uccellii's current projects focus on describing individual trajectories of school-relevant language development; on the design and validation of a research instrument to assess school-relevant language skills in elementary and middle school students; and on understanding how monolingual and multilingual speakers and writers learn to use a variety of discourse structures flexibly and effectively for diverse communicative and learning purposes. Uccelli studied linguistics at the Pontificia Universidad Católica del Perú and subsequently earned her doctoral degree in Human Development and Psychology at the Harvard Graduate School of Education. Being a native of Peru, she is particularly interested in Latin America where she collaborates with local researchers and often participates in research conferences and workshops.

Paul

Mr. Paul Wenger
Vice President
National Association of Elementary School Principals

Paul Wenger is Vice President of the National Association of Elementary School Principals. He is also the Principal at Jordan Creek Elementary in West Des Moines which is a Leader in Me school. Wenger was previously an elementary principal at Edgewood-Colesburg Community School District. He also served as President of the School Administrators of Iowa and has been a principal mentor, legislative committee member, and Iowa Leadership Academy Steering Committee member. Wenger has implemented professional learning communities, statewide voluntary preschool programming, PBIS, and multi-tiered systems of supports for students. Prior to working in school administration, Wenger taught elementary school in the Central Community School District for 13 years. He received his bachelor's degree in elementary education and physical education from Wartburg College and his master's degree in educational leadership from Iowa State University.

2025 NAEP Reading Framework Update Visioning Panel List

Victoria

Ms. Victoria Young
Director (retired), Reading, Writing, and Social Studies Assessments
Texas Education Agency

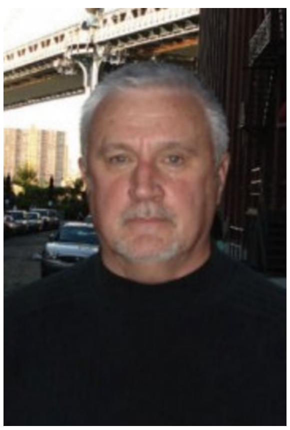
As the Director of Reading, Writing, and Social Studies Assessments for the state of Texas, Victoria Young was directly responsible for managing content development as well as for overseeing all activities related to the scoring of approximately four million compositions and short answer reading responses each year. During her 27-year career, she focused her efforts on designing state assessments that contributed to a fuller understanding of student achievement and instructional programs. Since her retirement in 2015, she has served in a leadership role in the development and implementation of new English language arts and reading content standards for Texas. She continues to be particularly interested in the ways in which coherent, vertically aligned reading and writing programs and authentic instructional literacy practices can increase the academic success of all students, both in the classroom and on state and national assessments.

2025 NAEP Reading Framework Update
Technical Advisory Committee

Derek

Dr. Derek C. Briggs
Professor, Research and Evaluation Methodology
University of Colorado, Boulder

Derek C. Briggs is a professor of quantitative methods and policy analysis and chair of the Research and Evaluation Methodology program at the University of Colorado Boulder. He is also the director of the Center for Assessment Design Research and Evaluation (CADRE). Dr. Briggs's research agenda focuses upon building sound methodological approaches for the measurement and evaluation of growth in student learning. He has a special interest in the use of learning progressions as a method for facilitating student-level inferences about growth and helping to bridge the use of test scores for formative and summative purposes. Other interests include the use and analysis of statistical models to support causal inferences about the effects of educational interventions on student achievement.



Howard

Dr. Howard Everson
Senior Principal Research Scientist
SRI International

Howard T. Everson is the Director of Assessment Design & Research in the Center for Technology in Learning at SRI International. He is also a Professor of Psychology at the Graduate School, City University of New York and former Director of the Center for Advanced Study in Education at the Graduate School, City University of New York. His research and scholarly interests focus on the intersection of cognition, technology and assessment. Professor Everson's measurement expertise is in the areas of item response theory, differential item functioning, learning analytics and cognitive diagnostic measurement models. Dr. Everson served as the Executive Director of the NAEP Educational Statistics Services Institute at the American Institutes for Research and was Vice President and Chief Research Scientist at the College Board. Dr. Everson is a Psychometric Fellow at the Educational Testing Service, and an elected Fellow of both the American Educational Research Association and the American Psychological Association, and a charter member of the Association for Psychological Science. Dr. Everson is the current editor of the National Council of Measurement in Education's journal, *Educational Measurement: Issues and Practice*.

2025 NAEP Reading Framework Update
Technical Advisory Committee

Joan

Dr. Joan Herman
Co-Director Emeritus
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)

Joan Herman is Director Emerita of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at UCLA. A member of the National Academy of Education and elected Fellow of the American Educational Research Association. Dr. Herman's. research has explored the effects of testing on schools and the design and use of systems of assessment to support school accountability and improvement. Her recent work focuses on the quality and effects of teachers' formative assessment practices, fairness in testing and the assessment of deeper learning. She also has wide experience as an evaluator of school reform.

Dr. Herman received her BA in Sociology from the University of California, Berkeley, was awarded an MA and Ed.D in Learning and Instruction from the University of California, Los Angeles, and is a member of Phi Beta Kappa.

Kristen

Dr. Kristen L. Huff
Vice President
Curriculum Associates

Since May 2016, Dr. Huff has been the Vice President of Assessment and Research at Curriculum Associates, where she leads a team of more than 20 assessment designers, psychometricians, and researchers. Curriculum Associates supports a system of online assessments integrated with personalized learning and whole-class instruction designed to help teachers teach more effectively and students reach their full learning potential. Dr. Huff's work focuses on ensuring the coherence of design, interpretation, use, and policy across formative, interim, and summative assessment to advance equity and high-quality education for all students. Dr. Huff received her Ed.D. in Measurement, Research and Evaluation Methods from the University of Massachusetts Amherst. Prior to her studies at UMass, Dr. Huff completed a master's degree in Educational Research, Measurement, and Evaluation from the University of North Carolina at Greensboro.

2025 NAEP Reading Framework Update
Technical Advisory Committee

Michael

Dr. Michael Kolen
Professor Emeritus in Educational Measurement
University of Iowa

Professor Michael J. Kolen is a Professor Emeritus in Educational Measurement at the University of Iowa. Dr. Kolen received his doctorate from the University of Iowa in 1979, his MA degree from the University of Arizona in 1975, and his BS degree from the University of Iowa in 1973. He served on the faculty at Hofstra University from 1979-1981, and he worked at ACT from 1981-1997, including being Director of the Measurement Research from 1990-1997. Dr. Kolen co-authored three editions of the book Test Equating, Scaling, and Linking: Methods and Practices, published by Springer-Verlag. He has published numerous articles and book chapters on various topics in educational measurement and statistics, including test equating and scaling. Dr. Kolen has been President of the National Council on Measurement in Education (NCME) and is past editor of the Journal of Educational Measurement. He is a Fellow of Division 5 of the American Psychological Association, a Fellow of the American Educational Research Association, and member of various other professional organizations. Dr. Kolen served on the 2014 Joint Committee on the Standards for Educational and Psychological Testing. Dr. Kolen received the 1997 NCME Award for Outstanding Technical Contribution to the Field of Educational Measurement and the 2008 NCME Award for Career Contributions of Educational Measurement.

Scott

Dr. Scott Marion
Executive Director
National Center for the Improvement of Educational Assessment

Scott Marion, Ph.D. is the President and Executive Director of the Center for Assessment. He is a national leader in designing innovative and balanced assessment systems to support both instructional and accountability uses. Dr. Marion coordinates and/or serves on state and district technical advisory committees (TAC) for assessment and accountability. Dr. Marion has served on multiple National Research Council (NRC) committees related to next generation science assessments, the issues and challenges associated with incorporating value-added measures in educational accountability systems, and on outlining the "best practices" in state assessment systems. Dr. Marion regularly presents the results of his work at national conferences and has published dozens of articles in peer-reviewed journals and edited volumes. A former field biologist and high school science teacher, Dr. Marion has a master's degree in Science Education from the University of Maine and a Ph.D. in measurement and evaluation from the University of Colorado, Boulder. Prior to joining the Center for Assessment in early 2003, Dr. Marion served as the Director of Assessment and Accountability for the Wyoming Department of Education. Finally, Dr. Marion has served on his local school board for 6 years in Rye, NH.

2025 NAEP Reading Framework Update
Technical Advisory Committee

Jennifer

Dr. Jennifer Randall
Associate Professor
University of Massachusetts

Jennifer Randall, Associate Professor, joined the University of Massachusetts faculty in 2007. She earned her BA (1996) and MAT (1999) from Duke University and Ph.D. in 2007 from Emory University. Prior to her graduate studies, Jennifer taught pre-school and then high school social studies for several years. Her research interests primarily reflect the measurement issues and concerns she encountered as a classroom teacher which include the grading practices/philosophies of teachers, particularly differential practices as they relate to students of color, first generation students, English learners, and students with disabilities. She is especially interested in the ways in which assessments (both large-scale & classroom-based) take into consideration, and impact, historically marginalized populations in the U.S. and abroad. Dr. Randall's research areas include applications of the Rasch model, to assess measurement invariance in high stakes reading assessment, grading practices, and test accommodations.

Willy

Dr. Guillermo Solano-Flores
Professor
Stanford University, Graduate School of Education

Dr. Guillermo Solano-Flores is Professor of Education at the Stanford University Graduate School of Education. He specializes in educational assessment and the linguistic and cultural issues that are relevant to both international test comparisons and the testing of cultural and linguistic minorities. He has conducted research on the development, translation, localization, and review of science and mathematics tests. He has been principal investigator on several National Science Foundation-funded projects that have examined the intersection of psychometrics, semiotics, and linguistics in testing. He is the author of the theory of test translation error, which addresses testing across cultures and languages. Also, he has investigated the use of generalizability theory—a psychometric theory of measurement error—in the testing of English language learners and indigenous populations. He has advised countries in Latin America, Asia, Europe, Middle East, and Northern Africa on the adaptation and translation of performance tasks into multiple languages and the development of assessment systems. Current research projects examine academic language and testing, formative assessment practices for culturally diverse science classrooms, and the design and use of illustrations in international test comparisons and in the testing of English language learners.

Revision of the
Item Development and Review Policy

The purpose of the Item Development and Review Policy is to articulate principles for quality and fairness of NAEP assessments to all NAEP audiences. The policy also outlines the Board's role in ensuring a fair assessment.

Since the policy's last revision in 2002, NAEP has shifted to digital administration, and the educational measurement field has embraced new professional standards and best practices in assessment. Collaboration protocols for NCES and the Governing Board have also evolved, partially in response to the iterative nature of item development for digital-based assessment. Consequently, there are opportunities to bring this policy up to date, while supporting new efficiencies.

In 2019, Governing Board staff hosted an expert panel meeting with assessment development leaders to raise issues that need be addressed in an updated Governing Board policy on item development and review (see attached meeting minutes). This conversation highlighted current best practices and potential directions for the policy update.

In this Assessment Development Committee (ADC) session, Michelle Blair of the Governing Board staff will provide a briefing on how this policy has been implemented over time, with a focus on possible ways for the Governing Board to exercise final authority over the appropriateness of NAEP items – a legislated responsibility of the Board. The ADC carries out this responsibility because the Governing Board has delegated this final authority to the Committee. A primary goal of this initial policy discussion session is to identify the Committee's consensus about how this final authority ***should*** be exercised, given the possible ways for exercising this final authority.

Central questions include:

- What does the Board need to review in order to stand behind the items on the assessment?
- How should reviews be conducted to support the Board's confidence in the assessment?

Notes of the Expert Panel Meeting* on NAEP Item Development and Review Policy August 12–13, 2019

Contents

* This meeting was held on August 12–13, 2019 at HumRRO's facilities in Alexandria, Virginia. Discussion was facilitated by Thanos Patelis of HumRRO and Michelle Blair of the Governing Board staff.

Summary of Expert Panel Meeting on NAEP Item Development and Review Policy

August 12–13, 2018

Purpose and Participants

A panel of experts convened to discuss best practices in item development and review to inform the revision of the Governing Board Item Development and Review Policy. Board staff have identified this policy revision as an opportunity to:

- reflect best practices,
- consider digitally-based assessment (DBA),
- remove procedurally-oriented detail that will be included in a procedures manual,
- improve through-lines between framework creation & revision and item development & review; and
- create efficiencies, where possible.

Panel members included **Joan Herman**, retired Director of University of California, Los Angeles's National Center for Research on Evaluation, Standards, and Student Testing (CRESST); **Stuart Kahl**, independent consultant; **Erika Landl**, Senior Associate, Center for Assessment; **Jeffrey Nellhaus**, independent consultant; **Marianne Perie**, President, Measurement in Practice; **Ed Roeber**, Director of Assessment, Michigan Assessment Consortium; **Deb Sigman**, Interim Program Director for the Comprehensive School Assistance Program (CSAP) at WestEd; and **Andrew Wiley**, Partner, ACS Ventures.

Key Features of the Current Policy

Last revised in 2002, the purpose of the Item Development and Review Policy is to articulate principles for quality and fairness of NAEP assessments to all NAEP audiences. The policy describes item development for NAEP at large, with a focus on the Board's role in ensuring a fair assessment. Accordingly, there are six principles in the current Item Development and Review Policy.

Principles 1 and 2 address <u>*fidelity to the content guidance*</u> provided by the Board in the form of frameworks, specifications, and achievement level descriptions.

> *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 NAEP Basic, NAEP Proficient, and NAEP Advanced performance shall be an important consideration in all phases of NAEP development and review.

Principles 3 and 4 operationalize the Board's legislative authority relevant to *review of assessment items*.

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 test questions that are to be administered in conjunction with any pilot test, field test, operational assessment, or special study administered as part of NAEP.

Principles 5 and 6 designate broader *criteria for assessment items' appropriateness* that must be satisfied, representing the final *outcomes* of item development and review processes.

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

Principle 6. All NAEP test questions shall be free from racial, cultural, gender, or regional bias, and shall be secular, neutral, and non-ideological. NAEP shall not evaluate or assess personal or family beliefs, feelings, and attitudes, nor publicly disclose personally identifiable information.

In engaging with Panelists, additional clarifications about the policy were provided:

- The current policy does not refer to scoring procedures. Instead, the focus is on accurate representations of the content in the assessment items and the associated scoring criteria for each item.
- The current policy includes several procedurally-oriented details that would be more appropriate for a procedures manual.
- Reflecting the matrix sampling design of NAEP, the current policy applies to all assessment items that are part of a large item pool. Each student takes only a portion of the items in the assessment, but the complete assessment includes the full item pool for the subject area. This differs from state assessments where the full item pool for a subject area is administered to each student.
- The current policy was developed early in the history of the Governing Board and before the transition to digital-based assessment.
- The current policy addresses assessment items, without explicit details regarding survey questionnaire items.

Current Implementation of the Policy

NCES calls on many constituents to help guide item development and review for each assessment, beginning with a rigorous review of the existing item pool to identify needs. New items undergo a series of testing, revisions, and reviews prior to submission to the Governing Board. These activities include:

- Item development contractor reviews by content experts on political sensitivity, bias, fairness, editorial, and language accessibility and translatability.
- Pretesting activities, such as:
  - playtesting, where an individual or small group works through and discusses an item or set of items with a facilitator or observer,
  - small and large scale item tryouts, where students work uninterrupted through a set of programmed items, and
  - cognitive labs, where students are instructed to think-aloud as they work through item sets during a structured one-on-one interview.
- Reviews by “standing committees” composed of teachers and other content experts, state and local education agency representatives, and content area researchers. Each standing committee reviews new items for:
  - the grade level appropriateness of the items for the particular grade;
  - the representative nature of the item set;
  - the alignment of the items with the framework and test specifications;
  - the quality and content accuracy of items and scoring rubrics, and
  - potential bias and sensitivity issues.
- An independent expert panel review to confirm mathematical accuracy in the case of the mathematics assessment; and
- NCES review of items and scoring rubrics.

Hence, NCES puts forth a rigorous effort when they submit items and other material for Governing Board review. After Board review, NCES attends to Board member comments, and may explain why a change will not be made based on evidence from cognitive labs, for example. While NCES includes external experts as part of their reviews of items before submission to the Governing Board, Panelists commented that the Board review adds greater stakeholder representation to the NAEP enterprise.

The law states that "the Governing Board shall have final authority on the appropriateness of all assessment items." Historically, the Governing Board has interpreted "final authority" over all items as a process that requires "signing off" on each item. Currently, the Governing Board, represented by the Assessment Development Committee (ADC), reviews assessment items for all subject area assessments before and after pilot testing. Reviews after pilot testing include item data, such as p-values. In recent years, the ADC has streamlined post-pilot reviews: re-reviewing pre-pilot items only if they are flagged by the Committee because of a concern. Assessment items for special studies are not typically reviewed by the ADC.

The Governing Board's reviews of assessment items are largely focused on appropriateness, while NCES engages in a series of reviews and processes to address content matters. As a demonstration of the emphasis of comments from assessment item reviews, the following table shows the types of item-level comments from the ADC over a three and a half year period.

January 2015 – June 2018: ADC Cognitive Item Specific Comments (N=965)

Comment Code	N	Percent
Typo & Grammar	32	3.32%
Language/ Wording	142	14.72%
Clarification Needed	134	13.89%
Expand Item	119	12.33%
Drop Item/ Passage	18	1.87%
Question Format	84	8.70%
Remove Redundancy in Stem and Response Options	8	0.83%
SES Bias	6	0.62%
Regional Bias	10	1.04%
Requires prior content knowledge	3	0.31%
Topic Appropriateness (e.g., sensitive, controversial)	15	1.55%
Graphics/ Layout	131	13.58%
Framework Target Mismatch	25	2.59%
Avoid Datedness	30	3.11%
Difficulty Concern (e.g., vocab., language load, etc.)	45	4.66%
Scoring Rubrics - Changes	34	3.52%
Scoring Rubrics - Positive	6	0.62%
Item Statistics	41	4.25%
Positive Engagement (e.g., relevant)	67	6.94%
Negative Engagement	55	5.70%
Grade appropriateness	38	3.94%
Affirmative (e.g., exceptional, engaging, etc.)	244	25.28%

For reviews of contextual questionnaire items, the ADC reviews items for subject-specific questionnaires, while the Governing Board Reporting and Dissemination Committee reviews items for the core contextual questionnaires.

As a demonstration of the emphasis of comments from questionnaire item reviews, the following table shows the types of item-level comments from the ADC over the same period above.

January 2015 – June 2018: Survey Item Specific Comments (N=179)

Comment Code	N	Percent
Typo & Grammar	18	10.06%
Language/Wording	60	33.52%
Clarification Needed	35	19.55%
Expand Item	49	27.37%
Drop Item/ Passage	6	3.35%
Question Format	27	15.08%
Remove Redundancy in Stem and Response Options	8	4.47%
Use Question for Other Grades	5	2.79%
Use Question for Other Subjects	4	2.23%
SES Bias	2	1.12%
Regional Bias	2	1.12%
Topic Appropriateness (e.g., sensitive, controversial)	3	1.68%
Graphics	0	0.00%
Avoid Datedness	18	10.06%
Item Statistics	3	1.68%
Positive Engagement (e.g., engaging, relevant)	0	0.00%
Negative Engagement	0	0.00%
Affirmative (e.g., exceptional, engaging, etc.)	6	3.35%
Applies to Other Respondents	3	1.68%

Summarizing the patterns in ADC review comments over this period, approximately a quarter of comments involved positive feedback. The most frequent negative feedback involved the need for clarity. ADC members have rarely expressed concerns about framework alignment or bias.

Balancing Important Considerations

An important outcome for this expert panel meeting was to surface central questions and issues that must be addressed in updating the Governing Board Item Development and Review Policy. Panelists' discussion surfaced a number of factors that must be balanced in updating the policy to meet the Board's current goals.

(1) **How to reflect the Board's accountability role.** Ensuring the appropriateness of NAEP appears to be very deliberately chosen in the legislation as part of the Board's role. The National Assessment of Educational Progress Authorization Act of 2002 (P.L. 107-279) states the Board has "final authority" over all items. To be accountable to the public, the Board has conducted an approval process that addresses all items. However, "final authority" could be interpreted and exercised in different ways. The procedure to evaluate appropriateness could be defined differently, with a range of options.

(2) **What level of detail is required for the policy?** Currently, the policy defines a granular level of detail for some processes. This does not seem entirely appropriate for a policy

document. For example, Panelists discussed that the procedures manual should address reviewer training and qualification, since there are different qualifications for stakeholders at different times in the item development and review process. Panelists noted that the literature on stakeholders in item reviews is dated, and current practice is documented in technical documents that are not always made public and vary in quality. Additionally, when there is a new framework, a member of the framework committee should be available to respond to questions about the framework or to clarify any confusion. The current Governing Board Item Development and Review Policy provides for this participation of former framework committee members.

(3) **How to best reflect a focus on critical outcomes from item development and review.** As much as possible, the policy should be focused on what NAEP needs to reflect to students, teachers, schools, and the public. The policy revision process should avoid being overly specific on the procedures required to achieve this outcome.

(4) **How to avoid unnecessary redundancy in Governing Board reviews.** Given the multiple item reviews by others in the NCES process, Panelists discussed that the role of the Governing Board should be to identify "fatal flaws."

(5) **How to assure that Governing Board reviews "add value**." NCES noted that more reviews help NAEP remain the gold standard. At the same time, Panelists encouraged the Board to be clearer about how Board reviews are intended to add value to NAEP.

(6) **How to assure an efficient review process.** It is important to streamline review processes in ways that maintain the Board's confidence in the assessment items.

(7) **What evidence does the Governing Board need to approve items?** Given the elaborate development process conducted by NCES, there are opportunities for the Board to review related evidence for appropriateness. For example, DBA comes with a number of new metrics to be considered. One question is: how should process data be used? Panelists agreed the procedures manual should address the validity evidence needed for item review. Information is needed to help answer questions from critics or the public. Procedures should address the different types of evidence relevant to reviews both before and after pilot testing. In all cases, Panelists agreed that reviewers need scoring rubric information to appropriately evaluate items.

(8) **How the policy should cite existing best practices for large scale assessment.** To what extent should the policy restate best practices that are articulated elsewhere versus merely referring to it? In many cases, the policy will need to account for differences between state assessments and NAEP. Relatedly, Panelists discussed overall quality and several resources relevant to universal design in assessment:

- National Center on Educational Outcomes has guidelines for students with disabilities (SDs)
- Center for Applied Linguistics has guidelines for English learners (ELs)
- WIDA project on eliminating academic language on assessments

(9) **How to assure that assessment items are fresh and innovative.** Should there be a policy principle about maintaining trend versus innovation or should this be largely subsumed under the Governing Board Framework Development Policy?

(10) **How the policy should align with other Board policies.** Some of the core inputs for item development and review are referenced in other policies. For example, the specifications are defined in the framework policy. Should additional details about the specifications be included here?

(11) **How the policy should use language from the NAEP legislation.** Panelists noted that some principles restate legislation. There may be better ways to signal alignment with the law.

(12) **How assessment block construction should influence item and assessment review.** Given the matrix sampling design of NAEP, a key question is how this design should shape the review that is conducted by the Board.

(13) **How the review process should mimic the students' experience.** Related to the previous factor (using assessment blocks), Panelists' discussion centered on how the review process should account for what the student experiences.

(14) **How framework fidelity should be addressed in Board policies and procedures.** Panelists discussed that the first two principles of the current policy (addressing fidelity to the framework and ALDs) are still relevant. Still, how should the Board determine the extent to which cognitive demand is represented in the item pool as called for in the framework, for example?

(15) **How to surface tensions between the framework and the item pool.** Tensions in the framework are typically identified by item writers, before the Governing Board reviews items. Alignment studies also help identify tensions in a framework. Additionally, Panelists suggested Board members receive summary information about the passages and items that are not successful (i.e., dropped after the pilot assessment or as a result of Board reviews).

(16) **How reviews should address new item types or new scenario-based tasks (SBTs).** New item types and other new items warrant special reviews.

(17) **How the policy should articulate the representation of balanced perspectives.** The current policy requires assessments to include balanced perspectives across the entire item pool for a given subject-area assessment at a grade, i.e., that the item pool is neutral and non-ideological. However, it is not entirely clear how NAEP item developers would create a summary of how well perspectives are balanced. What information would the Governing Board need to review and assess the balance of perspectives?

Policy Recommendations

While discussing how to balance the factors articulated above, Panelists put forward the following recommendations.

Review a designated subset of the item pool for each NAEP assessment. Panelists tended to agree that the Board could maintain final authority without approving every item, depending on their relationship with NCES and its contractors. Others suggested the only way to avoid fatal flaws might be to review every item. Panelists discussed ways in which the Governing Board could look at a subset of items to determine the extent to which they reflect the framework, are free from bias, and meet other criteria. For example, the Governing Board review could be focused only on:

- items where opinions differed across NCES reviewers;
- items with policy implications;
- items flagged for differential item functioning; and/or
- items where cognitive labs raised issues.

Review the assessment as a whole. Although specific items may be free from bias, it is possible that the item pool composed of such items could have biases. Panelists cautioned against leaning too heavily on item-level data versus assessment-level data. They discussed the need for the Board to provide a higher-level review at the item-pool level to ensure a balanced assessment. These reviews should center on fidelity to the framework, including alignment to the Board's NAEP specifications for content balance, cognitive demand balance, achievement level descriptor coverage, and item type balance. For example, the Every Student Succeeds Act of 2015 (ESSA) requires states to provide evidence to document adequate validity based on content for the State's general assessments. This includes an independent alignment study documenting that each assessment is aligned to its framework. Review of aggregate information for the entire pool will support construct validity.

Conduct an evidence-driven review. The Board's review should center on pieces of evidence rather each item. States need to produce a variety of evidence to meet peer review requirements, such as those addressing technical quality and validity. Other types of evidence to collect and review include item statistics, frameworks, specifications, item writing guidelines and training manuals, item writing checklists, and bias and sensitivity checklists (e.g., lists of taboo topics). In an example from Alaska, item developers created a matrix of passages and sets of items to identify the objectives measured and pilot statistics. The matrix included hyperlinks to the items. The process is also evidence of quality, and the Board needs to trust the process instituted by NCES. Panelists echoed the idea that the Board should trust but verify with evidence. A potential new principle could state that the Governing Board shall use evidence-based processes to ensure the quality of items and the quality of the assessment.

Use additional tiers of review. In discussion of streamlining and focusing Governing Board review, Panelists talked about assigning items to different tiers based on the level of confidence in the item. For example, new item types should receive more scrutiny. When states hire a new vendor, they should look for more evidence and conduct a more extensive review. Validity studies should be conducted when there is a new assessment. There might also be a validity review when releasing items, identifying how these new gaps should influence future assessment reviews. Future periodic alignment studies could also be restricted to the degree of alignment for new items only, rather than conducting another alignment study for the full item pool. Panelists also offered examples of post-pilot reviews as limited and focused on current event issues or hot topics, e.g., something that has occurred since the item was originally written that would trigger a sensitivity issue and potentially create a gap in framework coverage. However, the timing of previous reviews plays a role in this. If NCES reviews are only weeks prior to Governing Board reviews, then hot topic reviews should not be necessary.

Consider new mechanisms for conducting reviews. Some aspects of review at the state level are handled by independent review panels to meet legislative requirements or as best practice. The panel noted the importance of having independent reviews. In response to the need for broad representation among reviewers and the logistics of conducting reviews, states use remote reviews to include multiple perspectives and in-person review sessions to gather input from a smaller number of individuals.

Consider how reviews can prioritize English language learners and students with disabilities. Leading states have prioritized accessibility for English learners and students with disabilities. States include people with a broad range of experience across the accommodation spectrum in their item review processes. Given that NCES already conducts several such reviews, the Board should consider what information NCES can provide to the Board along these lines.

Consider how alignment studies can support Board reviews. Panelists discussed the role of alignment studies, one method for collecting validity evidence, in NAEP and state accountability assessments. Such studies might show that items do not match the expected cognitive demand or an inadequate coverage of the framework objectives. To establish overall validity, Panelists listed general types of evidence reviewed in alignment studies, including focus or balance, reach, match, and range.

Consider how changes in the policy should be communicated to other NAEP stakeholders. Panelists recommended proactively communicating changes in the policy to stakeholders, after it is revised.

Strengthen feedback mechanisms. Board staff currently attend several NCES review meetings. Board staff should proactively share the results of these meetings with the ADC as part of the Committee's assessment reviews. Similarly, how NCES addresses feedback from the ADC should be shared more proactively with the Committee.

Update policy to include consideration of digital-based assessment. The policy was written before DBA, and there are a number of related factors that should be addressed. For example, principle 5 (on accuracy of items) should include more information about new item types.

Articulate roles and responsibilities. Use the procedures manual to articulate who is responsible versus who is contributing to which outcomes and products. Clarify the resources needed to carry out the policy. Theory of action terminology could be a resource, e.g., start with defining guiding principles based on high priority conditions that must be maintained to uphold goals, including fairness and accuracy.

Provide more guidance on how framework objectives are to be sampled. Panelists cited a 2007 NAEP Validity Studies (NVS) Panel study of grades 4 and 8 mathematics. They concluded that the probability that an objective would be assessed in a given year ranges from .25 to 1.0. Given the sampling of framework objectives, there is a need to systematically sample content from the framework.

Incorporate questionnaire reviews into the policy. While a separate principle to address questionnaires specifically may not be needed, the policy should more explicitly include questionnaire reviews.

Overall, Panelists suggested reducing the size of the Board's review of items and increasing the review of the item pool. In addressing potential pitfalls of item-level reviews, Panelists noted that rephrasing words (i.e., “wordsmithing”) is a tendency in stakeholder reviews that must be avoided in the Board’s review.

Criteria for Assessments and Items

In delineating overall aspects of quality, Panelists emphasized the importance of fidelity to the framework (i.e., content and item type match) and accuracy of content as important contributors to overall quality. Consistent editorial and style requirements and elements of universal design also fit under this heading. Other important dimensions include technical quality (e.g., difficulty, discrimination); bias and sensitivity (including consideration of passage and item features, such as the protagonists); online rendering across devices and platforms (i.e., the student’s user experience); and accessibility of digital tools (e.g., equation editor, drag and drop).

Panelists noted there are fewer quality review features to address in assessment reviews, compared with item reviews. For the assessment, fidelity to the framework can be reviewed in terms of the specified percentage emphasis for each sub-content area and item type. The extent to which achievement level descriptors (ALDs) are represented can also be reviewed.

Panelists discussed what fairness means at the assessment level. Panelists agreed accessibility is not a typical assessment-level concern; however, they noted it is possible for speededness to impact a student’s experience and there may be an interaction between accessibility and delivery mode.

Structuring an Updated Policy

Panelists generally reached consensus to include overall definitions and purpose statements in an introduction that provides more details on governance protocols. This section should focus on the responsibility of the Governing Board while recognizing the responsibilities of NCES. The policy should focus on item as well as assessment qualities to include assurance of quality, exceptions to accountability, and guidelines for accuracy and other aspects of item quality.

Panelists discussed the benefits of separate principles for items and assessments versus broad principles that cover items and assessments together, with details at the item and assessment levels added in the procedures manual. Some Panelists delineated three categories of principles: (a) governance or responsible party, (b) item quality characteristics (e.g., bias, accuracy), and (c) reviews. Alternative top-level principles included (a) items and item pool, (b) content alignment with framework, (c) fairness, (d) bias, and (e) accessibility.

Panelists noted that stakeholders are included in NAEP legislation. So, this should be included in one of the principles or as a subprinciple of each principle.

Regarding evidence, the policy itself should not include a laundry list of evidence, but it should state that multiple sources of evidence are needed to establish that items and the assessment are high quality. Panelists noted that the sources of evidence should be addressed in the procedures manual. This will also allow the policy to stand the test of time and be flexible to account for changes over time.

Panelists debated including separate principles for items and assessments or covering both levels in each principle. There was no consensus. As examples, the following are two options for developing and organizing the policy principles.

Option 1

Guiding Principle 1. Individual items/tasks/scoring rubrics:

- Represent the content domain as defined within the framework
  - Items align to and measure the assessment objectives
  - Items are developed in consideration of the ALDs
- Are fair and accessible
- Are accurate in their presentation and free from error
- Include scoring criteria which should be clear, accurate, and explicit.

Guiding Principle 2. The operational pool (test forms):

- Reflects the breadth and depth defined within the framework and reflects the test blueprints
- Allows for inferences consistent with the expectations defined within ALDs
- Is fair
- Is accurate in presentation/content and free from error
- Provides for reporting at the level necessitated by score reports.

Guiding Principle 3. Stakeholders are appropriately embedded in all phases of the items and test development and review process.

Guiding Principle 4. Only data that are necessary will be collected and stored in a way to maintain Personally Identifiable Information (PII).

Option 2

Guiding Principle 1. Reflect the breadth defined within the framework.

- Item
- Assessment

Guiding Principle 2. Achievement levels are considered in all phase of development.

- Item
- Assessment

Guiding Principle 3. Stakeholders are involved in items assessment and review.

- Item
- Assessment

Guiding Principle 4. Test questions and forms, and associated scoring criteria, are accurate in their presentation and free from error.

- Item
- Assessment

Guiding Principle 5. All NAEP test questions shall be free from racial, cultural, and bias.

- Item

- Assessment

Guiding Principle 6. Only collect data that will be necessary to support the assessment goals.

- Item
- Assessment

Appendix A: Meeting Agenda and Attendees

Expert Panel Meeting on NAEP Item Development and Review Policy
National Assessment Governing Board Technical Support Project

August 12-13, 2019 | Agenda

Meeting Overview

The National Assessment Governing Board (Governing Board) oversees and sets policy for the National Assessment of Educational Progress (NAEP), which includes exercising final authority over all assessment items. Broadly speaking, the Governing Board Item Development and Review Policy serves to articulate the principles for quality and fairness of NAEP assessments to all NAEP audiences, which includes: teachers; curriculum specialists; content experts; assessment specialists; state and local administrators; policymakers; business representatives; parents; users of assessment data; researchers and technical experts; and members of the public.

The purpose of this meeting is to obtain expert guidance on plans to update the Governing Board Item Development and Review Policy. A copy of the recently updated Governing Board Framework Development Policy is attached as a secondary reference. Experts in large-scale assessment will discuss stakeholder involvement in item development and review processes, information to consider when developing policy, the role of assessment frameworks in policy development, and reactions to the Governing Board's current Item Development and Review policy, including consideration of which aspects should remain and where the policy may be improved to meet current goals.

DAY 1

9:00 – 9:15 **Welcome, Introductions, and Meeting Goals** — Thanos Patelis

Relevant Readings: Biographies

9:15 – 10:15 **I. Overview of NAEP Item Development and Review Policy** — Michelle Blair

Relevant Readings:
- Governing Board Policy Statement: Item Development and Review
- Governing Board Policy Statement: Framework Development

10:15 – 10:30 **Review & Revise Agenda Topics** — All

10:30 – 10:45 **Break**

10:45 – 12:00 **II. Criteria for Items** — All

Relevant Readings:
- Thought Paper: Involving Stakeholders in Item Review, by Marianne Perie

Guiding Questions:
1. What are the essential types of information to consider when evaluating items? Discuss the importance of the following, generally and in the context of digitally based assessment:
    a. Fidelity to the framework?
    b. Adherence to technical standards, e.g., the AERA/APA/NCME Standards?
    c. Quality of the items?
    d. Fairness?
    e. Accessibility?
    f. The extent to which items reflect the ALDs?
    g. Appropriateness of scoring criteria and procedures?
2. Which types of information should be used in stakeholder reviews of items? How do states ensure this information is considered in these reviews for state assessments?
3. What are criteria that can be established to determine which items may have bias and appropriateness issues, allowing for deeper stakeholder review of those items?

12:00 – 12:45	**Break for lunch**	
12:45 – 2:00	**Criteria for Items: Continuation of Discussion**	All
2:00 – 3:15	**III. Criteria for Assessments**	All

Relevant Readings:

- Thought Paper: A Proposed Policy for Establishing the Overall Validity of NAEP Assessments, by Jeffrey Nellhaus and Edward Roeber

Guiding Questions:

1. What are the essential types of information to consider when evaluating assessments, i.e., the full pool of items to be used in an actual NAEP assessment? Discuss the importance of the following, generally and in the context of digitally based assessment:
    a. Fidelity to the framework?
    b. Representativeness of the domain?
    c. Adherence to technical standards, e.g., the AERA/APA/NCME Standards?
    d. Quality of the assessment?
    e. Fairness?
    f. Accessibility?
    g. The extent to which ALDs are represented?
    h. Appropriateness of intended score interpretations?
2. What information could be built into "assessment reviews" that would help a policy board determine whether there are tensions in the framework or between the assessment and the framework before the operational assessment is administered? After the operational assessment is administered?
3. What evidence must the reviews generate to establish overall validity?
4. How can these reviews add additional evidence in support of validity arguments for NAEP?

3:15 – 3:30	**Break**	
3:30 – 5:00	**Criteria for Assessments: Continuation of Discussion**	All
6:00	**Meet for optional group dinner**	

DAY 2

9:00 – 9:15 **Review of Previous Day and Plan for Today** Thanos Patelis

9:15 – 10:15 **IV. Assessment and Item Review: Procedures and Use of Stakeholders** All

Relevant Readings:

- Thought Paper: Involving Stakeholders in Item Review, by Marianne Perie

Guiding Questions:

1. What are criteria to evaluate appropriateness of assessment and item review procedures?
2. How do states engage various stakeholders in item or assessment review processes?
    a. What types of stakeholders typically participate?
    b. How are these stakeholders identified?
    c. Can you describe attempts at engagement that have been particularly successful? Unsuccessful?
3. How do stakeholders learn how their comments were addressed?
4. Are there specific items or types of items presented to students that have not been subjected to a stakeholder review in state testing programs?
    a. Which type(s) of items?
    b. What is the justification for not reviewing them?

10:15 – 10:30 **Break**

10:30 – 11:45 **Assessment and Item Review: Continuation of Discussion** All

11:45 – 12:30 **Break for lunch**

12:30 – 1:45 **V. Policy Recommendations** All

Relevant Readings:

- Thought Paper: Criteria and Processes for Item Review: Background for Review of Governing Board Item Development and Review Policy, by Stuart Kahl

Guiding Questions:

1. What aspects of the current policy do you feel are important to maintain and why?
2. In what ways can the Governing Board strengthen the policy?
3. Which ideas are worthy of being stated as principles given the goals of NAEP assessments?
4. What advice do you have to move policy towards more of an assessment-level focus rather than the current item-level focus? What are some of the pitfalls we might encounter when making this change, and how might we mitigate them?
5. Are there additional references that reflect widely accepted professional standards?

1:45 – 2:00 **Wrap-up** Thanos Patelis

Attendees

Expert Panelists:
Joan Herman, CRESST
Stuart Kahl, Independent Consultant
Erika Landl, Center for Assessment
Jeffrey Nellhaus, Independent Consultant
Marianne Perie, Measurement in Practice
Ed Roeber, WIDA Consortium/WCER
Deb Sigman, WestEd
Andrew Wiley, ACS

Governing Board Staff:
Michelle Blair
Sharyn Rosenberg
Lisa Stooksberry

HumRRO:
Monica Gribben
Thanos Patelis
Sheila Schultz

NCES:
Holly Spurlock

Item Development and Review Policy Expert Panel Meeting Biographies

Panelists

Joan L. Herman is the Director of University of California, Los Angeles's National Center for Research on Evaluation, Standards, and Student Testing (CRESST). With experience as an evaluator of school reform, she has researched the effects of testing on schools and the design of assessment systems to support school planning and instructional improvement. Her recent work has focused on the validity and utility of teachers' formative assessment practices in science. A former teacher and school board member, Dr. Herman has authored a number of resource books for educators and researchers, the most notable being *Tracking Your School's Success: A Guide to Sensible School-Based Evaluation* and *A Practical Guide to Alternative Assessment*. In addition to publishing in research journals and books, she frequently speaks on evaluation and assessment topics to policy audiences. Currently, Dr. Herman is the editor of the academic quarterly, *Educational Assessment* served on the Joint Committee for the Revision of Standards for Educational and Psychological Testing and is a member NRC's Board on Testing and Assessment. She received her Ed.D. in Learning and Instruction from the University of California, Los Angeles

Stuart Kahl is currently an independent consultant. He was previously with Measured Progress, which he cofounded. contributes regularly to the thought leadership of the assessment community. In recent years, his particular interests have included formative assessment, curriculum-embedded performance assessment, and new models for accountability assessment programs. The Association of Test Publishers (ATP) awarded Dr. Kahl the 2010 ATP Award for Professional Contributions and Service to Testing. He regularly publishes research papers and commentaries introducing and analyzing current issues and trends in education, and as a frequent speaker at industry conferences, Dr. Kahl also serves as a technical consultant to various education agencies. Prior to Measured Progress, he worked for the Education Commission of the States, the University of Colorado, Clark University, and RMC Research Corporation, and was an elementary and secondary teacher. Kahl earned degrees from Johns Hopkins University and from the University of Colorado.

Erika Landl joined the Center for Assessment in July 2012.She has worked with several states to articulate coherent, defensible theories of action aligned to state goals and policy initiatives and has developed user-friendly resources that support the evaluation and refinement of educator evaluation and school accountability systems. She has consulted on the design of innovative assessments, including those for Career Technical Education; generated papers summarizing current practices related to the evaluation of educators in non-tested grades and subjects, and supported individual states and consortia in drafting detailed Requests for Proposal (RFPs). Most recently Erika developed a tools and process to support the evaluation of large-scale summative assessments against CCSSO's Assessment Quality Criteria. Dr. Landl is the lead on multiple technical advisory committees including Pennsylvania and Arkansas, and frequently develops white papers and presentations for technical and non-technical audiences at the national, state and local levels. Erika previously served as a Senior Research Scientist at Pearson, where she was lead psychometrician for a variety of state and national assessment programs. During her 13 years at Pearson, she was responsible for the planning, management and coordination of the full array of psychometric activities necessary to sustain a large-scale assessment program, including test design and development, scaling and equating, item and test analysis, parameter estimation, standard setting, the development of reliability and validity research, report design, and the creation of technical documentation. Dr. Landl received a Ph.D. in Educational Measurement and Statistics from the University of Iowa.

Jeffrey Nellhaus is currently an independent consultant with interests including developing test specifications, assisting with improvements in assessment and accountability, and advising state K-12 and higher education agencies in policy relating to college/career readiness. Dr. Nellhaus previously was with Achieve as the Director, PARCC Assessment. Before joining Achieve, Jeff spent nearly 25 years with the Massachusetts Department of Elementary and Secondary Education where he held the positions of Deputy Commissioner, Acting Commissioner, and Associate Commissioner for Curriculum and Assessment. While at the MA DESE, Jeff directed the design, development and implementation of the Massachusetts Comprehensive Assessment System (MCAS), and the development of the Massachusetts Curriculum Frameworks, which include the Common Core State Standards. For his work on MCAS he was awarded the Manuel Carballo Governor's Award for Excellence in Public Service. Jeff has served on the National Validity Studies Panel to National Assessment of Education Progress (NAEP) and on Technical Advisory Committees for the states of Maine, Kentucky and Rhode Island. He has also served on the Technical Advisory Committee on Standard Setting for NAEP and on the Growth Model Peer Review Panel for the U.S. Department of Education. Jeff holds a bachelor's in Chemistry from the University of Massachusetts, a master's in Science Teaching from Antioch Graduate School of Education, and an Ed.M. in Administration, Policy and Planning from Harvard Graduate School of Education.

Marianne Perie recently started the company Measurement in Practice, LLC, where she is president. Prior, she was a professor at the University of Kansas. Dr. Perie brings 20 years of experience and expertise in K–12 assessment, accountability systems, test validity evaluation and performance standards to CAARD. She has extensive experience providing technical assistance and assisting states in developing valid, reliable and equitable assessments and accountability systems. She has expertise in setting performance standards and has provided technical advice to several states and consortia. She also has taught courses and written extensively on standard setting and has considerable expertise in validity evaluation. Perie has been working to develop strong validity arguments for alternate assessments and has provided several technical assistance workshops to states through the U.S. Department of Education. She is a former director of AAI's Center for Educational Testing and Evaluation.

Ed Roeber currently serves as Senior Assessment Policy Advisor with the WIDA Consortium/WCER at the University of Wisconsin, Madison, WI. In this role, he advises states on improving English proficiency assessments, how to include English learners (ELs) in their academic assessment programs, how to improve the assessment of English proficiency for ELs, and assuring that ELs are included in educational reform efforts at the state and national levels. Before assuming this position in the fall 2012, Dr. Roeber was an adjunct Professor, Measurement and Quantitative Methods, in the Michigan State University College of Education, East Lansing, MI. In this capacity, he taught courses on educational measurement, worked to improve the assessment skills of prospective and current educators, conducted research on how teachers learn to use formative assessment strategies, and provided additional support for faculty and students on assessment. Previously, he was Senior Executive Director, Office of Educational Assessment & Accountability in the Michigan Department of Education from 2003 to 2007. He oversaw the assessments of general education students (in mathematics, science, language arts and social studies), students with disabilities and English language learners, as well as the accreditation and accountability programs. From 1998 to 2003, he was Vice President, External Relations for Measured Progress, a non-profit educational assessment organization located in Dover, New Hampshire. He worked with state policy leaders and staff of state and local education agencies to help design, develop, and implement quality assessment programs. Dr. Roeber earned his PhD. In Measurement and Evaluation from the University of Michigan.

Deb Sigman, a state and national education leader focusing on connecting assessments with enhanced student learning, is the Interim Program Director for the Comprehensive School Assistance Program (CSAP) at WestEd. She also serves as CSAP's Director of State Strategic Support where she collaborates with state, regional, and local education agencies to develop and support a culture of learning in which all students acquire 21st century learning skills. Sigman also is Co-Director of the California Comprehensive Center (CA CC),1 of 15 Regional Comprehensive Centers funded by the U.S. Department of Education. The CA CC provides technical assistance to extend the capacity of the California Department of Education and other state-level stakeholders. In addition, Sigman is the Senior Advisor to the Center on Standards and Assessment Implementation (CSAI), 1 of 7 National Content Centers. CSAI supports states' implementation of their standards and assessment systems. Sigman's experience in education policy, student assessment, and accountability supports professional learning opportunities on interpreting assessment results for educators and their students, and helps school leaders, teachers, and parents better use assessment data to support student learning and achievement. Sigman focuses on the development and delivery of technical assistance regarding the integration of curriculum, assessment, and instruction with an emphasis on training educators in the appropriate and constructive use of assessment results. Prior to joining WestEd, Sigman served in education leadership positions at the district, state, and national levels. She held several administrative positions in California districts and served for nearly 11 years at the California Department of Education, first as the State Assessment Director and then as the Deputy Superintendent of Public Instruction, overseeing assessment, accountability, and various federal grant programs. She holds degrees in psychology and counseling from the University of California and California State University.

Andrew Wiley is a Partner with ACS and specializes evaluation and operational support to ensure that validity evidence adheres to industry best practices. He received his Ph.D. in psychometrics from Fordham University. His experience as the research lead for nationally recognized assessment programs, such as the SAT and the PSAT/NMSQT meant that he was responsible for ensuring the reliability, validity, and fairness of these programs. Drew has also led independent investigations into highly-visible statewide assessment programs as well as classroom-based assessment tasked with evaluating consistency with industry best practices. Currently, Drew is partnering with several assessment stakeholders that are engaging in major redesigns of their existing assessment programs as well as introducing new assessment programs. Drew continues to be actively involved with the measurement community and has served on multiple committees for the National Council on Measurement (NCME) and the Association of Test Publishers (ATP), including a term as the Chair of the Board of Trustees. Drew's research interests include program redesign, product development, and quality assurance practices, and he enjoys the challenge of finding defensible practices with assessment programs considering product changes or introducing new assessments.

Governing Board Staff

Michelle Blair is Assistant Director for Assessment Development, serving as lead staff to the Board's Assessment Development Committee. Ms. Blair has overseen an array of activities for Board initiatives, including framework development for the NAEP Technology and Engineering Literacy assessment and special analyses for the Board's research on academic preparedness for postsecondary pursuits. She has also provided technical leadership on a number of Board projects. Ms. Blair participates in national and regional education conferences on an ongoing basis to showcase Board initiatives, as well as her own research. Ms. Blair joined the Governing

Board in 2007 and previously served as Senior Research Associate, providing technical assistance to the Board's three standing committees with a focus on research and policy matters. Prior to the Governing Board, Ms. Blair taught 6th and 7th grade English language arts in Broward County, Florida, and worked in the New York City Department of Education, addressing youth support programs as well as special initiatives for the Chancellor. Ms. Blair holds a masters degree from Harvard University in Public Administration and a masters degree from the University of Maryland, College Park in Mathematical Statistics.

Sharyn Rosenberg serves as the Assistant Director for Psychometrics, which entails serving the Governing Board's Committee on Standards, Design, and Methodology and providing technical and psychometric expertise to Governing Board projects, including achievement level setting activities. Dr. Rosenberg joined the Governing Board in 2013 with an extensive background in education and psychometrics, as well as NAEP. Her work experiences include Horizon Research, where she conducted complex data analyses and provided psychometric expertise for projects, and the American Institutes for Research (AIR), where she provided research and psychometric support for NAEP. At AIR, Dr. Rosenberg most recently served as the Project Director for the NAEP research and technical support team where she managed and conceptualized NAEP research studies, as well as responded to technical requests from the National Center for Education Statistics' Assessment Division. The focus of her graduate work at the University of North Carolina at Chapel Hill, where she also earned a Certificate in Survey Methodology from the Odum Institute, was on measurement and quantitative methods.

Lisa Stooksberry works with Board staff to support the Governing Board's congressional mandate to set policy for The Nation's Report Card, as Deputy Executive Director. Prior to joining the Governing Board, Dr. Stooksberry served as senior vice president for standards and assessment at the National Board for Professional Teaching Standards. In this capacity, she oversaw the National Board Certification program, a voluntary standards- and performance-based measure of accomplished teaching, available in 25 certificate areas. Dr. Stooksberry's deep background in overseeing the technical aspects of assessment includes item development, assessment alignment and validation studies. A former teacher and teacher educator, Dr. Stooksberry also worked at the American Association of Colleges for Teacher Education (AACTE), where she led partnership efforts among teacher preparation programs and organizations such as the Council of Chief State School Officers and the Stanford Center for Assessment, Learning and Equity at Stanford University. She led AACTE's efforts in a multistate initiative designed to improve teacher preparation through performance-based assessments of pre-service teachers.

HumRRO Staff

Monica Gribben is a Senior Staff Scientist at HumRRO. She has more than 32 years of applied research experience. Her education evaluation projects include evaluations of local, state, and national programs. She has led and participated in evaluations involving formative and summative analyses of the Prince George's County Public School's Alternative Middle School, Alternative High School, Biotechnology Program, and Science and Math Program. Dr. Gribben has worked with the team evaluating the California High School Exit Examination on behalf of the California Department of Education. Dr. Gribben supports the National Assessment Governing Board in conducting literature reviews and expert panels to inform policy decisions associated with the Nation's Report Card. She has conducted analysis of state assessment trends to study the impact of the No Child Left Behind Act for the Center on Education Policy. She has extensive experience in the development, administration, and analysis of surveys and focus groups. Her analytic skills include content analysis of focus group discussions, as well as meticulous data quality control. Dr. Gribben formerly taught high school mathematics and earned a bachelor's degree in psychology at Yale University and master's and doctoral degrees in industrial psychology at George Mason University.

Thanos Patelis is a Principal Scientist and has over 28 years of experience in educational research, measurement, program evaluation, statistical analysis, assessment product development, graduate-level education, and research management. Dr. Patelis regularly advises assessment and educational organizations, institutions, and agencies on assessment design, assessment systems, and psychometrics. He has provided advice and contributed to state accountability systems. He has undertaken many large-scale validation studies and program evaluation efforts in secondary and post-secondary settings. He has supervised teams of measurement experts, statisticians, researchers, and project management professionals including a project management office within a research and development department of a large testing and educational organization. He developed scale anchors and prototypes of reports for a national achievement test battery to provide more meaningful information to educators. He led the development of multiple assessments and indicators using principled approaches to test development. He has designed and implemented a program of evaluation research for educational initiatives implemented in multiple school districts across the country. He has developed and managed data systems, score reports, and data tools for assessment users for national assessment programs. He pioneered efforts to develop assessment literacy and use of classroom-based assessments with teachers; continues to contribute and advocate in support of assessment quality methods and procedures, especially on issues of fairness involving the use of assessments and the need for independent reviews of assessment systems. He has undertaken research to evaluate and enhance the use of statistical analyses of structural equation modeling and coefficient alpha. He has served as chair for the American Psychological Association's (APA) Committee on Psychological Tests and Assessment and he contributes to other professional organizations, including the National Council on Measurement in Education (NCME) and APA's Division 5, Quantitative and Qualitative Methods. He is associate editor for *Applied Measurement in Education, Educational Measurement: Issues and Practice, the Athens Journal on Social Science*, and a reviewer for several journals and professional organizations. He is a member of technical advisory committees for research and assessment. Dr. Patelis has been recognized for his excellence in teaching by APA's Society for the Teaching of Psychology and mentoring by NERA. He has achieved Fellow status for APA Division 5.

Sheila Schultz manages HumRRO's Educational Policy Impact Center (EPIC). She has approximately 30 years' experience conducting assessment, evaluation, and personnel research. Dr. Schultz has developed a variety of data gathering instruments, including paper-pencil and on-line surveys and she has facilitated workshops, focus groups, and interviews to conduct her research. Dr. Schultz has evaluated numerous programs to assess the impact of efforts to improve performance, measure the quality of services and their delivery, and guide the formulation of policies and procedures. She has triangulated assessment results to provide diagnostic feedback and to offer recommendations for performance improvement to both organizations and employees. Dr. Schultz directed HumRRO's NAEP quality assurance contract and she directed the effort to evaluate the high-quality of summative assessments for four programs. Dr. Schultz directed a study that examined alignment of Smarter Balanced's summative assessments to the Common Core State Standards. She led an internal effort to develop prototype instruments to help states/districts build new educator evaluation systems, and she directed an effort to develop and validate a nationally portable work readiness credential for entry-level workers. She directed a series of evaluation studies that examined implementation and outcome measures for Maryland's Prince George's County Public School's Biotechnology, Science and Math, Middle School Alternative, and High School Alternative

Programs. Much of Dr. Schultz's early work involved conducting job analyses for various occupations and integrating the information into selection and performance evaluation systems for agencies such as Drug Enforcement Administration; Federal Aviation Administration; Bureau of Alcohol, Tobacco, Firearms, and Explosives; and Federal Bureau of Investigation. Dr. Schultz obtained a doctoral degree in Educational Research and Evaluation from Virginia Tech and a master's degree in Industrial-Organizational (I-O) psychology from George Mason University. She is a former elementary and secondary school teacher.

Appendix B: NAEP Item Development and Review Policy

Adopted: May 18, 2002

National Assessment Governing Board

Item Development and Review

Policy Statement

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 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 Governing 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 National Assessment of Educational Progress Authorization Act of 2002 (P.L. 107-279) contains a number of important provisions regarding item development and review for 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 Governing 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 Governing Board shall have final authority on the appropriateness of all assessment items;"
- "the Governing Board shall take steps to ensure that all items selected for use in NAEP are free from racial, cultural, gender, or regional bias and are secular, neutral, and non-ideological;" and
- "the Governing 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 Governing 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 NAEP. 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 NAEP 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 widely accepted 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. (2004). Washington, DC: Joint Committee on Testing Practices.

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

Guiding Principles – Item Development and Review

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 test questions that are to be administered in conjunction with any pilot test, field test, operational assessment, or special study administered as part of NAEP.

Principle 5

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

Principle 6

All NAEP test questions shall be free from racial, cultural, gender, or regional bias, and shall be secular, neutral, and non-ideological. NAEP shall not evaluate or assess personal or family beliefs, feelings, and attitudes, nor 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 Governing Board, the framework for each assessment shall 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 that delineates the scope of the assessment relative to the content domain. The framework shall 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 Governing 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 Governing 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, the pool of test questions shall measure 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 shall reflect the objectives, and the item pool shall 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 shall be explained in supporting materials.

6. Supporting material submitted with the NAEP items shall provide a description of procedures followed by item writers during development of NAEP test questions. This description shall 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 and fairness reviews, and assessment development.

7. In submitting items for review by the Governing Board, NCES shall provide information on the relationship of the specifications and the content/process elements of the pool of NAEP items. This shall include procedures used in classifying each item.

8. The item types used in an assessment shall match the content requirements as stated in the framework and specifications, to the extent possible. The match between an objective and the item format shall 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 shall ensure that the development contractor engages a minimum of 20 percent 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 Governing 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 shall 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 Governing 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 shall be used to ensure a match between the descriptions and the resulting NAEP items. The achievement level descriptions shall 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 shall 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 shall 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 guiding statute, a primary duty of the Governing Board pertains to "All Cognitive and Noncognitive Assessment Items." Specifically, the statute states that, "The Governing Board shall have final authority on the appropriateness of all assessment items." Under the law, the Governing 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 Governing 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 Statistics (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 Governing Board policy.

3. Final review of all NAEP test questions for bias and appropriateness shall be performed by the Governing 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 any 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 Governing Board shall receive, in a timely manner and with appropriate documentation, all test

the cycle (from the initial field test pool), the Governing Board shall be informed of this process as well.

8. All NAEP test items shall be reviewed by the Governing Board in a secure manner via in-person meetings, teleconference or videoconference settings, or online via a password-protected Internet site. The Governing Board's Assessment Development Committee shall have primary responsibility for item review and approval. However, the Assessment Development Committee, in consultation with the Governing Board Chair, may involve other Governing Board members in the item review process on an *ad hoc* basis. The Governing Board may also submit items to external experts, identified by the Governing Board for their subject area expertise, to assist in various duties related to item review. Such experts shall 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 Governing Board prior to a second round of pilot or field testing.

10. Documentation of the Governing Board's final written decision on editing and deleting NAEP items shall be provided to NCES within 10 business days following completion of Governing Board review at each stage in the process.

Principle 5

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

Policies and Procedures

1. NCES, through its subject area content experts, trained item writers, and item review panels, shall examine each item carefully to ensure its accuracy. All materials taken from published sources shall be carefully documented by the item writer. Graphics that accompany test items shall be clear, correctly labeled, and include the data source where appropriate. Items shall be clear, grammatically correct, succinct, and unambiguous, using language appropriate to the grade level being assessed. Item writers shall 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 shall 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 shall accompany each constructed-response item. Such criteria shall 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.

questions that will be administered to students under the auspices of NAEP. These items include those slated for pilot testing, field testing, and operational administration.

2. The Governing 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 Governing Board shall not review items being administered as part of test development activities, such as small-scale, informal tryouts 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 Governing Board for review in accordance with a mutually agreeable timeline. Items shall 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 Governing Board shall examine all items prior to the special study, pilot test, or field test stage. In the case of the NAEP reading assessment, all reading passages shall be reviewed by the Governing Board prior to item development. For each reading passage, NCES shall 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 shall provide information and explanatory material on passages deleted in its fairness review procedures.

6. For its second review, the Governing Board shall examine items following pilot or field testing. The items shall 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 Governing Board shall receive a report from the calibration field test stage, which occurs prior to the operational administration. This "exceptions report" shall 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

3. Constructed-response scoring criteria shall 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 shall be expanded to include examples of actual student responses to illustrate each score point. Actual student responses shall be used as well, to inform scorers of unacceptable answers.

4. Procedures used to train scorers and to conduct scoring of constructed-response items shall be provided to the Governing Board, along with information regarding the reliability and validity of such scoring. If the technology becomes available to score student responses electronically, the Governing Board shall be informed of the reliability and validity of such scoring protocol, as compared to human scoring.

Principle 6

All NAEP test questions shall be free from racial, cultural, gender, or regional bias, and shall be secular, neutral, and non-ideological. NAEP shall not evaluate or assess personal or family beliefs, feelings, and attitudes, nor 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 could 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 any individual or group in a demeaning manner would be unacceptable.

2. In conducting bias reviews, steps shall 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 NAEP.

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:

- Secular – NAEP questions shall not contain language that advocates or opposes any particular religious views or beliefs, nor shall 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.”

- Neutral and Non-ideological - Items shall 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;
- analyze the themes of Franklin D. Roosevelt’s first and second inaugural addresses;
- identify the purpose of the Monroe Doctrine;
- select a position on the issue of suburban growth and cite evidence to support this position;
- provide arguments either for or against Woodrow Wilson’s decision to enter World War I;
- 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 Governing 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 Governing Board shall review both stimulus materials and test items to ensure adherence to the NAEP statute and the policies 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 shall not ask a student to reveal personal or family beliefs, feelings, or attitudes, or publicly disclose personally identifiable information.

Adopted: March 3, 2018

National Assessment Governing Board

Framework Development

Policy Statement

It is the policy of the National Assessment Governing Board to conduct a comprehensive, inclusive, and deliberative process to determine and update the content and format of all assessments under the National Assessment of Educational Progress (NAEP). The primary result of this process shall be an assessment framework (hereafter, "framework") with objectives to guide development of NAEP assessments for students in grades 4, 8, and 12 that are valid, reliable, and reflective of widely accepted professional standards.

The Governing Board, through its Assessment Development Committee, shall monitor the framework development and update processes to ensure that the final Governing Board-adopted framework, specifications, contextual variables documents, and their development processes comply with all principles and guidelines of the Governing Board Framework Development Policy.

Introduction

Since its creation by Congress in 1988, the Governing Board has been responsible for determining the content and format of all NAEP assessments. The Governing Board has carried out this important statutory responsibility by engaging a broad spectrum of stakeholders in developing recommendations for the knowledge and skills NAEP should assess in various grades and subject areas. From this comprehensive process, the Governing Board develops a framework to outline the content and format for each NAEP assessment at grades 4, 8, and 12. Development of a framework for a new assessment is guided by the schedule of NAEP assessments adopted by the Governing Board.

Under provisions of the National Assessment of Educational Progress Authorization Act of 2002 (P.L. 107-279), Congress authorized the Governing Board to continue its mandate for determining the content and format of valid and reliable assessments based on widely accepted professional testing standards and active participation of stakeholders. This mandate aligns with the purpose of NAEP, which is to provide fair and accurate measurement of student academic achievement.

Given this mandate, the Governing Board must ensure that the highest standards of test development are employed in framework development to support the validity of educational inferences made using NAEP data. The Governing Board Item Development Policy details principles and guidelines for NAEP assessment items, and the Governing Board has final authority on the appropriateness of all assessment items.

By law, NAEP assessments shall not evaluate personal beliefs or publicly disclose personally identifiable information, and NAEP assessment items shall be secular, neutral, and non-ideological and free from racial, cultural, gender, or regional bias.

To develop the recommended framework for Board adoption, the Governing Board convenes stakeholders to identify the content and design for each NAEP assessment.

In this process, involved stakeholders include:

Teachers	Policymakers
Curriculum Specialists	Business Representatives
Content Experts	Parents
Assessment Specialists	Users of Assessment Data
State Administrators	Researchers and Technical Experts
Local School Administrators	Members of the Public

This Policy complies with the National Assessment of Educational Progress Authorization Act of 2002 (P.L. 107-279) and the documents listed below which express widely accepted technical and professional standards for test development. These standards reflect the agreement of recognized experts in the field, as well as the policy positions of major professional and technical associations concerned with educational testing. A procedures manual shall provide additional detail about how this Policy is implemented.

The Standards for Educational and Psychological Testing. (2014). Washington, DC: American Educational Research Association, American Psychological Association, and National Council on Measurement in Education.

Code of Fair Testing Practices in Education. (2004). Washington, DC: Joint Committee on Testing Practices.

National Center for Education Statistics (NCES) Statistical Standards. (2012).

Principles for Framework Development

Principle 1: Elements of Frameworks

Principle 2: Development and Update Process

Principle 3: Framework Review

Principle 4: Resources for the Process

Principle 5: Elements of Specifications

Principle 6: Role of the Governing Board

Guidelines for the Principles

Principle 1: Elements of Frameworks

The Governing Board is responsible for developing a framework for each NAEP assessment. The framework shall define the scope of the domain to be measured by delineating the knowledge and skills to be tested at each grade, the format of the NAEP assessment, and the achievement levels.

Guidelines

a) The framework shall determine the extent of the domain and the scope of the construct to be measured for each grade level in a NAEP assessment. The framework shall provide information to the public and test developers on three key aspects of the assessment:
   - *What* is to be measured, including definitions of the constructs being assessed and reported upon and descriptions of the purpose(s) of the assessment;
   - *How* that domain of content is most appropriately measured in a large-scale assessment, including the format requirements of the items and the assessment, the content and skills to be tested at each grade, sample items for each grade to be tested, the weighting of the item pool in terms of content and cognitive process dimensions, and any additional requirements for the assessment administration unique to a given subject area, such as provision of ancillary materials and uses of technology; and
   - *How much* of the content domain, in terms of knowledge and skills, should students know and be able to do at the basic, proficient, and advanced levels in achievement level descriptions for each grade to be tested. The achievement level descriptions shall be based on the Governing Board's policy definitions for basic, proficient, and advanced achievement and shall incorporate the content and process dimensions of the assessment at each grade.

b) The framework shall determine the construction of items for each NAEP assessment. The achievement level descriptions in each framework shall also be used in the level-setting process.

c) The framework shall focus on important, measurable indicators of student achievement to inform the nation about what students know and are able to do without endorsing or advocating a particular instructional approach.

d) Content coverage in each subject and grade shall be broad, inclusive of content valued by the public as important to measure, and reflect high aspirations for student achievement. (See *Principle 4* for more detail on the factors balanced in content coverage.)

e) Frameworks shall be written to be clear and accessible to educators and the general public. The framework shall contain sufficient information to inform all stakeholders about the nature and scope of the given assessment. Following Governing Board adoption, the framework shall be widely disseminated.

Principle 2: Development and Update Process

The Governing Board shall develop and update frameworks through a comprehensive, inclusive, and deliberative process that involves active participation of stakeholders.

Guidelines

a) In accordance with the NAEP statute, framework development and update processes shall be fair and open through active participation of stakeholders representing all major constituents in the various NAEP audiences, as listed in the introduction above.

- *Framework panels* shall reflect diversity in terms of gender, race/ethnicity, region of the country, and viewpoints regarding the content of the assessment under development.
- *Public comment* shall be sought from various segments of the population to reflect many different views, as well as those employed in the specific content area under consideration.

b) Framework development and update processes shall be executed primarily via two panels: a Visioning Panel with a subset of members continuing as the Development Panel. This process shall result in three documents: a recommended framework, assessment and item specifications, and recommendations for contextual variables that relate to the subject being assessed. For each framework,

- *The Framework Visioning Panel* shall formulate high-level guidance about the state of the field to inform the process, providing these in the form of guidelines. The major part of the Visioning Panel work will be at the beginning to provide initial guidance for developing a recommended framework. The Visioning Panel shall be comprised of the stakeholders referenced in the introduction above. At least 20 percent of this panel shall have classroom teaching experience in the subject areas under consideration. This panel may include up to 30 members with additional members as needed.
- *The Framework Development Panel* shall develop drafts of the three project documents and engage in the detailed deliberations about how issues outlined in the Visioning Panel discussion should be reflected in a recommended framework. As a subset of the Visioning Panel, the Development Panel shall have a proportionally higher representation of content experts and educators, whose expertise collectively addresses all grade levels designated for the assessment under development. Educators shall be drawn from schools across the nation, including individuals who work with students from high-poverty and low-performing schools, as well as public and private schools. This panel may include up to 15 members, with additional members as needed.

c) In addition to a recommended framework, the framework development or update process shall result in assessment and item specifications (see *Principle 5*) and recommendations on related contextual variables to be collected from students, teachers, and school administrators. Recommendations shall take into account burden, cost, quality of the data to be obtained, and other factors. (See the Governing Board Policy on Collecting and Reporting Contextual Data.)

d) The scope and size of a framework development project shall determine the size of framework panels and the number of panel meetings needed. A framework update project may require smaller panels and fewer meetings if a smaller scope is anticipated for recommended revisions. Each project shall begin with a review of major issues in the content area. For a framework update, the project shall also begin with an extensive review of the current framework, and the Visioning Panel shall discuss the potential risk of changing frameworks to trends and assessment of educational progress. (See *4.b*).

e) Framework development and updating shall be comprehensive in approach and conducted in an environment that is open, balanced, and even-handed. Panels shall consider all viewpoints and debate all pertinent issues in formulating the content and design of a NAEP assessment, including findings from research. Reference materials shall represent multiple views.

f) For each project, protocols shall be established to support panel deliberations and to develop a unified proposal for the content and design of the assessment. Written summaries of all hearings, forums, surveys, and panel meetings shall be made available in a timely manner to inform deliberations.

Principle 3: Framework Review

Reviews of existing frameworks shall determine whether an update is needed to continue valid and reliable measurement of the content and cognitive processes reflected in evolving expectations of students.

Guidelines

a) At least once every 10 years, the Governing Board, through its Assessment Development Committee (ADC), shall review the relevance of assessments and their underlying frameworks. In the review, the ADC shall solicit input from experts to determine if changes are warranted, making clear the potential risk of changing frameworks to trends and assessment of educational progress. The Board may decide based on the input that the framework does not require revision, or that the framework may require minor or major updates. To initiate updates, the ADC shall prepare a recommendation for full Board approval. Minor updates include clarifications or corrections that do not affect the construct defined for the assessment. Major updates shall include the convening of a Visioning Panel (see *Principle 2*). Framework revisions shall also be subject to full Board approval.

b) Within the 10 year period for an ADC review, major changes in the states' or nation's educational system may occur that relate to one or more NAEP frameworks. In this instance, the ADC will determine whether and how changing conditions warrant an update and the Governing Board via recommendation may convene a Visioning Panel to revise or replace the framework. Before framework panels are convened, special research and analysis may also be commissioned to inform the updates to be considered.

c) If the Visioning Panel recommends major updates, then a subset of panel members shall continue as the Development Panel to develop the draft framework and assessment and item specifications, in accordance with *Principle 2*. Regular reports will be provided to the ADC and the recommended framework update shall be subject to full Board approval.

d) When a framework update is conducted, framework Visioning and Development Panel recommendations shall describe the extent to which adjustments in the achievement level descriptors (see *1.a*) and contextual variables (see *2.c*) are needed. (See the Governing Board Policy on Achievement Levels and the Governing Board Policy on Collecting and Reporting Contextual Data for additional details.)

Principle 4: Resources for the Process

Framework development and update processes shall take into account state and local curricula and assessments, widely accepted professional standards, exemplary research, international standards and assessments, and other pertinent factors and information.

Guidelines

a) The NAEP framework development and update processes shall be informed by a broad, balanced, and inclusive set of factors. The framework shall reflect current curricula and instruction, research regarding cognitive development and instruction, and the nation's future needs and desirable levels of achievement. This delicate balance between "what is" and "what should be" is at the core of the NAEP framework development process.

b) An initial compilation of resources shall summarize relevant research, advantages and disadvantages of the latest developments, and trends in state standards and assessments for the content area. This compilation shall also summarize how stakeholders have used previous NAEP student achievement trends in the assessment area. The compilation may include public comment. Using this compilation as a springboard, framework panel deliberations shall begin by thoroughly identifying major policy and assessment issues in the content area.

c) The framework panels shall also consider a wide variety of resources as deliberations proceed, including but not limited to curriculum guides and assessments developed by states and local districts, widely accepted professional standards, scientific research, other types of research studies in the literature, key reports having significant national and international interest, international standards and assessments, other assessment instruments in the content area, and prior NAEP frameworks, if available.

d) Technical experts shall be involved to uphold the highest technical standards for development of the NAEP framework and specifications. As a resource to the framework panels, these experts shall respond to technical issues raised during panel deliberations.

e) In balancing the relative importance of various sources of information, framework panels shall consider direction from the Governing Board, the role and purpose of NAEP in informing the public about student achievement, the legislative parameters for NAEP, constraints of a large-scale assessment, technical assessment standards, issues of burden and cost-effectiveness in designing the assessment, and other factors unique to the content area.

Principle 5: Elements of Specifications

The specifications document shall be developed for use by NCES as the blueprint for constructing the NAEP assessment and items.

Guidelines

a) The assessment and item specifications shall produce an assessment that is valid, reliable, and based on relevant widely accepted professional standards. The specifications shall also be consistent with Governing Board policies regarding NAEP design, such as groupings of items, test administration conditions, and accommodations for students with disabilities and English language learners (see the Governing Board Policy on NAEP Testing and Reporting on Students with Disabilities and English Language Learners). The specifications shall be reviewed by technical experts involved in the process, prior to submission to the Governing Board.

b) The primary audience for the specifications, or assessment blueprint, shall be NCES and the contractor(s) responsible for developing the assessment and the test questions.

c) The specifications shall evolve from the framework and shall be written in sufficient detail so that item writers can develop high-quality questions based on the framework objectives for grades 4, 8, and 12, where applicable. The specifications shall include, but not be limited to detailed descriptions of:

- the content and process dimensions, including the weighting of those dimensions in the pool of questions at each grade;
- types of items;
- guidelines for stimulus material;
- types of response formats;
- scoring procedures;
- achievement level descriptions;
- administration conditions;
- ancillary or additional materials, if any;
- considerations for special populations;
- sample items, including a substantial number and range of sample items with scoring guidelines for each grade level; and
- any unique requirements for the given assessment.

d) Special studies, if any, to be conducted as part of the assessment shall be described in the specifications. This description shall provide an overview of the purpose and rationale for the study, the nature of the student sample(s), and a discussion of the instrument and administration procedures.

Principle 6: Role of the Governing Board

The Governing Board, through its Assessment Development Committee, shall monitor all framework development and updates. The result of this process shall be recommendations for Governing Board action in the form of three key documents: the framework; assessment and item specifications; and contextual variables that relate to the subject being assessed.

Guidelines

a) The Assessment Development Committee (ADC) shall be responsible for monitoring framework development and updates that result in recommendations to the Governing Board on the content and format of each NAEP assessment. The ADC will provide direction to the framework panels, via Governing Board staff. This guidance shall ensure compliance with the NAEP law, Governing Board policies, Department of Education and government-wide regulations, and requirements of the contract(s) used to implement the framework project.

b) When a framework Visioning Panel is to be convened, the ADC shall develop a charge for the panel, and the charge shall be subject to full Board approval. The charge will outline any special considerations for an assessment area.

c) The ADC shall receive regular reports on the progress of framework development and updates.

d) In initiating a framework update, the Governing Board shall balance needs for stable reporting of student achievement trends. Regarding when and how an adopted framework update will be implemented, the Board may consider the NAEP Assessment Schedule, cost and technical issues, and research and innovations to support possibilities for continuous trend reporting.

e) At the conclusion of the framework development or update process, the Governing Board shall take final action on the recommended framework, specifications, and contextual variables. The Governing Board shall make the final decision on the content and format of NAEP assessments.

f) Following adoption by the Governing Board, the final framework, specifications, and contextual variables shall be provided to the National Center for Education Statistics (NCES). These documents, which include the achievement level descriptions for Basic, Proficient, and Advanced performance definitions, are provided to NCES to guide development of NAEP test questions and questionnaires.

See material sent under separate cover

Changes to NAEP Mathematics, Reading, Civics, and U.S. History Questionnaires

On or before February 26, 2020, the Assessment Development Committee (ADC) will receive material ***under separate cover*** regarding proposed changes to contextual questionnaires for the NAEP Mathematics, Reading, Civics, and U.S. History questionnaires.

Recent contextual questionnaire data suggests it is possible that student burden may be too high. As a result, students may not have enough time to complete all the items included in the questionnaire. Responsively, NCES has proposed several items be "rotated off" to support streamlining of the questionnaires. These items may be considered for possible reinstatement in future administrations, but trend reporting will be disrupted for these particular questionnaire items.

James Deaton of NCES will summarize the proposed changes and address the Committee's questions. As the ADC reviews the detailed proposed changes (in the February 26, 2020 memo; sent under separate cover), a central issue for this discussion is: How can the Board ensure that the most critical questionnaire items for contextualizing student achievement are retained?

ADC Activities in the Strategic Vision

The ADC develops recommendations for what NAEP should assess and exercises final authority over all NAEP items. Several activities in the Governing Board Strategic Vision call for the ADC's leadership. These projects involve informing educators, updating policies, and exploring new approaches to framework updating, as well as projects to review and update frameworks as needed.

Notable Accomplishments under the Current Strategic Vision

- SV5: Revised Board policy on Framework Development (approved March 2018)
- SV5: Conducted framework reviews for NAEP Reading and Mathematics Frameworks
- SV5: Implemented a NAEP Mathematics Framework Update (to be presented for full Board action at the upcoming November 2019 Board meeting)
- SV5: Streamlined reviews of NAEP assessment items, paving the way for a portfolio of work devoted to NAEP frameworks

Remaining Priorities for the Current Strategic Vision

- SV3: Develop a set of principles to guide questionnaire revisions in ways that reflect the Board's expectations for how NAEP data should be used (in conjunction with the Reporting and Dissemination Committee)
- SV5: Implement a NAEP Reading Framework Update (to be presented in August 2020)
- SV5: Revise the Board Item Development and Review Policy with additional linkages to framework reviews
- SV8: Determine how a review other countries' assessment programs should inform frameworks, framework processes, contextual data, and reporting

A working draft of ADC's project plans is attached, along with a summary of common elements for each framework project and ongoing committee discussions relevant to the Strategic Vision. At the conclusion of the ADC's March 2020 session, Chair Dana Boyd will invite questions on these resources.

WORKING DRAFT* PLAN: ALL ADC STRATEGIC VISION (SV) ACTIVITIES

UPDATES SINCE NOVEMBER 2019 ARE HIGHLIGHTED

ACTIVITY	START	FINISH	STATUS
Identify NAEP Resources & Information for Educators (*SV #3 Expanding NAEP Resources* and *SV #6 Contextual Variables*)	May 2017	Nov 2021	ADC discussed NAEP Questions Tool and contextual variables in 2017. Suggestions for new or refined NAEP resources can be shared with R&D for Board outreach. In March 2019, the ADC discussed development of a set of principles to guide questionnaire revisions in ways that make them actionable, reflecting the Board's expectations for how NAEP data should be used. To be determined: when/how to develop ADC recommendations. In August 2019, the Committee discussed the Questions Tool and the NAEP Data Explorer as resources for educators.
Update Framework Development Policy	Jun 2017	Mar 2018	ADC began revising policy in Summer 2017. Board discussion continued in November 2017. Board adopted the revised policy in March 2018.
Review & Update **Mathematics** Framework for 2025 Assessment	Aug 2017	Mar 2025[1]	State math standards review began in August 2017. Results were shared in May 2018 ADC Framework Review, which also engaged external expert commentary. ADC prepared a framework recommendation for Board action, and it was unanimously adopted in August 2018. The framework contractor[2] for the Math Framework Update project was secured in Summer 2018. The Board reviewed a draft framework when public comment was collected in Spring 2019 and continued discussion in August 2019. The Board adopted the framework in November 2019 and Board action on the specification is slated for March 2020, allowing NCES to conduct development leading to a 2025 administration of the updated assessment. The framework will be posted online with a summary of the update process in Spring 2020 and a special note from the Governing Board, based on the Board's 2019 policy deliberations.

* All timelines are estimated. This draft will be updated based on Board policy decisions. All activities address *Strategic Vision Priority #5 Updating Frameworks*, unless otherwise noted. Factors contributing to the sequencing of framework projects include how recently the last framework update was conducted, staff capacity, timing of the next administration on the NAEP Assessment Schedule, and urgency of the update.

[1] Timeline includes administering the assessment.

[2] The mathematics framework project will be implemented by the same contractor as the reading framework project, on staggered schedules so that most of the mathematics project is completed by the time the reading project begins.

WORKING DRAFT* PLAN: ALL ADC STRATEGIC VISION (SV) ACTIVITIES

UPDATES SINCE NOVEMBER 2019 ARE HIGHLIGHTED

ACTIVITY	START	FINISH	STATUS
Review & Update **Reading** Framework for 2025 Assessment	Oct 2017	Mar 2025[1]	ADC Framework Review was held in March 2018 to inform development of recommendations for a Fall 2019 framework update project launch. In August 2018, the ADC prepared a draft framework recommendation, which the Board adopted in March 2019. The Board will review policy issues at the conclusion of the public comment period in August 2020. Board action on the framework is slated for November 2020, allowing NCES to conduct development leading to a 2025 administration of the updated assessment.
Explore New Approaches to Framework Update Processes (also *SV #8 International Assessments*)	Nov 2017	Aug 2023	The Board's Technical Services contractor is developing several resources to assist in exploring innovations in how NAEP assessment updates are implemented. Framework Update Projects will review other countries' assessment programs to inform frameworks, framework processes, contextual data, and reporting.
Update Item Development and Review Policy	Aug 2018	Mar 2021	The ADC began discussing goals for the policy revision in August 2018. In 2019, an expert panel was convened to gather insights regarding best practices in assessment development. An initial policy discussion is slated for March 2020.
Review & Update **Civics** and **U.S. History** Frameworks	Mar 2018	TBD	Discussion of outreach began in March 2018, with suggestions to develop options for the ADC to consider. In August 2018, ADC review of the current NAEP item pools indicated that framework revisions did not need to be fast-tracked. Framework reviews will begin after the framework review for the NAEP Science Framework, based on the NAEP Schedule of Assessments.
Review & Update **Science** and **Technology & Engineering Literacy (TEL)** Frameworks	Mar 2018	TBD	Discussion of outreach began in March 2018, Tentative next steps: learn more about standards in NGSS non-adopter states and learn whether stakeholders view that some or all of the TEL subarea on Technology & Society addresses student achievement goals in Civics or U.S. History. Framework reviews for the NAEP Science Framework are slated for Fall 2020.
Review & Update **Writing** Framework	TBD	TBD	Initial discussion regarding the Writing Framework is slated for 2021.

WORKING DRAFT* PLAN: ALL ADC STRATEGIC VISION (SV) ACTIVITIES

UPDATES SINCE NOVEMBER 2019 ARE HIGHLIGHTED

ACTIVITY	START	FINISH	STATUS
Develop Content Descriptions for the Long-Term Trend (LTT) Mathematics and Reading Assessments (*SV #7 Long-Term Trend*)	TBD	TBD	March 2018 Executive Committee deliberations on LTT called for ADC to develop content descriptions of the assessments to support LTT item development, as well as updates to the Governing Board LTT policy and improved explanations of LTT assessment goals. ADC requested these descriptions also illuminate knowledge and skills of lower performing students, if possible. NCES has already developed a list of measurement objectives for LTT Mathematics, and similar work may be possible for Reading. Board staff is using these inputs to begin development of the LTT content descriptions.

Common Elements of Each Framework Update Project

Based on the revised Framework Development Policy, several milestones address all NAEP assessment framework projects. Framework update projects engage stakeholders and content experts to identify needed revisions, via subject-specific factors including:

- Evolution of discipline and implications for NAEP frameworks
- Relevance to students' postsecondary endeavors
- Student achievement trends in terms of contextual factors
- Digital-based assessment issues
- International content and measurement trends

MILESTONES: ALL FRAMEWORK PROJECTS

ADC Discussion with External Experts in the Subject Area(s)
ADC Recommendation for Updating Assessment
Board Action on Charge
Framework Contractor Selection
Trend Scan & Resource Compilation
Panel Meetings (3 to 6)
Full Board Review & Public Comment
Framework Draft Finalized
ADC Final Review of Framework
Board Action
Assessment Administered

As a first step, the ADC conducts a framework review, where content experts are invited to a Committee session to provide reflections on the state of the discipline and the extent to which the relevant NAEP framework should be updated. Studies and additional outreach is pursued, as needed, to inform the ADC's recommendation about the type of framework update that is required. Next, the ADC brings its recommendation to the full Board for approval. In the case of an anticipated framework update, the recommendation includes a charge to stakeholders who will serve on the panels convened to draft recommendations for the ADC's consideration.

After Board discussion of the ADC recommendation, the Board will take action on the charge. Concurrently, Board staff will identify a contractor to execute the framework update process.

The framework contractor will launch the project by identifying individuals to serve on the framework panels and by compiling and developing resources to support the meetings of these stakeholders. A subset of these resources will include the Governing Board's charge to the framework panels as well as documents used to inform the Board's development of the charge. The first meeting of stakeholders will be for the *Visioning Panel* to discuss the major issues to be addressed in the framework. A subset of the Visioning Panel will continue on as the *Development Panel* to develop an updated framework. This panel will also develop the recommended updates to the Test and Item Specifications, as well as the Contextual Variables.

The ADC monitors the framework contractor's work via regular project updates. A draft of the panels' recommended framework will be shared for full Board review and public comment, as well as review by the Board's Committee on Standards, Design and Methodology. This feedback will allow the Development Panel to address concerns and finalize the draft framework, specifications, and contextual variables for the ADC's final review and Board action. The adopted framework, specifications, and contextual variables are given to NCES to begin assessment development, piloting, and finally administration of the operational assessment based on the new framework.

Ongoing Committee Discussions

Recent ADC discussions have raised several issues for ongoing consideration as the Committee leads Strategic Vision activities and prepares content recommendations for Board deliberation and action (November 2019 additions highlighted):

- The optimal role of NAEP for each content area.
- How Board and Committee priorities should be reflected in upcoming framework updates.
- How to clarify the context of each framework and how the Board has chosen to navigate this context.
- Extent to which current frameworks are flexible enough to adapt as needed.
- How to improve the transition periods between old and new frameworks for a subject area, e.g., providing guidance on how to gradually reflect framework updates in the assessment.
- How to provide strategic guidance about how a framework's objectives are sampled, when they cannot all be covered in a single administration of the assessment.
- The level of specificity in assessment results that is most useful to policymakers, researchers, and educators.
- How future NAEP items will be a resource for the field.
- Expected gains and losses to the field for each NAEP framework decision.
- How to ensure that Governing Board framework policies and procedures are followed.
- How to establish and maintain partnerships that highlight actionable aspects of results, e.g., teacher access to released NAEP items and contextual information.
- How to develop viable options for new configurations of NAEP assessment content in ways that balance expertise, outreach, research, and trends in curricular standards.
- How to incorporate how other countries think about changing what they assess.
- Whether to more deeply assess an existing content area or add new content areas.
- Whether streamlining of NAEP frameworks is an appropriate goal.
- How to be intentional about content overlap between different assessments, while fulfilling statutory requirements, e.g., biennial reading and mathematics assessment.