Assessment Development Committee

July 13, 2020

WebEx, Eastern Time

AGENDA

3:00 – 3:50 pm	**Briefing: Reaping the Rewards of Reading for Understanding** *Amy I. Berman, National Academy of Education* *Annemarie S. Palincsar, University of Michigan* *P. David Pearson, University of California, Berkeley* *Gina Biancarosa, University of Oregon* *Panayiota Kendeou, University of Minnesota*	*Attachment A*
3:50 – 4:40 pm	**Update: 2025 NAEP Reading Framework** *Mark Loveland and Cynthia Greenleaf, WestEd*	*Attachment B*
4:40 – 5:00 pm	**Action: 2025 NAEP Mathematics Assessment and Item Specifications** *Dana Boyd, Committee Chair* *Mark Miller, Committee Vice Chair* *Michelle Blair, Assistant Director for Assessment Development*	*Attachment C*
5:00 – 5:30 pm	**Discussion: The Next Strategic Vision** *Dana Boyd and Mark Miller*	*Attachment D (Under separate cover)*

Briefing: Reaping the Rewards of Reading for Understanding

On June 9, 2020, the National Academy of Education released a new report: *Reaping the Rewards of the Reading for Understanding Initiative* (also attached). This report synthesizes findings from scholarship conducted over the past decade on reading comprehension. The report was developed through funding from the U.S. Institute of Education Sciences (IES), supporting six research teams. Culling and reviewing over 200 scholarly articles, the six research teams summarized findings and common themes about the nature and development of reading comprehension, as well as how it is assessed and how it is reflected in curriculum and instruction in pre-K through grade 12.

There were two major precursors to this report:

1. IES launched the *Reading for Understanding Initiative* to respond to concerns that children's improvement in reading comprehension had leveled off over the previous few decades.
2. Scholars observed that research on reading comprehension had sufficiently matured, and much of this research was a direct result of the *Reading for Understanding Initiative*.

Together, these developments encouraged IES to support funding for a major effort to synthesize research findings, yielding this groundbreaking and comprehensive report. Report highlights include the importance of: emphasizing comprehension in pursuit of knowledge and insight; redoubling our efforts to enhance language development, both oral and written, for students across the age-span; and changing the culture of classrooms to emphasize collaboration, deep comprehension, critique, and use of comprehension. There is also an additional paper that is a conceptual review of technology-related reading comprehension research, as well as research related to multimodal meaning-making (both digital and nondigital) and reading comprehension in out-of-school contexts.

On July 13, 2020, the Governing Board's Assessment Development Committee (ADC) will host the author team of this timely report for a briefing. Notably, the report's author team includes 2025 NAEP Reading Framework Panel Chair P. David Pearson and others illustrious reading scholars.

ADC members will have the opportunity to ask questions about the report as well as questions about how the public comment draft of the 2025 NAEP Reading Framework relates to the report's key takeaways.

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, in accordance with the Governing Board Achievement Levels Policy
- 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 Driving Panel Discussion for the Reading Framework Update

As noted previously, the Panel has focused on drafting framework recommendations in response to several major issues:

- Reflecting new theoretical and research-based understandings;
- Updating texts and tasks to reflect contemporary aspirations;
- Maintaining separate NAEP Reading and NAEP Writing assessments per NAEP legislation, while addressing the increasingly integrated instruction and assessment of reading and writing;
- Accounting for the interplay between knowledge and reading comprehension;
- Optimizing the use of digitally-based assessment; and
- Representing students' reading achievement more equitably.

Preparing for Full Board Policy Discussions on the Framework Update

On June 22, the 2025 NAEP Reading Framework Update Project entered the public comment phase. A draft of the Framework Development Panel's recommendations has been posted for public feedback.

Before the Panel begins the next round of revision for the framework draft, the Panel must understand: what is the Board's consensus on major policy issues? This will enable the Panel to revise the draft framework to reflect the Board's consensus, while they also revise the draft framework to reflect the consensus from public comments regarding purely content matters. These revisions will pave the way for the Board to be able to support the framework draft and eventually adopt it.

In this session, the ADC will be briefed on the Development Panel's progress since March 2020, the public comment phase of the framework update project, and how the public comment draft is positioned on policy issues that have emerged to date, namely:

1. Reflecting components of social emotional learning as part of the construct of reading comprehension and the contextual questions in the reading-specific survey questionnaires;
2. Integrating questionnaire items into the assessment, rather than solely presenting these items in separate questionnaires;
3. Providing choice to students as they take NAEP, e.g., in selecting passages, tasks, etc.;
4. Using commissioned texts (authored by test developers) to augment authentic texts (found in the public domain);
5. Considering unintended uses of new subscales for the NAEP Reading Assessment (i.e., reporting subscores for reading in literature, science, and social studies)
6. Including more race by socioeconomic status breakouts in initial reporting of NAEP results; and
7. Disaggregating English learners into more categories (including a new "former EL" grouping).

Additional policy issues will likely be raised in public comment, which closes on July 23.

On July 31, the ADC will lead a full Board to discussion to determine: what guidance does the Governing Board need to provide to the Panel on the policy issues that have been raised?

The July 13 ADC session and a planning meeting on July 29 are opportunities for the Committee to determine how to lead the July 31 full Board policy deliberation on major policy issues for this framework update.

Quarterly Report

2025 NAEP READING FRAMEWORK UPDATE

WestEd

July 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. To complete this work, WestEd is partnering with the Council of Chief State School Officers (CCSSO), who will assist 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 will also come from project collaborators: the Literacy Research Association (LRA), the International Literacy Association (ILA), and the National Council of Teachers of English (NCTE).

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 will formulate guidelines for developing a recommended framework, based on the state of the field. Seventeen members of the Visioning Panel constituted 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. The Development Panel has met four times to date, with a fifth and final meeting scheduled for September, to complete revisions to the framework in response to public comment and related Board policy guidance.

Upon completion of draft versions of the framework documents in June 2020, the project has transitioned to the public comment phase, to be conducted primarily by WestEd and CCSSO, with assistance from collaborating organizations. Feedback on the draft documents will come from member organizations represented on the two panels, other stakeholder organizations, and the public. WestEd staff will tabulate feedback, make initial recommendations for revisions addressing the feedback, and coordinate the development of final versions of the framework documents, to be submitted to the Governing Board by November 2020.

Progress to Date

Panel Activities

Panel activities have been successfully conducted around the Visioning Panel meeting in October 2019 and the first four Development Panel meetings. The most recent Development Panel meeting, conducted virtually on March 16-18, 2020, focused on resolving issues in order to complete drafts of individual framework chapters, developing initial draft achievement level descriptions (ALDs), and organizing the panel for the remaining work required to meet the timeline for developing the Public Comment Framework draft. Following the Development Panel meeting, panelists worked in individual chapter writing teams to complete final drafts of each framework chapter. In late April 2020, chapter writing teams submitted drafts of each chapter for review and then conducted chapter reviews across the framework. Following an editorial review of the Framework by the project team and panel chair, an initial framework draft was reviewed by Governing Board staff and the National Center for Education Statistics (NCES) in late May/early June 2020. The Development Panel met virtually in early June 2020 to discuss feedback on the initial draft and proposed revisions. The Public Comment Framework draft was developed and posted for public review on June 22, 2020.

The Development Panel is currently engaged in the public comment period. During the first week of public comment, Visioning Panel members (inclusive of Development Panel members), met to discuss the draft and to begin collecting feedback. Panelists are also participating in outreach webinars as co-presenters.

TAC Activities

The TAC has met on five occasions, to provide feedback on the project deliverables and to respond to questions from the first four Development Panel meetings. TAC feedback on these deliverables and questions were reported at subsequent Development Panel meetings. Most recently, the NAEP Reading TAC meetings #4 and #5 were convened on May 5 and June 2, 2020. The primary focus of the recent TAC meetings was to review and provide feedback on the complete draft NAEP Reading Framework and ALDs. Feedback and responses from the TAC informed the development of the Public Comment Framework draft in June 2020.

Next Steps

Panel Activities

The Development Panel will continue to engage in outreach activities during the public comment period, from June 22 to July 23, 2020. Following the public comment period, the Development Panel will meet again in a series of virtual meetings in August and September to consider public feedback, related Board policy guidance, and suggested revisions to the framework draft and to resolve any remaining issues needed to finalize the framework documents for submission to the Governing Board in mid-November. Panelists are also contributing to the development of the Assessment and Item Specifications, which will be drafted in July-October 2020 and finalized following Governing Board and NCES review in November-December 2020.

Drafts of the Updated Framework

The next drafts of the framework will include revisions in response to feedback received during the public comment period and the Governing Board's July 31 policy discussion. These next drafts will be the subject of the final Development Panel meetings in August and September 2020. The final version of the 2025 NAEP Reading Framework will be submitted to the Governing Board in mid-November 2020.

Outreach

Outreach activities are underway. Outreach activities will be conducted through July 23, 2020 and will serve 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 will aim 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 are in the process of soliciting written feedback from their member organizations and participating in outreach webinars. WestEd is actively soliciting feedback from additional stakeholder organizations through outreach webinars aimed at specific stakeholder audiences. Each outreach webinar is co-hosted by a stakeholder organization and includes discussion prompts to stimulate engagement from each of the target stakeholder groups. In all instances, groups will 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 is also conducting outreach webinars to solicit feedback on draft versions of the framework documents through its extensive membership network.

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	March 2020 – June 2020
Gather Public Comment	June 2020 – July 2020
Develop Final Versions of Framework Documents	July 2020 – November 2020
Develop Final Versions of Specifications Document	November 2020 – January 2021
Submit Final Process Report	February 2021

Reading Framework for the 2025 National Assessment of Educational Progress

****Public Comment Draft****

National Assessment Governing Board
U.S. Department of Education

Table of Contents

NAEP Overview and History

The National Assessment of Educational Progress (NAEP), often called The Nation's Report Card, is the largest nationally representative and continuing assessment of what students in public and private schools in the United States know and are able to do in various subjects. Since 1969, NAEP has been a common measure of student achievement across the country in mathematics, reading, science, and many other subjects. NAEP report cards provide national, state, and some district-level results, as well as results for different demographic groups. NAEP is a congressionally mandated project of the National Center for Education Statistics (NCES), located within the U.S. Department of Education's Institute of Education Sciences. The National Assessment Governing Board, an independent, bipartisan organization made up of governors, state school superintendents, teachers, researchers, and representatives of the general public, sets policy for NAEP.

The Education Sciences Reform Act of 2002 (P.L. 107-279) is the governing statute of NAEP. This law stipulates that NCES develops and administers NAEP and reports NAEP results. Under the law, the Governing Board is given responsibility for setting the assessment schedule, developing the frameworks that provide the blueprint for the content and design of the assessment, and setting achievement levels. The NAEP Reading Assessment is given every two years in English to students in grades 4 and 8, and every four years at grade 12. The assessment measures reading comprehension by asking students to read selected grade-appropriate materials and answer questions based on what they have read. The results present a broad view of students' reading knowledge, skills, and performance over time.

NAEP assessments are unique in the United States in that they are not reported by student, by school, or by district, with the exception of the 27 districts that participate in the NAEP Trial Urban District Assessment (TUDA). 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, for districts participating in TUDA, and for various demographic groups.

To address the increased role of technology in education and society, NCES began transitioning NAEP from paper and pencil to digitally based assessments, with the first digital assessment of NAEP Reading administered in 2017. NCES is utilizing established best practices for NAEP to remain at the forefront of innovation for large-scale assessments and exploring new testing methods and question types to reflect the growing use of technology in education. NCES furnishes all needed hardware as well as a tutorial to ensure equity across administrations.

Development of the 2025 NAEP Reading Framework

The process of developing the 2025 NAEP Reading Framework Reading Framework was guided by Governing Board policies that specify that the work be undertaken by a Visioning Panel of educators and other key stakeholders in education. From this group, a subset of members continued as the Development Panel to finalize a document to recommend to the Governing

Board for approval (See Appendix X). After a review of the current Reading Framework adopted in 2005, the Board determined that a framework update was needed. In 2019, the Board adopted a charge for the Visioning and Development Panels that would develop recommendations for the framework update. The charge was as follows: "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. 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."

In their deliberations, the Visioning Panel and the Development Panel considered the remarkable developments in the literacy world that have taken place since the most recent revision of the NAEP Reading Framework in 2009. Accordingly, the Visioning Panel set guidelines for the Development Panel to design a new Framework that would

- Expand the construct of reading;
- Expand the definition of text;
- Extend the range of comprehension tasks that require knowledge application, including writing from sources;
- Augment and expand the cognitive targets and the approach to reporting performance on them;
- Expand how language structures and vocabulary are defined and measured; and
- Include, measure, and report on the role of engagement in reading performance.

In addressing these guidelines, the Development Panel also considered frameworks for other literacy assessments, such as those for the Programme for International Student Assessment (PISA) and the Progress in International Reading Literacy Study (PIRLS), and responded to other educational and societal developments. These include advances in technology; an explosion of digital and multimodal texts; recent research in disciplinary literacy; greater attention to affective and motivational dimensions of reading; increased diversity of student populations; the adoption of new standards; innovations in digitally based assessments; and new understandings of human development and learning, including sociocultural perspectives and universal design for learning (UDL). These developments, briefly discussed below, guided the creation of the new Framework.

Technological innovations have brought about changes in the format of texts as well as approaches to reading. Sabatini (2020) underscores the scope of this change: "[d]igital forms of literacy are reshaping the genres and nature of literacy practices, and consequently the construct of reading comprehension in the 21st century" (p. 1). Researchers are identifying the ways that online reading capability is both similar to and distinct from reading text printed on paper (Coiro, Lankshear, Knobel & Leu, 2014; Singer & Alexander, 2017). The proliferation of information sources requires students to exercise critical judgment about source relevance, rhetorical effectiveness, trustworthiness, and perspective.

Additionally, 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 language and vocabulary in such literacy and learning (LaRusso, et al, 2016). Furthermore, **affective 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). Recent work on factors such as self-efficacy, growth mindsets, metacognition, and self-regulation impacting performance demonstrate that these factors are also be relevant and important to measure (Dweck & Molden, 2005; Farrington, et al., 2012; Hall, 2016; Taylor, Oberle, Durlak & Weissberg, 2017).

Importantly, 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 (Business Roundtable, 2017; National Center on Education and the Economy, 2013). At the same time, texts are inevitably cultural and political in nature, drawing on frames of reference that may not be universally shared (Lafontaine, Baye, Vieluf & Monseur, 2015; Wexler, 2018). Recent studies demonstrate that readers draw on multiple dimensions of language and vocabulary knowledge, extending understandings of the role of language and vocabulary in meaning making (e.g. LaRusso, et al., 2016). And new understandings of translanguaging (Pacheco & Miller, 2016; Pacheco & Smith, 2015)—how meaning-making engages multiple linguistic and cultural processes for bilingual and biliterate readers—have provided insights about how NAEP reading assessment might ensure a more equitable assessment for all U.S. children and youth.

Most prominent in **new standards** (e.g., Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects; College, Career, and Civic Life (C3) Framework for Social Studies State Standards; Next Generation Science Standards; Texas Essential Knowledge and Skills) 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, lending support for increasing opportunities for students to write both short and extended constructed responses on the NAEP reading assessment. New standards also uniformly emphasize the multimodal nature of reading, including using a variety of text types to conduct research, critique sources, and communicate understanding (Breakstone, McGrew, Smith, Ortega & Wineburg, 2018; Leu, Kinzer, Coiro, Castek & Henry, 2017). Similarly, in reading literature, students are expected to analyze and appreciate how authors use literary devices and elements of craft to achieve literary goals such as analyzing multiple interpretations of a text (e.g., story, drama, poem).

As standards have been updated, a number of **new reading assessments** have been developed to assess them. Smarter Balanced, the Partnership for Assessment of Readiness for College and Careers (PARCC), PISA, PIRLS, and the Global, Integrated Scenario-Based Assessment (GISA) are examples of tests that made efforts to instantiate new standards. (GISA was developed under the U.S. Department of Education's Reading for Understanding Initiative.) Unique features of this generation of new assessments include synthesis across multiple texts, technology enhanced items, items with multiple correct answers, and multimodal features. 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.

Digitally based assessments offer new possibilities for the range of texts and the types of tasks used in reading assessment and allow universal design features to be made available to all students (NCES, 2020). In 2017, the NAEP Reading Assessment introduced real-world, problem-based scenarios that include dynamic texts, videos, animation, and innovative item types and formats. These developments also include building avatar-enriched social contexts for reading and presenting tasks with specific purposes (e.g., propose a school policy, write an article for a school newspaper) for student engagement in reading across multiple texts. Digitally based assessments are also able to capture richer information about student performance, including how students approach and solve problems, locate information, and use their time (NCES, 2018).

Perhaps most significantly, recent **research on human learning and development** in a wide range of fields has highlighted the social and cultural nature of learning and development and the unique assets each learner brings to the learning task (National Research Council, 2018). Further, advances in **universal design for learning** have enhanced opportunities for all students to participate in learning and assessment (CAST, 2018).

New Emphases and Features of the 2025 NAEP Reading Framework

Based on these developments, the 2025 NAEP Reading Framework rests on a revised definition of reading, which is outlined in Chapter 2. Following from that definition, the Framework calls for new emphases and features in the assessment. These include:

An Emphasis on Equity

The Framework takes an asset-based orientation to readers and a "best-foot-forward" stance in designing and administering the assessment, including increasing opportunities for a diversity of students to "see themselves" in text selections and tasks. Efforts to optimize all students' performance as readers and learners focus on selecting relevant purposes, activities, texts, and tasks, as well as providing a range of supports, both ordinary and technology enhanced, to better replicate the reading activities in which readers engage in today's world and to allow for students' own resources to be leveraged while they engage in active meaning making from text. Increased attention to students who are English learners (EL) will be systematic throughout the design of the 2025 NAEP Reading Assessment and is evident in the reporting of EL students' results. Accordingly, to the extent possible within a constrained assessment environment, NAEP reading activities are designed to reflect the multiple and diverse communities, purposes, and texts with which readers engage in their school, work, and everyday lives.

Sociocultural Model of Reading Comprehension

The Framework is based on a sociocultural model of reading that states that, as a human meaning-making activity, reading is always situated in social and cultural contexts that shape every aspect of readers' engagement with text and influence how readers respond to and learn from the experience of reading. The 2025 NAEP Reading Assessment construct reflects this understanding by transitioning to testing blocks that are even more highly contextualized than in previous NAEP Reading Assessments. The 2025 NAEP Reading Assessment assumes and attempts to build on the cultural assets (knowledge, skills, and experiences) that all students bring.

Activity Structure and Purposes

A major shift for the 2025 NAEP Reading Framework is the infusion of explicit purposes into the assessment activities to guide students' reading and responses as they take NAEP. Testing blocks are intentionally and transparently situated within both an activity structure that implies a broad purpose and a disciplinary context. Activity structures are organized by two overarching purposes:

- Reading to Develop Understanding
- Reading to Solve a Problem

Activity structures for the assessment are also organized by disciplinary contexts:

- Reading to Engage in Literature
- Reading to Engage in Science
- Reading to Engage in Social Studies

The 2025 NAEP Reading Assessment simulates disciplinary contexts of reading by providing students with activities to respond to literary, science, or social studies texts that are read both in school and outside of school and that include everyday texts. The activities also include tasks students undertake to address the purpose and result in consequences that can be internal to the reader (e.g., I learned something that makes me feel more knowledgeable) or a product (e.g., I created an informational poster about food safety).

Nature and Characteristics of Texts

Building on the range of literary and informational texts students have encountered previously in NAEP Reading, the Framework for 2025 greatly expands the text types and textual environments with which students will interact in the NAEP digital platform. Multimedia and multimodal texts are increased within static, dynamic, and complex textual environments, including single and multiple texts.

Scaffolds

A dramatically expanded component in the 2025 NAEP Reading Assessment is the systematic use of scaffolds to support students in navigating the assessment and performing to their potential on the assessment. These include features already in use, such as look backs, avatars for scenario-based tasks, and resetting. Scaffolds provide support for knowledge, for metacognition (the ability to monitor one's reading) and strategy, and for motivation.

An Expanded View of Vocabulary

The 2025 NAEP Reading Framework introduces a new and expanded construct, *Language Structures and Vocabulary*, that goes beyond measuring knowledge of individual word meanings to also include knowledge of language structures. This means that instead of measuring only knowledge of single words or expressions, the 2025 NAEP Reading Framework focuses also on readers' understanding of how different words and structures are connected to advance the ideas in a passage. To that end, the NAEP will assess students' understanding of:

> (a) discourse structures (relations across words and phrases) as used in a passage to advance ideas. For example, this might include readers' understanding of which logical

relation between two ideas is communicated in a passage through a particular expression (e.g., as a consequence); or whether two expressions in a passage refer to the same person, (e.g., the king and the monarch);

(b) meanings of individual words or expressions as used in a passage (as traditionally measured by NAEP in the past); and

(c) word structures (word parts) as used in a passage to convey related but distinct meanings. This might include readers' understanding of how morphologically related words in a passage convey meaning (e.g., naturalist and nature as conveying related but different meanings).

This expanded view of language understanding is included in the Framework so that systematic attention is paid to this crucial component of reading comprehension in the design of assessment items. However, performance on items that measure Language Structures and Vocabulary will not be reported independently. Instead, across the four comprehension targets, items may be coded for whether they focus on Language Structures and Vocabulary or not. This will allow for the examination of the contribution of language knowledge to reading achievement, while maintaining an assessment design that views language and comprehension as inextricably interrelated (Nagy & Townsend, 2012; National Research Council, 2010; Uccelli, et al., 2015).

Comprehension Targets

The cognitive targets of the previous framework have been reconceptualized as *Comprehension Targets* and expanded from three to four targets. In keeping with the sociocultural model, the terminology has been changed to signal the broader array of processes that comprise reading and meaning making. The Comprehension Targets ***Locate and Recall*** and ***Integrate and Interpret*** retain the labels from the previous assessment. The former target Critique and Evaluate is renamed ***Analyze and Evaluate*** to more accurately reflect students' processes in completing tasks reflecting this aim. The new target ***Use and Apply*** was added to reflect contemporary understandings that comprehension is best viewed as a series of productive processes that culminate in forms of personal and social production—what is referred to as outcomes or consequences across the framework. Use and Apply items are designed to provide information about how readers transform their understanding of previously read texts into new knowledge that can be acted on and presented to others.

Reporting of Results

The 2025 NAEP Reading Assessment continues the affordances of the prior reporting system, but calls for several expansions, all of which enhance the explanatory capacity of NAEP:

- Reporting assessment data by literary, science, and social studies contexts, highlighting the prominence of the disciplinary grounding of reading
- Identifying ELs as current or former ELs consistent with federal legislation to better reflect the variability of English language proficiency within this population in reporting
- Expanding explanatory variables (formerly called contextual variables) to include block-specific measures and enhanced analysis of process data. (See Measuring Explanatory Variables below.)

These proposed changes to the NAEP Reading Assessment reporting system are intended to allow for analyses that lead to better supports for students and improved learning opportunities.

Including Explanatory Variables

Another major shift in the 2025 NAEP Reading Assessment is greater emphasis on explanatory variables in the analysis and reporting of NAEP reading performance. The aim is to *explain* student performance, in addition to reporting trends over time. This shift presents an opportunity to understand students' performance in terms of the diverse circumstances and the contexts in which they live and learn. Such a shift is consistent with sociocultural framing rather than deficit characterizations of the performance of particular student populations. The Framework calls for two sets of explanatory variables, both of which are important:

- **Reader attributes** index the knowledge, interest, motivation, engagement, habits, attitudes, language competence, and skills/strategies that individual students bring to the reading act
- **Environmental variables** provide perceptions about the contexts that influence individual student performance, some emanating from home and community contexts (e.g., home language, family SES, parent education, participation in community activities) and others related to the school environment (e.g., opportunities to learn, school and classroom supports for learning, peer relationships)

Increased emphasis on explanatory variables can occur in three locations in the assessment:

- **Survey Responses** – Continued use of survey items as a part of a common questionnaire to gather both general and reading specific views from students about their in and out of school experiences
- **Block-Specific Measures** – A new form of student survey items to assess an array of reader attributes related to performance within a specific assessment block
- **Process Variables** – Data collected (such as keystroke trajectories, navigation through the passages and items, and the amount of time spent on particular pages or images) used to draw plausible inferences about motivation and engagement and metacognitive behaviors

Organization of the Framework

The following chapters outline the Framework in detail, including the theoretical basis of reading comprehension—specifically the sociocultural model—to be assessed and more detailed information about the assessment construct, design, and reporting.

- Chapter 2 presents the **sociocultural model** of reading comprehension for the 2025 NAEP Reading Assessment. As described in chapter 1 and based on recent research from an array of fields, the model begins with the assumption that reading comprehension is a fundamentally sociocultural communication process.
- Chapter 3 builds upon the sociocultural model from Chapter 2 to define and describe the key components of the **assessment construct** and defines what is in the NAEP Reading Assessment.
- Chapter 4 describes how the elements fit together in a coherent **assessment design** and details how the components interact with one another to form assessment activities to be administered to students.
- Chapter 5 summarizes how the results of the NAEP Reading Assessment will be reported in 2025 and how **reporting** will provide greater explanatory value based on contextual variables gleaned from students' testing behaviors and surveys.

Exhibit 1.1. Key Similarities and Differences Between the 2009-2019 and the 2025 NAEP Reading Frameworks

	2009–2019 NAEP Reading Framework	**2025 NAEP Reading Framework**
Theoretical Framework	Reading comprehension as a dynamic cognitive process	Reading comprehension as a dynamic cognitive process expanded to a sociocultural model that positions the reader, the text, and the activities in a sociocultural context.
Definition of Reading Comprehension	Reading is an active and complex process that involves: ● Understanding written text. ● Developing and interpreting meaning. ● Using meaning as appropriate to type of text, purpose, and situation.	Reading comprehension is a sociocultural process in which individuals use language, knowledge, and foundational skills to extract, construct, integrate, and critique meaning as they engage with a wide range of texts for purposes shaped by home, community, and school experiences.
Purposes for Reading	No explicit purposes assumed for all assessment tasks.	Purpose-driven assessment includes two broad purposes: ● Reading to develop understanding ● Reading to solve problems
Contexts for Reading	Practical, academic, and other contexts drawn from grade-appropriate sources spanning the content areas.	● Reading to engage in literature ● Reading to engage in science ● Reading to engage in social studies
Content (Type of Texts)	● Literary text ○ Fiction ○ Literary nonfiction ○ Poetry ● Informational text ○ Exposition	● Literary texts ● Science texts ● Social studies texts The range of text types includes the textual elements

	○ Argumentation and persuasive text ○ Procedural text and documents	that characterize texts in each disciplinary context. See exhibit 4.7.
Cognitive Processes	Cognitive targets distinguished by text type: ● Locate/recall ● Integrate/interpret ● Critique/evaluate	Comprehension targets distinguished by context: ● Locate and recall ● Integrate and interpret ● Analyze and evaluate ● Use and apply
Language Structures and Vocabulary	Systematic approach to vocabulary assessment with potential for a vocabulary subscore.	Systematic approach to vocabulary expanded to go beyond measuring knowledge of individual words' meanings to also include knowledge of language structures. The construct includes three dimensions: ● Discourse (relations across words and phrases) ● Semantic (words) ● Morphological (word parts) Assessment items may be double scored for both 1) comprehension and 2) language structures and vocabulary; no subscore for language structures and vocabulary is proposed.
Passage Source & Selection	Use of authentic stimulus material plus some flexibility in excerpting stimulus material. Expert judgment and use of at least two research-based readability formulas for passage selection.	Criteria for including texts in the NAEP reading assessment, regardless of the discipline in which a given block is situated, is: ● Authenticity ● Engagingness ● Social and cultural diversity ● Developmental appropriateness

		● Degree of content elaboration ● Disciplinary appropriateness ● Complexity ● Quality and coherence Flexibility to include some commissioned texts if it is impossible to find naturally occurring texts. Both disciplinary expertise and deep knowledge about the nature and structure of text to be used in the text selection process. Evaluation of text complexity based on a combination of quantitative and qualitative measures, as well as reader attributes and navigational complexity.
Passage Length	Grade 4: 200–800 words Grade 8: 400–1,000 words Grade 12: 500–1,500 words	Grade 4: 200–800 words Grade 8: 400–1,000 words Grade 12: 500–1,500 words
Role of Technology	Transition to digital platform beginning in 2017. No detailed description of how technology should be used relative to the construct.	Digital platform for the entire assessment and affordances of digital interface woven into development of the construct. Real-world, problem-based scenarios that include dynamic texts, videos, animation, and innovative item types and formats. These developments also include building avatar-enriched social contexts for reading and presenting purposeful tasks.

		Text structures include single static on screen text, single dynamic text, and multiple texts (or complex textual environments).
Scaffolds		Three types of scaffolds to support all students within the digital platform: ● Knowledge scaffolds ● Metacognitive and strategy scaffolds ● Motivational and social scaffolds
Item Type	Selected-response and both short and extended constructed-response items included at all grades.	Selected response items, short and extended constructed-response items, and dynamic response items at all grade levels.
Reporting	Reporting subscales for literary and informational texts	Expansion to include: ● Reporting subscales for literary, science, and social studies contexts, highlighting the prominence of the disciplinary grounding of reading ● Further disaggregating students by English language proficiency into three reporting categories, including current, former, and never English learners
Explanatory Variables	Contextual information enriches reporting of results. Contextual variables selected to be of topical interest, timely, and directly related to academic achievement. They	Greater emphasis on explanatory variables organized in two sets: ● Reader attributes related to the knowledge, interest, motivation, engagement, habits, attitudes, language

	may reflect current trends, such as use of technology.	competence, and skills/strategies that individual students bring to the reading act ● Environmental variables related to contexts that influence individual student performance, some emanating from home and community settings (e.g, funds of knowledge, home language, family income, parent education, participation in community activities, and the like) and others related to the school environment (opportunities to learn, school and classroom supports for learning, peer relationships) Variables can be assessed in three ways: ● Core and Reading-specific survey responses ● Block-specific measures ● Process variables

Chapter 2 lays the foundation for the 2025 NAEP Reading Framework by describing the theoretical model of the reading process on which the framework is based. This model will inform all aspects of the NAEP Reading Assessment detailed in the chapters that follow—from the assessment construct (Chapter 3), to its design (Chapter 4), to its reporting and interpretation (Chapter 5).

A Sociocultural View of Human Learning

The 2025 NAEP Reading Assessment frames reading comprehension within a sociocultural context. This framing is the natural outgrowth of recent understandings about the social and cultural nature of all learning and human development. A broad consensus has emerged across the multiple disciplines of the learning sciences—including psychology, developmental studies, anthropology, linguistics, cognitive science, and even biology—recognizing the central role of culture in the lifelong learning process. A report of the National Academies highlights the sociocultural nature of learning.

> "Not surprisingly, embrace of sociocultural theory led to one of the most important recent theoretical shifts in education research: the proposition that all learning is a social process shaped by and infused with a system of cultural meaning (Nasir & Hand, 2006; National Research Council, 2009; Tomasello, 2016). … Human development, from birth throughout life, takes place through processes of progressively more complex reciprocal interactions between the human individual (an active, biopsychological organism) and that individual's immediate physical and social environments. Through these dynamic interactions, culture influences even the biological aspects of learning." (National Academies of Sciences, Engineering, and Medicine, 2018, pp. 27-28)

The sociocultural model of reading comprehension informing the 2025 NAEP Reading Framework is not only warranted by scientific understandings of human development and learning, it is the outgrowth of earlier and current work in reading comprehension (Anderson & Pearson, 1984; RAND Reading Study Group, 2002; Pearson, et al., 2020) and previous NAEP reading frameworks. Like all human learning, learning to read is a meaning-making activity imbued with socially and culturally specific characteristics and practices. Scholars have generated a considerable amount of evidence supporting the view that literacy, including reading, is carried out for very specific social and cultural purposes (Heath, 1983; Lee, 1997, 2007, Moll & Gonzalez, 1994; New London Group, 1996; Rand Reading Study Group, 2002; Scribner & Cole, 1981). These foundational investigations have contributed new insights into the varied ways in which literacy is learned, used, and adapted within families, communities, and schools. Understanding literacy as a social practice has complicated earlier views conceptualizing literacy as an individually acquired discrete set of cognitive skills (Street, 1984).

A key insight in this work is that students are able to demonstrate greater literacy achievement when they are able to draw on funds of knowledge and social practices acquired in their homes and communities, contexts rich with varied forms of knowledge, practices, and competencies (Moll, 2014). A wide range of students in U.S. schools are able to read and achieve at higher levels when invited to enact literacy learning experiences reflecting their home and community funds of knowledge. In studies by Moll and his colleagues (Moll & Gonzalez, 1994; Gonzalez, Moll, & Amanti, 2005), students learning to read in English were able to demonstrate higher comprehension when offered opportunities to draw on their language resources in Spanish in the process of comprehending English texts. Similarly, Lee (1997, 2006, 2007) conducting research within a cultural modeling framework, has demonstrated that when African American students can draw on the skills they use to read and interpret the texts of their everyday lives, they are able to engage in sophisticated interpretation and analysis of literature.

Scholars concerned with the language and literacy development of culturally and linguistically diverse youths have drawn attention to the fact that language practices are continuously changing. These changes occur in response to shifting cultural demographics of local communities and through digital interactions in multilingual global contexts (New London Group, 1996; Paris, 2009). Likewise, scholars have pointed out the new social, functional, and political demands that individuals and families encounter as a result of immigrant and transnational lifestyles that help generate a wealth of linguistic, intercultural, and sociopolitical world knowledge (Dorner, Guan, Nash, & Orellana, 2016; Enciso, Volz, Price-Dennis, & Durriyah, 2010; Sanchez & Orellana, 2006; Skerrett, 2015). Recent scholarship has described the social knowledge-building practices constituting the academic disciplines and the specific role that literacy plays in them (e.g. Goldman, et al., 2016). Moje (2015) argues that students must learn to capitalize on their own "funds of knowledge" as they learn to navigate the social and cultural norms of disciplinary literacy practices in school and beyond.

Taken together, this scholarship suggests that such multilingual, multicultural, and transnational knowledge constitute resources that can be drawn upon to teach and support student reading development and achievement. In arguing for a human development perspective toward reading comprehension, and by extension its sociocultural nature, Lee states that "…understanding how and under what circumstances humans comprehend and interrogate texts within and across contexts…" is essential. Absent such a highly contextualized framework, Lee goes on to argue, "it is much easier to conceptualize … diversity as a problem rather an opportunity" (Lee, 2020, p. 39).

The students in U.S. schools live and learn in a wide range of contexts—urban, rural, or suburban--and bring a wide spectrum of experiences and knowledge to reading comprehension practices. The students who take the NAEP Reading Assessment built from the 2025 NAEP Reading Assessment Framework will represent a wide range of communities of different ethnic, cultural, and linguistic strengths and in-and out-of-school experiences. Therefore, acknowledging the sociocultural perspective in the construction of the assessment will optimize students' ability to draw on what they know and can do in this measure of their reading comprehension.

Overview of the Model

The sociocultural model of reading comprehension identifies three key components—**reader, text, and activity**—and situates them within a sociocultural context to acknowledge the multiple ways all three elements are shaped by the affordances and constraints operating in the contexts in which reading takes place. This new model:

- Illustrates how what **readers** know, do, and understand from reading is tied to the variations in knowledge, skills, and experiences they bring to their reading from experiences at home, in their communities, and in school.
- Embraces the sociocultural nature of **texts**, including who reads and writes them and under what circumstances, how they are generated, their varied forms, and the ways they are used.
- Emphasizes reading comprehension **activities** that appreciate how readers integrate reading with their other communication practices—especially writing, listening, and speaking—to address wide ranging purposes and contexts.

Defining Reading

The 2025 NAEP Reading Assessment rests on a **sociocultural model** of reading, and defines reading comprehension as follows:

> Reading comprehension is a sociocultural process in which individuals use language, knowledge, and foundational skills to extract, construct, integrate, and critique meaning as they engage with a wide range of texts for purposes shaped by home, community, and school experiences.

Readers draw on a range of resources in making sense from text: (a) what they know about a topic, (b) what they know about texts and how they work, and (c) many internal processes needed to render text sensible, including foundational letter- and word-level processes, such as phonological awareness and letter-sound knowledge, along with cognitive processes, such as attention, working memory inferential reasoning, and comprehension monitoring, and (d) socially and culturally situated practices that emerge from the contexts of home, community, and school.

This sociocultural model seeks to account for the dynamic relations between people's participation in cultural practices and the array of resources they draw on for meaning making with texts. In this model, readers participate in a range of cultural communities, including the family, peer affinity groups, community settings, and schools, each of which both constrains and expands the practices students use to understand what they read.

The definition of reading for the 2025 NAEP Reading Assessment assumes that reading comprehension changes across settings and over time for readers as a function of their purposes for reading, the demands of reading activities they are asked to complete, and the resources they have acquired from all the contexts in which they live their lives. Recognizing such differences in readers and their experiences, the sociocultural model was developed to explain how readers orchestrate all of these affordances and constraints to optimize their performance, including their performance on assessments like NAEP. Informed by the sociocultural view of reading, the 2025 NAEP Reading Assessment seeks to provide a valid and relevant assessment experience for the nation's increasingly linguistically and culturally diverse groups of students.

Shift in Perspective for NAEP 2025

This new framework builds on and expands the previous model of reading that served as the basis for the NAEP Reading Assessment administered from 2009 through 2024 by adding a sociocultural perspective. The 2009-2024 framework emphasized the dominant role that readers' internal cognitive processes played in the reading process. The 2025 NAEP Reading Assessment continues the emphasis on readers and texts, but it places more emphasis on the central role of the context in which reading activity unfolds, including the assessment context itself. These influences include the "sets of values, beliefs, experiences, communication patterns, teaching and learning styles, and epistemologies inherent in the students' cultural backgrounds and the socioeconomic conditions prevailing in their cultural groups" (Solano-Flores & Nelson-Barber, 2001, p. 555). Pertaining to reading, these influences include a reader's history of engagement with text (i.e., how often they read, type of texts they read, purposes for reading, etc.) as well as history of participation in cultural communities. Most influential in developing these personal histories is how those communities value and integrate reading and texts into their lives for different purposes (e.g., share stories, recite poetry, reading for information) and in different ways (e.g., in silence, jointly).

The definition for the 2025 NAEP Reading Assessment is consistent with those of other large-scale reading assessment frameworks. For example, PIRLS "…highlights the widespread importance of reading in school and everyday life and acknowledges the increasing variety of texts in today's technological world" (Mullis & Martin, 2019, p. 6). Likewise, the 2018 PISA Reading Framework states, "Reading is no longer considered an ability acquired only in childhood during the early years of schooling. Instead it is viewed as an expanding set of knowledge, skills, and strategies that individuals build on throughout life in various contexts, through interaction with their peers and the wider community" (OECD, 2019, p. 27).

Key Assumptions of this Model

The uniqueness of readers is, along with context, the central focus of the 2025 NAEP Reading Assessment. Each reader uses reading to engage with and make meaning from texts found in school, work, and everyday life. Each reader's text interactions are influenced not solely by the immediate situation, but also by unique combinations of in-school and out-of-school experiences with reading, writing, and language that they have accumulated over time, shaping their practices and identities. Moreover, each reader brings a diverse—and unique—collection of cognitive, linguistic, socio-emotional, and metacognitive resources gathered from their participation in multiple social and cultural communities (Cervetti & Wright, 2019; González, et al., 2005; National Academies of Sciences, Engineering, and Medicine, 2018; Pearson & Cervetti, 2015).

Readers Leverage Personal and Cultural Resources. Readers draw from their own abilities, knowledge, experiences, and identities to consider how these resources are relevant for thinking and learning from texts as they make meaning from them. For instance, bilingual readers may use what they know about reading in one language to read in their other language (García & Godina, 2017). Readers also select and use tools and resources around them to access information and acquire and demonstrate their knowledge (Gutiérrez, et al., 2009).

Readers Engage with Diverse Texts. Texts can take many forms, and differences in texts pose varying challenges to readers. Any reading experience may, for example, include texts from different disciplines (e.g., literature, science, social studies), genres (e.g., a poem, an essay, a story), modalities (e.g., static print or onscreen texts, non-linear hypertexts with embedded links, interactive online texts), and cultural traditions, which vary in terms of the people, points of view, and experiences that are and are not represented, advantaging some readers and disadvantaging others in accessing, comprehending, and learning from those texts (Bråten, Braasch, & Salmerón, 2020; Singer & Alexander, 2017). Ideally, as readers grow, they develop skills that allow them to comprehend and use texts that are not well aligned with their knowledge and experience (Lee, 2005).

Readers Engage in Reading Activities for Varied Purposes. Readers complete reading comprehension activities to address purposes that vary according to readers' understandings of the contexts within which they find themselves. Social and cultural practices shape the reading activities readers experience and take part in (Vygostky, 1978). These reading activities have varied purposes that entail particular tasks and processes, which give rise to specific outcomes of the reading and consequences for the reader (Pearson, Valencia, & Wixson, 2014; Roth & Lee, 2007). An especially important facet of activity is the set of processes and practices that readers employ as they carry out these activities. Readers make meaning while employing a range of strategies for making sense of text, including bringing their knowledge of the world to their reading. They share meaning in ways that reveal what they have understood and learned. Depending on their resources and motivations, they monitor and adjust meaning-making strategies so that they can address specific purposes and contexts (Kintsch, 1998, 2019; Skerrett, 2020).

The Model Underlying the 2025 NAEP Reading Framework

This section elaborates on the definition of reading comprehension for the 2025 NAEP Reading Framework outlined above. The following sections describe the reader, text, and activity components of this model in more detail.

The Sociocultural Nature of Reading Comprehension

As noted earlier, recent research on learning and development provides a clear warrant for the claim that learning is enhanced when students can leverage familiar social and cultural practices in school-based learning. That is, learning is inherently social; individuals learn through interactions with others. Learning is also cultural; families, peers, religious affiliations, and other cultural groups all provide students with resources--funds of knowledge--that shape learning. But the influences do not stop there; learning, including opportunity to learn, is shaped by the economy and the political environment as well. Even individuals' perceptions of how they are treated in connection with their reading shape their identities and agency (Cantor, et al., 2019; Hruby & Goswami, 2019; Hruby, et al., 2011; Immordino-Yang & Gotlieb, 2017; Johnson,et al., 2016; National Academies of Sciences, Engineering, and Medicine, 2018; Osher, et al. 2020).

In sum, reading is a human meaning-making activity that is inescapably shaped by sociocultural forces and fundamental to human development. Sociocultural contexts profoundly affect the development of students' reading comprehension, including how they perform on reading comprehension assessments like NAEP. These forces shape every aspect of the act of reading: the resources (in the form of knowledge, experience, and cultural practices) that readers bring to reading, as well as what readers read, how and why they read, and what they do when they engage in meaning-making.

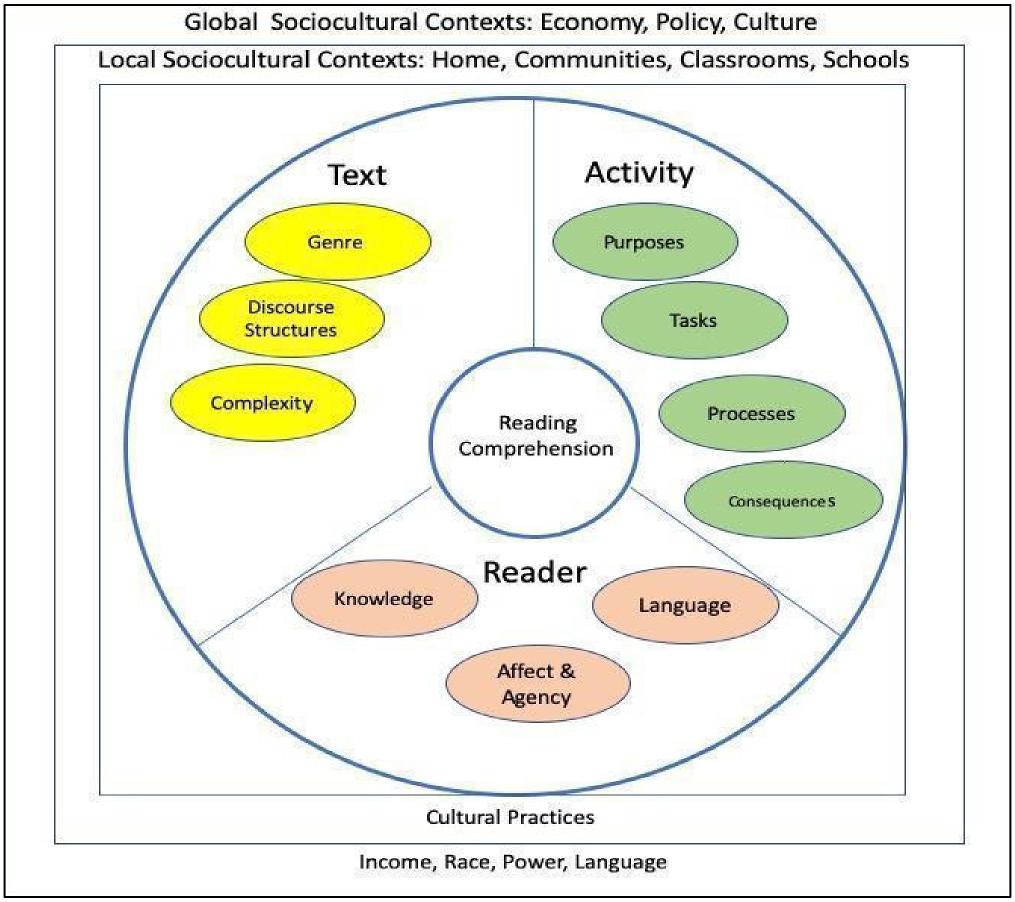
Current research documents how reading comprehension is a socioculturally constructed literacy practice (Frankel Becker, Rowe, & Pearson, 2016; Purcell-Gates, Duke, & Stouffer, 2016). It is a dynamic process involving readers in drawing on what they know to understand, interpret, and use information from texts to complete activities that are shaped by and situated in sociocultural contexts. Valencia, Wixson, and Pearson (2014, pp. 272-273) provide an example of this dynamism:

> A text such as Steinbeck's *Of Mice and Men,* with its bursts of conversation expressed in short, choppy sentences and reliance on the conversational register of ordinary folk, has a quantitative readability of second/third grade. Even so, educators would not use it before middle or high school because of the complex nature of the themes of friendship, loyalty, and intellect described by Steinbeck. Eighth-grade students might readily provide a plot summary of the work, but an analysis of the authorial tools that Steinbeck uses to unveil those complex themes would prove difficult for even sophisticated middle school students (pp. 272-273).

On the other hand, it is possible that a student in grade 4 whose family are migrant workers might be well positioned to critique the novel's themes in ways that more experienced readers in higher grades would not. In addition, 8th graders who have been taught to identify structural elements of literature may bring resources that enable them to weigh the value of Steinbeck's authorial tools.

As this example shows, reading comprehension involves a dynamic interaction among the reader, the texts they read, and the activities they complete as they read—all of which are socioculturally influenced and visually depicted in Exhibit 2.1. In the remainder of Chapter 2, this dynamic interaction is described in detail.

Exhibit 2.1. The Sociocultural Model of Reading



The sociocultural model underlying NAEP 2025 demonstrates how readers, texts, and activities are all influenced by sociocultural contexts and interactively engaged when reading comprehension occurs.

Sociocultural View of the Reader

Central to the NAEP sociocultural model of reading is the reader. Each reader uses reading to engage in school, work, and everyday life. Readers' interactions with texts are responsive to immediate situations: the particular contexts (studying for a physics test), texts (it is a formula-laden set of directions for understanding physical forces), and tasks (I need to turn in a two-page summary of a 10-page section of the chapter) they face in the moment. But those encounters are also shaped by the reader's tool kit—the diverse collection of sociocultural, cognitive, linguistic, affective, and metacognitive resources that individual readers bring to each reading situation. These resources are inherently cognitive, drawn on as readers mentally engage in comprehending text.

Readers' text interactions are also decidedly sociocultural, acquired from and refined by participation in the many communities of practice in which readers live and work (Cervetti & Wright, 2019; Moll, et al., 1992) These resources develop in relation to the unique set of experiences with reading, writing, and language that each reader accumulates over time. All readers select and use tools and resources encountered in community spaces to build their knowledge in response to texts they encounter in and out of school (Gutiérrez, et al., 2009).

Readers bring multiple resources to their reading. These include an interplay of biology (the capacity of the brain to process and archive human experience), cognition (the capacity of the brain to store and process the knowledge archived in memory), and culture (the values and practices of communities). Those factors that most influence the process of reading comprehension are knowledge, language, and agency, described in more detail in the sections that follow.

Readers' Knowledge. While readers' knowledge resources are sociocultural in origin, helpful distinctions can illuminate types of knowledge used in reading comprehension. Based on the reading task or goal, and on the reader's development, multiple sources of knowledge may be operating at any given moment as students read, including declarative, procedural, conditional, disciplinary, and epistemic. Texts also reflect particular cultural experiences and knowledge: the more familiar readers are with the experiences and knowledge inscribed in texts, the greater the opportunity for readers to comprehend these texts (Afflerbach, et al., 2020; Paris, Lipson & Wixson, 1983). In school, declarative and procedural knowledge receive the bulk of attention, declarative representing the content that students must learn across school subjects (the what), and the procedural representing the strategies used by readers to construct meaning from text (the how). But the others, which are more recent entries into theoretical accounts of the reading process, are equally important in reading development (Goldman et al, 2016; Afflerbach, et al, 2020).

Declarative Knowledge is the knowledge readers possess about the world—the stuff of school subjects surely, but also the stuff of everyday life—bus schedules, routines for self-improvement, managing relationships with others, or a recipe for pizza. Declarative knowledge is also known as world or background knowledge (Anderson, 2019; Anderson & Pearson, 1984), one's knowledge about how the world works (Alexander, Kulikowich, & Schulze, 1994; Schallert, 2002). Also, declarative knowledge constitutes one aspect of the funds of knowledge all learners bring to a reading task (Moll, et al., 1992).

Procedural Knowledge reflects knowledge of meaning-making processes that are inherently social and cultural in nature. These include the cognitive skills and strategies that readers use to construct meaning (Duke & Pearson, 2002; Serafini, 2004). Procedural knowledge also includes word-level strategies and skills (e.g., decoding and word meaning and usage) as well as knowledge of language, texts, and concepts to reason about ideas in texts (e.g., locating, identifying, and recalling important information, integrating and interpreting information, analyzing and evaluating information, and using information to build knowledge and complete tasks) (Ilter, 2019; Kendeou & O'Brien, 2016; Pressley & Afflerbach, 1995; Vaughn, et al., 2016). Recent research describes the use and effectiveness of procedural knowledge for reading multiple texts (Bråten, et al., 2011), and navigating multimodal texts (Killi, et al., 2008) and Internet reading (Cho, 2014). Procedural knowledge is also front and center in moral reasoning (assessing the ethical bases of human behavior in literature, history, or biography, for example) and social reasoning (inferring the internal states of others, including story characters, about what the others know, believe, or intend).

Conditional Knowledge involves readers' mindfulness and metacognition, or the ability to reflect on and monitor one's learning (Veenman, et al., 2006), which is used to plan, manage and successfully complete acts of reading for specific purposes. Thus, conditional knowledge is operating when readers set or take up social or personal goals for reading (Dinsmore & Parkinson, 2013), judge their learning (Wever & Kelemen, 1997) monitor their progress at reading (Markman, 1979; Massey, 2009), identify challenges and fix them (Michalsky, et al., 2009), and bring an act of reading to completion. Conditional knowledge also figures in readers' awareness of their attitudes toward, and effort given to, reading (List & Alexander, 2018). While metacognition in reading focuses on use of strategies, conditional knowledge helps readers keep account of their resources given their purposes, their emotions, their histories of success or failure, and their motivation and engagement. The mindfulness that is reflected in metacognition serves readers' understanding and use of their cognitive strategies and skills and knowledge, their purposes, their motivation and engagement, and their agency. That is, mindful readers understand how marshaling motivation, effort, and perseverance may influence reading performance and outcomes (Efkildes, 2006).

Disciplinary Knowledge reflects the reader's understanding of how reading, writing, and language are used to build knowledge and communicate key ideas in disciplines (Goldman, et al., 2016; Lee, 2014). These literacies entail specific practices to meet socially defined purposes, which include forms of texts and communications as well as cognitive and affective activity (Moje 2015). For example, readers of science, history, and literature learn to draw on reading practices that have been shaped within and across disciplines. These include such strategies as considering sources and chronologies of ideas, comparing claims about causality, and interpreting rhetorical devices. These practices reflect how and why knowledge is used and communicated within disciplines. Learning to use these practices is central to developing disciplinary understandings.

Epistemic Knowledge focuses on how readers know what they know, and it assists readers in developing a stance toward both texts and associated tasks (Lee, 2016 or so; Hofer, 1997; Afflerbach, et al., 2020). Epistemic knowledge develops through social interactions that support readers to move from understanding knowledge as a set of facts to understanding that the same information can be interpreted in multiple ways and, ultimately, to understanding knowledge as constructed judgments that can be evaluated based on socially determined criteria (Kuhn & Weinstock, 2002). Epistemic knowledge plays an important role in the analysis and evaluation of complex texts (Barzilai & Zohar, 2012) and in evaluating the degree to which the claims that authors make can withstand scrutiny on criteria of relevance and trustworthiness. That is, readers with more developed epistemic knowledge are able to consider and evaluate multiple interpretations as they critically reason about ideas in real-world and discipline-specific texts (Barzilai & Zohar, 2016).

Interactions Among These Five Types of Knowledge. Interactions are essential to successful reading, depending on the text, task, and context of reading. Moreover, the different types of knowledge act dynamically and interdependently when readers succeed. Finally, these families of knowledge are not simply cognitive; they do not magically appear in readers' brains. They are inherently sociocultural because they come to individuals marked by the cultural practices of the communities in which they were forged and developed. The challenge, then, in assessing reading is how to structure activities that situate readers' varied cultural resources as assets for their understanding and use of what they learn from texts (Gutiérrez, et al., 2009).

Language. In the past quarter century, scholarly accounts of reading comprehension have elevated the role of language in explaining, assessing, and facilitating reading comprehension. Different facets of language provide strong explanations for the nature and quality of reading comprehension at different levels of development. Early on, in kindergarten through grade 1, sub-word processes like letter-sound knowledge and phonemic awareness tend to explain the majority of the variance in reading performance, while more meaning-based language variables, including receptive and expressive vocabulary, explain increasing proportions of the variance as students move into grades 2 and 3. As students move through the intermediate grades into middle and high school, more sophisticated facets of language, what is appropriately labeled as knowledge of language structures and vocabulary, assume an increasingly important role in explaining and improving advanced levels of reading comprehension (Bunch, Walqui, & Pearson, 2014; Cervetti, 2020; Fang, Shleppegrell, & Cox, 2006; Nagy & Townsend, 2012; National Research Council, 2010; Proctor, et al., 2020; Schleppegrell, 2004; Uccelli, et al., 2015).

Language and Diversity. Readers encounter and acquire multiple languages, discourses, and vocabulary as they participate in different cultural contexts and communities across their lifetimes. Many readers in the U.S. acquire more than one language or language variety with which they engage in literacy activities. The linguistic repertoires that readers develop reflect the distinct aspects of the languages and language varieties used in a variety of settings—home, neighborhood, places of worship, or school, for example. Readers bring these linguistic repertoires to bear on the meaning-making processes through which they engage with texts (Cazden, 2001; Harris & Schroeder, 2013; Heath, 1983; 2012). In particular, bilingual or multilingual readers who are acquiring English as an additional language bring an amplified linguistic repertoire that they may employ as both assets and constraints to the task of English reading (Jiménez, García, & Pearson, 1995, 1996). Their bilingualism or multilingualism affords a metalinguistic advantage in understanding similarities and differences across languages that could be employed in the service of reading (Daniel & Pacheco, 2016; García, 2009), underscoring the value of the home language (Aguilar et al., 2020; Francis, Lesaux, & August, 2006). Their ability to read English texts well may depend on their reading competence in the first language as well as their level of competence in English. In addition, bilinguals and multilinguals have to know how and when to tap into their biliterate or multiliterate knowledge to confront the lexical, syntactic, semantic, discourse, and knowledge demands inherent in reading in a less familiar language (August & Shanahan, 2006; García & Godina, 2017).

Language, Text, and Diversity. Social and cultural dimensions of reading involve the type of texts readers encounter but always within a particular contextual setting. Texts employ language that reflects different genres, discourse traditions, disciplines, perspectives, and worldviews, which cannot possibly correspond with all readers in the same way. For English language learners, English texts may present substantive challenges as they may encounter unfamiliar and inaccessible language structures, specialized vocabulary, and/or discourses. As noted previously, texts reflect particular cultural experiences and knowledge domains. Additionally, the reading situation reflects a particular sociocultural context as readers must understand what is being asked, consider time limits, try to read texts that may be of little interest, make strategic decisions about the correct answer, read without help, and so on. This situation is not unique to bilingual learners, though. The language of school texts offers a particular way of using language that is often unfamiliar to large proportions of students. In a sense, all students, even students who speak English as their first language, continue to be language learners throughout their school careers when they encounter the specialized language of the disciplines of science, history, and even literary analysis.. School literacy involves, to a large extent and for most students, the continuous learning of new sociocultural ways of using language for particular, and often unfamiliar, purposes.

Additionally, sociocultural perspectives highlight that the texts assigned in schools and on assessments may create special barriers for students whose sociocultural resources may be of little help as they encounter texts and reading tasks that minimally reflect their experiences, knowledge, and epistemology (Abedi, 2004; Solano-Flores & Nelson-Barber, 2001). To mitigate some of these challenges, schools and assessments could employ a wider range of text choices, ample representations of cultural and linguistic diversity in texts, broader opportunities for readers to demonstrate their comprehension and understanding on reading tasks, and scaffolds that direct attention to the salient features of the texts, activities, and tasks readers encounter in assessments. Indeed, the 2025 NAEP Reading Assessment aims, to the extent possible, to incorporate these ways of addressing these challenges rather than leaving them to chance.

Language and Reading Disciplinary Texts. Recent work in reading within disciplinary contexts (Goldman et al. 2016; Uccelli, et al., 2015) suggests that language acquisition—beyond communicative competence—has a facilitative effect on reading. Passage comprehension entails, to a large extent, readers' comprehension of the *language* of text, and this entails the particular language usages and discourses of disciplinary contexts such as literature, science, and social studies. A text is, after all, language inscribed on paper or a screen, or sometimes even a transcript of oral discourse. As readers attempt to make meaning of a text in a particular reading situation, multiple language histories intersect with one another, especially the language histories of both the reader and the writer of a text.

Readers bring the language resources learned throughout their own history of socialization into reading situations. These resources affect a reader's comprehension of the language of the texts read in school, on a test, or in the workplace. This consideration is key for linguistically diverse populations of students: their histories of engagement with texts in multiple languages are brought to bear in their meaning making around texts. Additionally, authors' language uses and choices are also influenced by their histories with particular discourse traditions (e.g., disciplinary traditions of genres and discourse structures). Moreover, the language traditions of both readers and writers are shaped by broader socio-political histories and power relations that interact dynamically as readers make sense of a text. For example, school texts often reflect mainstream experiences and discourses that might be unfamiliar to some readers, which could impede reading comprehension. Thus, a reading situation is often also an intercultural space as readers' language resources come into contact with the language resources and expectations deployed by the author or privileged by the discipline.

Readers' Affect and Agency. Research in human development highlights the dynamic relations among identity and emotional health as correlates of cognitive work such as that involved in reading. While the cognitive and linguistic aspects of reading receive the bulk of instructional and assessment attention, readers' affect (feelings) and agency (willingness and ability to take action), continually operate to influence student development and achievement; reading is no exception to this influence. Reading requires engagement, persistence, focus and effort. And it is in and across multiple settings that readers learn how to be engaged, to persist, to be focused, and to put forth effort as they construct meaning with texts. Equally, in and across multiple settings, readers develop perceptions: of the self, of others, of settings, and of tasks. Students' sense of agency as readers is highly related to how they perceive reading tasks, perceive themselves as readers of particular kinds of texts, perceive their efficacy, and perceive the resources and the salience of the resources available to them in their schools and communities.

Perceptions of the self involve many aspects of identity such as individuality, membership in families and social communities, and participation in academic communities (including communities of readers). In the context of schooling, and specifically experiences in schooling that entail learning to tackle texts, perceptions of positive relationships with others (e.g., teachers, fellow students), perceptions of the structure of schools as organizations and as classroom settings, and perceptions of tasks (what am I expected to do as I engage with texts that are difficult) all matter for readers' goals, effort, engagement, and persistence. These perceptions influence readers' motivation and engagement and their sense of self-efficacy and agency.

Motivation and Engagement. These play a key role in readers' success. When a reader's motivation is combined with reading strategies and skills, engagement with the reading task may result in comprehension of the text (Wigfield, et al., 2008). Motivation may be context-dependent and influenced by readers' perceptions of their personal goals, their safety, and their self-efficacy in specific settings. Motivation to read can be both task-specific and general (Latini, et al., 2019). Readers may be habitually motivated, based on their prior reading experiences and understanding of the value of reading (Guthrie, et al., 1999). Or they may be motivated to read broadly and deeply in history, but not in science. They may also be motivated to read by particularly interesting texts, topics, or activities, but unmotivated to exert effort on texts and topics they find less interesting. Negative reading experiences, on the other hand, are unmotivating to readers, leading to reluctance and disengagement.

Readers may be motivated to partake in a wide range of reading activities for aesthetic enjoyment, knowledge gain, and interactions with others. Motivation to read provides increased opportunities for students to practice reading strategies (Stanovich, 1986) and add to their stores of knowledge, which in turn can feed the reader's self-efficacy. Motivation can function as a consequence and cause of increased reading performance (or both), as the two factors serve to reinforce each other (Wigfield et al., 2016). Finally, motivation and engagement are malleable; a series of positive interactions in reading can increase reader motivation, just as negative experiences with reading can lead to a reader's lack of motivation.

Self-Efficacy. This is an individual's belief in their own capacity to act in a manner necessary to reach specific goals (Bandura, 1986). Self-efficacy also indicates individuals' perceived control; that is, they feel drawn to and motivated by activities in which they feel control and accomplishment, and they tend to avoid activities devoid of control and accomplishment. As such, self-efficacy is crucial to reading development and reading success

(Carroll & Fox, 2017; Peura, et al., 2019); without belief in self, readers have a diminished incentive to initiate reading, or to persevere when facing reading challenges (Bandura, 2006). The sociocultural shaping of readers' self-efficacy is clear; readers may demonstrate greater or lesser self-regulation and effort, depending on their self-efficacy for completing a particular reading task (Pajares, 1996) or their comfort in particular social circumstances. Readers' perceptions of classrooms, sense of school belonging, and sense of their communities' social support for reading shape the agency and self-efficacy with which they approach reading tasks. In turn, readers' sense of agency and efficacy influences their engagement and motivation to read and the effort they put forward during reading.

Thus, readers' affect and agency—realized in their motivation, engagement, and self-efficacy—have a strong influence on reading development as well as on readers' comprehension in general and their performance on specific reading tasks and assessments. Importantly, when readers are able to exercise some degree of choice about topics, reading materials, and tasks, their self-efficacy and agency is enhanced, leading to greater motivation and effort (Carroll & Fox, 2017; Hall, 2016). When assessments acknowledge the power and contribution of affect and agency, they may better represent the students' reading achievement.

Development Over Time. While all of these competencies—knowledge, language, and affect and agency—influence reading comprehension, some are more salient than others at different stages in students' development over time (Cervetti, 2020). Acquisition of these competencies can be enormously influenced by instruction and other social and cultural experiences. As acknowledged earlier, the language backgrounds of young children, including letter- and word-level skills, are most closely associated with reading performance. By grade 3, language skills such as vocabulary and syntax become more prominent, accounting for more of the variance in reading performance. By the upper elementary grades, however, background knowledge, motivation, and understanding of strategies required for inferencing and critique, as well as academic language skills, play progressively greater roles in readers' ability to comprehend what they read. Readers' comprehension grows even more sophisticated as they proceed through secondary school, refining identities, interests, and disciplinary expertise by engaging in varied literacy practices in and out of school (Goldman et al., 2016; Kim et al, 2018).

Thus, relations between readers' prior knowledge assumed by texts and the demands of tasks they are expected to carry out when reading and responding to such texts influence displays of competence in reading comprehension. Sociocultural perspectives highlight that the NAEP assessment may create barriers for students whose rich sociocultural resources may be of little help as they encounter texts and reading tasks that minimally reflect their experiences, knowledge, and epistemologies (Abedi, 2004; Solano-Flores & Nelson-Barber, 2001). At the same time, there are complex relations among the kinds of knowledge readers bring to aid comprehension. On the one hand, students can learn from texts about which they have limited prior knowledge when they have sufficient skill sets around cognitive processes and/or there is scaffolding to facilitate this learning. As disciplinary texts become more complex, the combined requirement of disciplinary background knowledge and language become increasingly important. On the other hand, there is a baseline of requisite prior knowledge that supports comprehension. Thus, there is not a simplistic, linear set of relationships to predict performance in reading comprehension. Assessments that account for this complex interplay are likely to better represent students' actual reading performance.

A Sociocultural View of Text

Text consists of symbolic representations, including multiple characters, symbols, and/or language that people can interpret. They relay ideas via sentence-level arrangement of characters, words, and punctuation and by the organization of discourse. The text features and content, as well as the sociocultural perspectives from which a text is written, read, and made sense of, affect readers' comprehension of the text. In recent decades, the proliferation of digital texts (nonlinear hypertext, multimedia, multimodal, multi-genre) has vastly expanded how readers navigate and respond to texts. Texts are now understood to draw on multiple semiotic modalities, including oral, print, graphic, and video forms, as well as observed discussions and other enactments (Kress & van Leeuwen, 2001). Youth have been observed communicating via tagging, 'zine writing, and purposefully selecting apparel to communicate ideas (e.g., Finders, 1997; Moje, 2000; Schultz, 2002). Though texts are brought to life (animated) when people read and respond to them, they can be animated in many different ways by readers who bring unique prior knowledge, experiences, perspectives, and resources to the reading process (Kintsch, 1998, 2019).

Social, cultural, and historical issues are at play in how texts are written, what they contain, and how they are read. Like readers and their knowledge and resources, texts are not culturally or politically neutral. Texts vary according to conventions of disciplines and communities; an article in a science journal differs from a historical document, which in turn differs from a novel. Novels, in turn, differ from one another. Even what counts as text varies by disciplinary community. Over time, differences in ideas that authors try to communicate through texts have resulted in a wide array of genres and modalities. Today's texts include print and multimodal forms, including all sorts of images, both static and dynamic. Variation by discipline is common; they include models and diagrams in science, illustrated timelines and maps in history, graphic novels in literature, text messages, layers of hypertext, and video in everyday communication. Regardless of the form of a text, authors bring cultural experiences to bear when they set pen to paper or fingers to keyboard (Kress, Jewitt, & Tsatsarelis, 2000; Smagorinsky, 2001).

Text Types. While virtually any genre, no matter its origin, can appear in almost any discipline, variability across disciplines is the norm. Different disciplines tend to invoke different text genres. These genres communicate ideas according to disciplinary conventions that have evolved over time to address authors' desire to communicate ideas. For example, expository texts that describe historical periods use sequencing conventions to describe how events unfolded and rhetorical conventions to persuade readers that their account is credible. Texts describing scientific phenomena may be structured to describe how phenomena are situated in the world, including diagrams, photographs, or even links to video to help explain scientific terminology and other aspects of phenomena. Authors of novels draw on metaphors, foreshadowing, and other literary devices to reveal characters' perspectives, motivations, or major themes within their narratives. And in all disciplines, multimodal texts incorporate hyperlinks and other navigational tools to enable readers to tap into additional sources or explore their own interests. The digital age has augmented disciplinary texts to include models, illustrations, diagrams, simulations, and explanations in a variety of media and multimodal forms (Beach & Castek, 2015; Giroux & Moje, 2017; Kress, 2013; Manderino, 2012), along with evolving conventions of argumentation using these sources (Goldman, et al., 2016; Newell, et al., 2011). Readers' previous experiences with genres, including their instructional experiences, influence their ability to comprehend such texts.

Text Structure. The structures of texts vary by genre. But they also vary within genres. Young children are generally aware from stories they hear and view how narrative texts are structured within cultures to include characters, settings, problems, and solutions. But as they move from stories to encounter expository texts, readers learn that texts are organized differently to accomplish new purposes— to describe, persuade, compare or sequence ideas, or explain causal relationships. With experience and teacher guidance, students come to anticipate these devices—and even cue words (e.g., therefore, however, in contrast, because) that reveal the text's structure. Readers who know about these organizational devices and cues are better situated to exploit them in transforming the information in text into knowledge that promotes understanding.

The number of ideas or claims, overall length, and even font size and spacing facilitate or impede reading comprehension, depending on experiences and knowledge assets brought to reading by a reader. Digital text forms may be static, which means that print, diagrams, and illustrations do not change, while others may be more dynamic, requiring readers to follow movement between print and video or across locations (e.g., clicking a link that moves you to another section). Digital texts offer added complexity because they require menu navigation, alternative forms of annotation, and following hyperlinks to read multiple documents, sometimes under the guise of a single task (Fletcher, 2015).

Types of texts that are also determined by sociocultural factors include figurative language; specific language choices, such as formal or informal register or tone; and organization, including chapters, stanzas, and headings, and other special features, such as dialogue, charts and tables, stage directions, hyperlinks, or visuals. Particular genres, including poetry and narrative, descriptive, and argumentative texts, follow structural conventions that continue to evolve to address changing sociocultural purposes. Distinct communities differ in perspectives toward ideas and the ways the communities draw on genre conventions to argue, tell stories, give directions, or describe procedures.

Text Complexity. Complexity is a function of a text's features (e.g., words, sentences, length, discourse structure, the difficulty and conceptual density of the topic); the knowledge and purposes a reader brings to reading; and the tasks, situations, and supports available to structure and support the reading activity. Given the many ways texts can challenge readers, researchers have developed several different ways to estimate text complexity, recognizing, at the same time, that complexity varies by reader and task. There are formulas to estimate how difficult a text is to read by considering factors such as sentence length and structure, word length, and the relative commonality or rarity of words. These formulas, while useful, often fail to account for subtle aspects of texts, such as complex themes presented in seemingly simple prose. To account for these factors, the quantitative formulas are supplemented by other approaches, which weigh text difficulty more qualitatively by looking at how print and other symbols are situated on the page or screen, and how readers' abilities to interpret such arrays are tied to their knowledge of and experiences with such texts and the resources needed to interpret them (Pearson & Hiebert, 2014).

The influences of text complexity on reading comprehension are shaped by sociocultural forces leading to readers' familiarity with certain kinds of texts and their topics. Particular text elements will resonate more for some readers than others. Readers who are experienced with using particular text features are more likely to be able to comprehend, interpret, and evaluate the texts that use them. The degree to which readers have experiences reading and comprehending varied kinds of texts contributes mightily to their reading comprehension, including during assessment. Readers can acquire such experiences through their cultural background or through instruction in school. The more familiar readers are with the experiences and knowledge described in texts and the structures of texts, the greater the opportunity for readers to comprehend these texts. For linguistically diverse readers who are learning English as a new language, and indeed for many students in U.S. schools, texts may present substantive challenges as they may encounter unfamiliar and inaccessible structures, language, and discourses (Goldman & Lee, 2014).

It is important to emphasize the most general point about examining all aspects of text though a sociocultural lens—and here it does not matter whether it is the ideas, the linguistic structures, or the sources of complexity under consideration: Text does not exist solely on the printed page or digital display. Text is the by-product of the cultural practices of the community that produced it, read it, and critiqued it. In all of these processes, the community animated the text, gave it life, most often through the talk that surrounded it, when it was composed, when it was comprehended, when it was analyzed and critiqued, and even when it was revered. That is why the sociocultural model insists on considering any one of its three key features (reader, text, and activity) in relation to the other two—and always within the sociocultural context.

A Sociocultural View of Activity

Readers read to complete varied reading comprehension activities. To complete a reading activity, a reader sets specific purposes, carries out reading tasks in particular ways to accomplish these purposes, engages in the mental processes needed to complete the task, and realizes consequences or outcomes of these actions. Reading is always driven by purposes that reflect the social and cultural contexts for reading, and activity encompasses the purpose-driven tasks, processes, and outcomes involved in reading, including the outcomes measured on NAEP reading. An activity is also influenced by how it is situated in a sociocultural context, including disciplinary context, as well as the sociocultural experiences the reader brings to the activity. These components are intimately interconnected and fully integrated (Calfee & Miller, 2005; Beach, 2000). These components are addressed in unique ways in reading assessment situations in which readers work to understand what is being asked for, must consider time limits, read texts that may hold little apparent interest, and identify correct answers.

Purposes. Purposes are the reasons readers have for reading. Purposes can include acquiring knowledge, making decisions, solving problems, or taking action; enjoyment or emotional involvement; or demonstrating expertise, as in responding to school- or test-based tasks and items. Reading purposes may be personal (e.g., I want to acquire mastery over this body of knowledge), social (e.g., I want to become a member of this community), or both (e.g., I want to outperform my peers on this test to get into an advanced placement course). Purposes may be intrinsic to the reader (e.g., add to their storehouse of knowledge, feel more confident about reading skills, or gain a new perspective on the world) or extrinsic (e.g., following directions for a task or reporting on something learned). A single activity might address more than one purpose, which can operate simultaneously and even evolve as the reader completes the activity (Coiro, 2020).

Because reading can involve multiple purposes, including the aims of the reader, teacher, and/or author, as well as the goals put forward in plans for curriculum, instruction, or assessment, purposes should be adaptable, fluid, and nimble. As readers read, their purposes for doing so may shift as they discover new information and recalibrate their understanding of the content or task.

Tasks. Generally speaking, a task is a piece of work to be completed; a reading task is the action a reader undertakes in order to fulfill a purpose; thus completion of reading comprehension tasks depends on a reader's purpose. Tasks may involve engagement with a single text or multiple texts, as well as additional actions beyond the reading, such as reflecting on or synthesizing what was read, writing a response to what was read (summary, evaluation, or critique), or producing an artifact (perhaps a PowerPoint presentation) that demonstrates one's understanding of and ability to use information from the text or texts. Tasks can be scaffolded or more open-ended. They can involve actions undertaken while reading or actions taken after reading—as a result of reading a single text or across texts with different points of view. Tasks may also involve interactions with other people (book club discussions, completion of shared research projects). In a very real sense, a task cannot be judged easy or difficult until and unless one knows the conditions under which it is completed, the texts in play, and the scaffolds, including other readers, available to assist in task completion.

Processes. Successfully addressing the purposes for reading depends on the execution and combination of a large number of cognitive, metacognitive, and affective processes that press into service sociocultural assets, such as language, knowledge, and experience, as readers make meaning of texts (Afflerbach, Pearson, & Paris, 2008; Stanovich, 2013; Paris, Lipson, & Wixson, 1983; Yeari & van den Broek, 2011). They draw on sub-word processes such as phonological awareness, letter-sound correspondences, monitoring one's comprehension while reading (metacognition), and semantic interpretation. These representations of symbols in the text allow a reader to understand both what texts say (what psychologists have dubbed the *text base*) and what texts mean (what they call the *situation model*). Knowledge jumpstarts this process. As noted above, readers use the knowledge they bring from memory to trigger text comprehension; they relate what's new to what they already know. Barriers to successful comprehension can occur with sub-word processes or knowledge. Even though the incidence is relatively low, it is still the case that a small percentage of low comprehension scores for older students can be traced to low levels of decoding skill (Wang, Sabatini, O'Reilly, & Weeks, 2019) and low levels of topical knowledge are consistently implicated as an explanation of low comprehension (O'Reilly & McNamara, 2007; Pearson, Hansen, & Gordon, 1979; O'Reilly, Wang, & Sabatini, 2019).

As comprehension unfolds, and readers build both a text base and a situation model, they run the gamut of specific processes such as locating specific ideas, integrating and interpreting those ideas in relation to one another, critically analyzing and evaluating the validity of those ideas, and even using those ideas to accomplish specific goals. This continuum is not unlike an age-old saying about comprehension: (a) you read the lines (locating and recalling information), (b) you read between the lines (integrating and interpreting ideas), and (c) you read beyond the lines (analysis and critique). And finally, readers can actually use the ideas to do something else (using and applying what they read).

These processes are influenced by sociocultural factors such as readers' familiarity with topics and tasks, the features of the text itself and the reader's familiarity with the genre or medium, and the social interactions and environment (institutional, communal, familial) in which reading occurs. When readers comprehend, they may gain access to new information that holds the possibility of changing the knowledge structures in their memory, which is then available for the next round of comprehension and so on (Kintsch, 1998, 2019).

However, this virtuous cycle—knowledge begets comprehension begets knowledge—breaks down when students are confronted with texts for which they cannot use their personal (culturally and socially informed) background knowledge to create that all-important situation model. In such circumstances, teachers can assist students in building needed connections between their existing knowledge and experiences and the less familiar circumstances and topics reflected in the text. Teachers can also help students build stamina and persistence for working through challenges they may encounter in texts.

Consequences. Consequences are the results of reading, and they are shaped by all the other elements of the activity—purposes, tasks, and processes—as well as by the reader's motivation to take part in the activity (Pearson, Valencia, & Wixson, 2014). Engaged readers read in order to accomplish a purpose, and the consequences provide them with insights, as well as concrete evidence, that allow them and others to gauge whether they have accomplished this purpose—and perhaps achieved other aims as well. These consequences may be tightly linked to a single text (locating and recalling facts); they may involve connections between texts (integrating and interpreting ideas and language); they may involve drawing on prior knowledge (which can lead to another kind of integration or to evaluating ideas); and they may involve producing new knowledge or artifacts (using and applying what was read to produce something new).

Consequences may be intrinsic to the reader or more extrinsic. That is, a reader might gain greater insight into the natural world or produce a product that demonstrates greater understanding or solves a particular problem. Even when a reader produces a tangible product, less tangible consequences often accompany the artifact—a reader adds to their storehouse of knowledge, feels more confident about reading skills, gains a new perspective on the world, and/or feels good about what they did.

Individual readers engage, coordinate, monitor, evaluate, and adjust what they are doing in pursuit of meaning (Ajzen, 2002). Their reading activities—the purposes, tasks, processes, and consequences—that students encounter in school and life can vary in the degree to which they can determine and initiate them, hold interpretive authority in them, and exercise agency and control in carrying them out. When readers engage in activities they choose, for purposes that they find meaningful and interesting, they are more likely to bring a variety of resources, including knowledge and agency, to their reading. They may also make greater demands on themselves to understand texts when reading activities invoke their interests, curiosities, and goals (Alexander & Grossnickle, 2016, Guthrie & Klauda, 2014; van den Brock, et.al., 2011).

What Happens When a Reader Reads

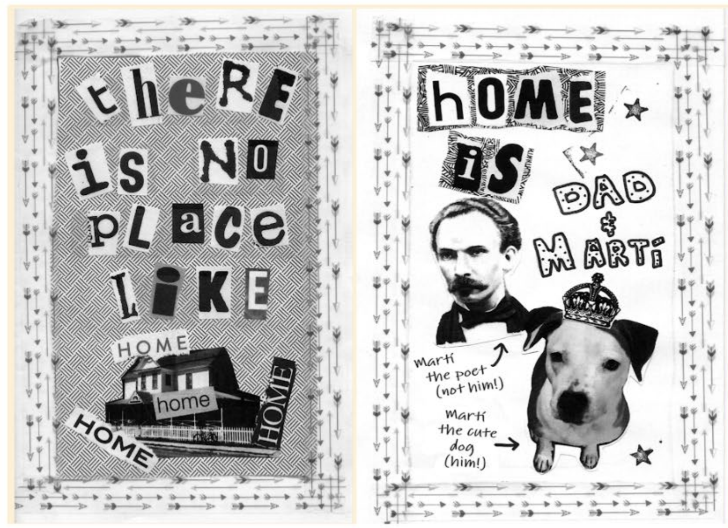
To see how readers bring their knowledge of texts and the sociocultural world to a text, consider the information in the text excerpt presented in Exhibit 2.2 below, Celia C. Pérez's (2017) popular multimodal novel, *The First Rule of Punk.* Depending on their experience with reading such texts, some readers would approach their reading with attention to the narrator's story of moving with a mother away from a father for two years, wondering about the character and why they were moving without attention to the inserted illustrations the narrative calls "zines." (Zines are defined by the narrator as "home-made magazines" that can "be about anything"). Readers who are zine authors, like the narrator, might instead skip through the print narrative and look carefully at the zine pages included in the excerpt, wondering who Marti is and whether the narrator's father, house, and dog looked like those in the photographs. Readers of any age who dislike fiction might set the text aside for more interesting pursuits.

Readers with limited experience reading and comprehending narrative fiction might decode the print—it is calculated with readability estimates ranging from third to fourth grade—but make little sense of the story, including the narrator's zines. Others might not be able to decode most of the print. Readers who see the author's name and recognize its Spanish origin might start to look for other hints in the story about the cultural setting and characters. A 4th grader who can read and understand the words might be intrigued by the storyline but not understand why a mother and child would move away from a father, even temporarily. An 8th grader might resonate with a storyline that resembles a friend's life experiences but decide this story seems simplistic. A 12th grader might be intrigued by the text's layout and illustrations but not at all interested in the story.

Exhibit 2.2. A Text Segment from Celia C. Pérez's (2017) Popular Multimodal Novel, *The First Rule of Punk.*

I started making zines earlier this year when I discovered Dad's collection of punk music zines from his high school days. Zines are self-published booklets, like home-made magazines, and they can be about anything—not just punk. There are zines about all kinds of topics, like video games and candy and skateboarding. A zine can be a tribute to someone or something you love and nerd out about or a place to share ideas and opinions. Dad said they're also a good way to write about what you're thinking or feeling, kind of like a diary that you share with people. Mine are mostly about stuff I find interesting or want to know more about. But ever since Mom told me we were moving, a lot of my zines had become about that.

Mom made it seem like this move was no big deal because we'd be back when her new job contract expired. But two years might as well be forever. Two years might mean all of middle school. And I couldn't even imagine what two years away from Dad would feel like. It was a *very* big deal. So for the next hour I wrote and cut and pasted a final plea to Mom. I glued the last letter onto a page just as the doorbell rang to signal that my time was up.



If the excerpt was used as part of a 4th grade reading comprehension assessment, readers might be asked to explain why the main character was making this zine. Readers might also be asked whether the narrator seemed angry or happy about moving away from the father, and to provide evidence from the story to explain their opinion. Readers might also be asked to evaluate whether the narrator was right to be angry with the mother. Finally, readers might be asked to imagine themselves in a book club that was comparing the feelings of the main character in Pérez's text to the main character in the text *Walk Two Moons*, by Sharon Creech. A reading comprehension assessment task that included the preceding might recognize differences in readers' experiences by offering the support of a prereading video on zines to ensure that all readers, and not just zine writers, could perform their best on the assessment.

The coherence of readers' mental representations of texts depends not only on the quality of the original text and the reader's accuracy at decoding that text, but also on the generation of local bridging inferences to create cohesive representation of the text (Cote, Goldman, & Saul, 1998). If more of the text was included in the Exhibit 2.2 excerpt so that readers knew the narrator was female, they might form different mental representations of the text than if that information was not revealed. Readers' mental representations might change further if they learn characters' names and story settings. Without such information, they must draw more inferences to create the cohesive ties needed for a situation model of the text's ideas or decide that the text does not provide enough information to build such a model.

Revisiting the Sociocultural Model, with an Eye Toward Assessment

There, in as brief a form as possible, is the model of reading that underlies the 2025 NAEP Reading Framework. It is the consistent sociocultural framing of how readers, texts, and activities interact that distinguishes this model from its predecessor. It is inherently sociocultural in character because all of the key components of the model—reader, text, and activity—are situated in both highly specific contexts, such as classrooms, homes, or digital spaces, and more general contexts, like communities, social networks, and nations.

The conceptual model for the everyday literacy practices described in this chapter cannot be fully enacted within the confines of any single assessment of reading comprehension. The 2025 NAEP Reading Framework elicits slices of readers' performance that at best serve as proxies for the components of more expansive, purposeful, discipline-specific, authentic reading comprehension activities. In particular, it attempts to invoke readers' demonstration of processes that have long been taken as evidence of reading comprehension.

As will be described in more detail in the chapters that follow, the 2025 NAEP Reading Framework draws on principles of universal design for learning, including offering learners access to multiple ways of acquiring needed background information; multiple means of engagement to tap into learners' interests, desire for challenge, and motivation; and multiple means of expression to provide learners alternatives for demonstrating what they know. These design principles attempt to account as much as is technically possible for readers' varied sociocultural sensibilities and the ever-evolving nature of texts and their uses in today's world. The 2025 NAEP Reading Framework aims to present a precise and accurate picture of what students in the nation's schools know and can do by inviting them to draw on the linguistic, cognitive, and epistemological strengths that individuals develop across their diverse communities of practice and bring to their reading. In this way NAEP strives to maximize the participation and optimize the performance of the widest possible population of students in the NAEP Reading Assessment.

The key question for this framework is this: What does it mean to build a reading comprehension assessment upon the foundation of this sociocultural model of reading? That is the question to be answered in Chapter 3, which describes the assessment construct for this framework.

The purpose of Chapter 3 is to demonstrate how the sociocultural model in Chapter 2 serves as the basis for the 2025 NAEP Reading Framework and the assessments that will be constructed using the framework.

Representing the Sociocultural Model of Reading Comprehension in the NAEP Reading Assessment Construct

Moving from the sociocultural model of reading comprehension to the assessment construct, it is important to highlight key relationships between the two. Even though the assessment construct is based on the sociocultural model, it is not the same as the sociocultural model. Granted, there is nothing in the assessment construct that is not present in (or strongly implied by) the sociocultural model of reading, but the assessment construct of virtually any process or domain does not reflect all the elements within a model of the process/domain (Mislevy, 2016). Reading is no different. The NAEP Reading Assessment construct is necessarily more constrained and less elaborate than the sociocultural model on which it is based. Why?

First, not everything in the sociocultural model can be easily assessed or even accounted for by the assessment. Some important facets of the sociocultural model are very difficult to assess because of logistical, political, and consistency issues. For example, NAEP cannot measure socio-economic status (SES) by asking participating students about their family incomes because (a) the question is intrusive and (b) the answer is not known by many students. Instead, NAEP and other large scale assessments in the U.S. often use proxy measures for SES, such as student reports of the highest levels of education attained by their parents and school-level reports of students' eligibility for free or reduced price lunch (Merola, 2005). Similarly, it might be useful to know whether students feel supported by their teachers, but such data are very sensitive—and leave open the possibility that the data could be used to unfairly blame teachers for students' scores. An example of inconsistency between a formal model and what assessment can describe based on that model is identifying the levels of English competence of individual English learners (ELs). While states use tests to determine EL performance, different states and consortia use different scales, categories of competence, and decision rules to determine English competence. What it means to have exited a bilingual or ESL program in one state is not the same as in another state.

Second, some features of the sociocultural model are emphasized more than others because NAEP begins at grade 4 and extends to grades 8 and 12. Although some struggling readers in grade 4 and later still have problems with foundational skills, such as decoding (Sabatini, Wang, & O'Reilly, 2019), the majority of the students at this grade level do not. Key areas that need to be taught and assessed in grades 4, 8, and 12 are how to comprehend texts in disciplinary contexts and apply what has been comprehended. For these reasons, NAEP Reading in grades 4, 8, and 12 is a test of students' English reading comprehension.

Even among the important components of the sociocultural model, some have higher priority than others. For example, there are a host of purposes for reading, as suggested in the section on "activity" in Chapter 2, but NAEP cannot include all of them in the assessment. Thus, NAEP must set priorities by selecting a subset of important purposes in creating tasks for the NAEP

Reading Assessment. In the 2025 NAEP Reading Framework, two reading Purposes (Reading to Develop Understanding and Reading to Solve a Problem) were determined to be the most encompassing and hence the most important.

Third, a standardized assessment can only approximate the contextualized and social nature of reading. The social and cultural contexts that surround reading in life include connected networks of people, experiences, and environments. They are part of the ongoing fabric of readers' experiences and typically involve qualities of regularity and familiarity. A standardized assessment like NAEP can attempt to represent versions of these contexts, but it cannot reflect the community, classroom, and home contexts that characterize students' usual reading lives.

Fourth, ensuring that the assessment honors diversity consistent with the asset orientation outlined in Chapter 2, so that all test-takers have an equal opportunity to show their comprehension strengths, is an elusive goal—one fraught with dilemmas. For example, the quest for texts that reflect students' diversity of experience and knowledge is constrained by the fact that NAEP Reading can sample only a limited number of texts. In a best case scenario, with an average of three texts per assessment block (a block is a 30 minute time period for reading and responding to test items) and 15 blocks per grade level at grades 8 and 12 (with 12 at grade 4), the sample would include 45 texts at grades 8 and 12 (with 36 at grade 4). This means that a particular student's cultural assets, funds of knowledge, experiences, and interests may not be reflected in any given NAEP activity. Similarly, the diverse manner in which students interact with texts in homes and communities, the specific text-related activities they engage in around texts, cannot be fully represented either. These challenges, however, should not prevent designers of the NAEP Reading Assessment from doing everything possible to achieve the diversity required to be able to call it an asset-driven assessment.

For all these reasons, there is some contrast between what the sociocultural model describes and what the NAEP Reading Assessment construct can deliver in relation to that description.

Moving from the Sociocultural Model to the Assessment Construct

Although some elements of the sociocultural model cannot be faithfully represented in a standardized assessment, the purpose of 2025 NAEP Reading Framework is to update the framework that guides assessment development so that it reflects as many aspects of the model as possible and takes advantage of the major shifts in how ideas are represented in text. A principal question in describing the move from the model to the construct is: What does it mean for the features of the sociocultural model to be reflected in the assessment construct? There are several ways in which features of the model are represented in the construct.

1. **A feature of the sociocultural model can be directly measured as an *outcome* in the assessment construct**. For example, test items are developed using the Comprehension Targets (i.e., Locate and Recall, Integrate and Interpret, Analyze and Evaluate, Use and Apply) to measure how well students have understood the texts, thus providing evidence about how well students are engaged in the processes described in the activity section of the sociocultural model. These outcomes emanate from the description of consequences in the sociocultural model.
2. **A feature of the sociocultural model can be directly measured, not as an outcome, but as an explanatory variable.** Explanatory variables provide additional information about the outcomes observed on NAEP. Thus a mini-probe querying students' level of

interest in the topic of a particular text might explain why some activities (an activity is defined in NAEP as the text(s) that students read plus the questions they answer) yield higher scores than others. Or, as has already been included in NAEP, a survey item measuring students' self-perception of how hard they work to succeed in school might explain differences in scores between groups of students.

3. **A feature of the sociocultural model may not be measured,** but it can influence student performance and how results are interpreted if that feature serves as a **scaffold** to guarantee student access to a resource required to demonstrate comprehension. For example, readers might view a short video about the use of weather balloons before reading a passage about weather forecasting; in such an instance the scaffold provides both necessary knowledge and a degree of access for all students.
4. **A feature of the sociocultural model is not measured directly, but is "taken into account" in the design process.** For example in selecting the texts and designing the activities in the assessment, test designers might be directed to screen passages and review items for broad cultural representation in order to increase the chances that students from all cultural backgrounds will "see themselves" in the assessment or find culturally relevant tasks to complete. This aspect of the assessment, while not directly reported, is nonetheless a critical aspect of how NAEP is created. It helps NAEP maintain a reputation as a credible and representative assessment tool. If and when educators, -makers, and the general public can see that the texts and tasks on NAEP broadly policy represent school curriculum, everyday reading experiences, and a range of cultural assets, they are more likely to view NAEP results as both trustworthy and valuable.

The Relationship Between the Sociocultural Model and the Assessment Construct

Exhibit 3.1 provides a visual metaphor for illustrating the plan for mapping the all-important relationship between the sociocultural model and the construct. The sections that follow will first provide a description of how each feature of the sociocultural model —context, reader, text, and activity, and in that order—is represented in one or more of the components of the assessment construct. This is followed by a description of how each component of the construct—activity structure and purpose, scaffolds, texts, items, and explanatory variables, and in that order—includes facets of each feature of the sociocultural model. Chapter 3 concludes with a large version of Exhibit 3.2, labeled Exhibits 3.3 and 3.4, with most if not all of the cells filled in to convey a complete sense of the mapping.

Exhibit 3.1. The Relationship Between the Sociocultural Model and the Assessment Construct

Components of the Assessment Construct	Features of the Sociocultural Model			
	Context	Reader	Text	Activity
Activity Structure & Purposes				
Scaffolds				
Texts				
Items				
Explanatory Variables				

The Overarching Importance of the Sociocultural Context. The most fundamental principle of the sociocultural model of reading is that, as a human meaning-making activity, reading is always situated in social and cultural contexts that shape every aspect of readers' engagement with text and influence how readers respond to and learn from the experience of reading. The Assessment construct reflects this understanding by using testing blocks that are highly contextualized. The 2025 NAEP Reading Framework assumes and attempts to build on the cultural assets (knowledge, skills, and practices) that all students bring to the assessment.

In terms of the many purposes that students pursue in reading, two very broad Purposes are ideally suited for a national assessment like NAEP that is charged with monitoring progress across the wide range of standards and curriculum frameworks developed by 50 states, each charged with the responsibility of providing education to its students:

1. Reading to Develop Understanding
2. Reading to Solve a Problem.

Similarly, texts offer students access to a world of ideas and insights to be understood, examined, and transformed into knowledge to use in all aspects of their lives. Texts are situated within bodies of knowledge and practice—in school these are known as content areas, school subjects, or, more recently, disciplines. Of the many possible sociocultural contexts for reading, 2025 NAEP Reading Framework will assess reading comprehension activities in three broad Contexts in which text is a salient feature.

1. Reading to Engage in Literature
2. Reading to Engage in Science
3. Reading to Engage in Social Studies

This emphasis on contextualization is present from the moment readers begin the 2025 NAEP Reading Framework assessment. For example, at the outset of an assessment activity, readers will be introduced to what will be called an activity structure. That introduction will specify a simulated context for traversing an entire 30 minute block, including

- A simulated social setting (a community setting or a classroom and perhaps some avatar classmates and even a teacher) and an explicit role for the reader
- A purpose for engaging in the entire activity (an activity-specific instantiation of one of the two overarching Purposes (Reading to Develop Understanding or Reading to Solve a Problem)
- The disciplinary Context in which the activity is situated (Literary, Science, or Social Studies)

The remainder of the tasks completed during the 30 minute block will be framed by this initial step of situating both the reader and the activity within the activity structure.

Contrast this model for the 2025 NAEP Reading Framework with previous iterations of the assessment developed from the 1992 and 2009 frameworks. In those versions of NAEP Reading (and countless other standardized assessments),which were considered to be the state of the art in their time, readers might have been presented with a stand-alone passage and directed to read it and answer the questions following it, with little or no guidance about what purpose understanding the passage might serve. Granted, it represented an implicit purpose — reading to demonstrate how well one can perform on the test — but it did not connect with any activity readers might engage with outside a testing environment. Thus, it is important to keep the sociocultural context in mind as the features of the model are applied to the Assessment Construct. How then do the key features of the sociocultural model (reader, text, and activity) emerge in the components of the assessment construct in the 2025 NAEP Reading Framework?

Readers. In the sociocultural model, the extent to which a reader succeeds at particular reading tasks is dependent on many factors related to the reader's experiences, knowledge, motivations, linguistic repertoires, and perceptions of self. In a strictly cognitive model of reading, measuring reading comprehension as an independent outcome would suffice. But in order to reflect a sociocultural model, it is important to gather information about how these additional factors affect performance, and, to the extent possible, to measure them or to develop scaffolds that optimize comprehension performance for every reader. Specific design considerations include the following:

- Given the significant impact of topic and domain knowledge on reading comprehension and the likelihood that students will vary in their knowledge of any topic, NAEP Reading Assessment designers can implement three practices: (a) design testing blocks around sociocultural contexts and particular topics that are likely to be broadly familiar to readers, (b) develop knowledge scaffolds that provide all students with a baseline level of knowledge regarding the context and topic, and (c) measure topical knowledge where appropriate. For example, in a reading block activity that features stories about coming of age in a Latin American community in the U.S., eighth-grade readers might begin the block by watching a video that includes an explanation of the role of the Quinceanera as a culturally specific example of the broadly familiar coming of age ceremonial tradition around the world. For any given text or activity block, test designers might select one or more of these practices. The sociocultural model describes how readers' identities and experiences in the world shape their access to and interpretation of text. In light of this, NAEP Reading Assessment designers will redouble efforts to ensure that the full range of diversity of the population of students is taken into account in selecting the texts and designing the activities in the assessment, thus increasing the chances that readers will find connections between their own experiences and the texts and tasks they encounter on in the NAEP Assessment.

- Given the significance of readers' motivation and perceptions of self and task in their reading comprehension performance, NAEP Reading Assessment designers will work on two fronts to enhance the role of engagement in the assessment: (a) by developing more engaging and relevant testing blocks and (b) by including block-specific and overall probes that yield indices of students' engagement in the activity blocks as well as their overall engagement in academic tasks. These block-specific probes might include quick ratings of readers' levels of interest in the topic of a particular testing block (e.g., how interesting, or difficult, or similar to your ELA class did you find the task you just completed?) or an overall survey item measuring students' self-perceptions of how hard they work to succeed in school. These probes might explain differences in scores between groups of students and might explain why some blocks (those with higher interest ratings) yield higher scores than others (those with lower interest ratings).
- Standardized reading assessments have historically been an isolated and individual endeavor; we even proctor consequential assessments to ensure that students are showing their own work, and without assistance of any sort. By contrast, the sociocultural model views human meaning-making activity as fundamentally social, with much of the reading that individuals do involving social purposes, cultural practices, and social engagement. Under the 2025 NAEP Reading Framework, assessment designers will create an activity structure for each block that contains elements designed to bring a stronger social dynamic to readers' work as they move through the block activity. For example, readers may "meet" an avatar guide through a video presentation at the opening of a block, they may be asked to engage with contemporary social issues, and they may collaborate with avatar peers as they work through the tasks within the overall block activity. It even may be possible for NAEP to consider real partners for NAEP reading assessments. More recently, assessment experts have recommended that designers of large-scale reading assessments consider including another live reader to collaborate with the test taker on reading tasks (see Coiro, Sparks, Kiili, Castek, Lee, & Holland, 2019; Fiore, Warta, Best, Newton, & Laviola, 2017). Over time, NAEP should examine the benefits of offering avatar partners, chat room partners, and even live partners as means of enacting the full set of collaborative options in NAEP. Some might question the use of these social options as a threat to the fundamental construct of reading. But they are a threat to that construct only if one adopts an individualized, static view of reading—exactly, as suggested earlier, the view implicit if not explicit in most standardized reading assessments. If one looks at reading as it occurs in the everyday world of homes, communities, and even schools, reading is as likely to exhibit a social character as it is an individualized character. Often readers can and do ask for assistance from others, consult with peers, talk about text and word meanings, and even argue about what a text means. In short, when the assessment construct is based on a sociocultural model, social interactions are a key part of the process.

Texts. The sociocultural model conceptualizes text as deeply varied in form and structure and as fundamentally reflecting the goals and beliefs of particular authors, contexts, and communities. This includes disciplinary communities that one might find in classrooms as well as informal special interest clubs that one might find in community centers. The goal of the assessment construct is to capture this view of text in the selection and development of texts and in the ways that readers are asked to engage with individual texts and sets of texts. Specific design considerations include the following:

- To respond to the sociocultural model's broad view of text, designers of 2025 NAEP Reading Assessment will draw from a wide range of genres, an array of media (e.g., prose and graphical representations), and a variety of dynamic text formats (e.g., hyperlinked documents). In addition, the texts will include content, organization and structures that characterize reading within different disciplinary contexts—as well as across those contexts.
- In line with the sociocultural model's recognition of the culturally based nature of texts, the 2025 NAEP Reading Assessment will include texts with different perspectives and different interests with respect to their representation of the world. Readers will be invited to engage with texts that represent an array of authorial experiences, worldviews, social purposes, and disciplinary contexts. Readers will also be asked to assume various perspectives to make sense of text and to Analyze and Evaluate the information presented in texts for utility, relevance, credibility, and intention.
- Importantly, the sociocultural model describes readers as having differential access to the particular information, genres, and language presented in texts. As such, designers working under the guidance of 2025 NAEP Reading Framework will select texts for accessibility based on multiple parameters, one being the likelihood that the topics are familiar to a wide range of students. In addition, scaffolds will be used to increase equity of access. For example, readers may have access to information about the meanings of potentially unfamiliar words (when word knowledge is not an assessment target), and they may be supported with background knowledge, as described above.

Activities. The sociocultural model describes activity as the purpose-driven tasks, processes, and outcomes involved in reading. From this perspective, reading is always driven by purposes that reflect the social and cultural contexts of reading. The locus of decisions about the purposes, tasks, processes, and outcomes can be external to the reader (e.g., a teacher assignment) or driven by internal goals and motives (e.g., a personal interest to learn more about forest fires). Although assessment tasks are necessarily shaped by external constraints to a greater degree than reading events in other settings, the 2025 NAEP Reading Assessment construct is designed to maximize the authenticity of the reading activities to the extent possible and to provide some level of reader choice. Specific design considerations include the following:

- Given the sociocultural model's emphasis on socially and culturally driven purposes and activities, the 2025 NAEP Reading Assessment testing blocks will situate reading inside of disciplinary Contexts (Literature, Science, and Social Studies) and broad Purposes (Reading to Develop Understanding and Reading to Solve a Problem) that readers might encounter both in and out of school. In doing so, the hope is to stimulate more internally motivated engagement on the part of the readers and to activate their knowledge and experiences outside of the assessment situation as resources for their work within the assessment.
- The 2025 NAEP Reading Assessment will provide opportunities for readers to make choices about particular pathways through the assessment experience. For example, they might choose (a) which of two or three texts to read, (b) how to answer a question

(multiple choice, drag and drop, or constructed response) or in what order to answer the questions, (c) to work with or without the avatar in a scenario-based assessment, perhaps even (d) whether to respond to a constructed response item in English or their home languages. With digital transcription of speech to print, as well as written translation programs now available, it is possible for students to either (a) speak using voice recognition to create a written draft or (b) compose in their home languages and translate it to English before it is scored. Such options should be made available to all students as a scaffold—not just to English learners as an accommodation. To this end, NAEP should undertake a special study on these choice options. NAEP should undertake these initiatives cautiously because the research on choice in a standardized assessment (Gribben, Patelas, & Schultz, 2020) is inconsistent in its findings. Findings range from enhancing student performance, to retarding student performance, to suggesting that choice benefits some students (i.e., higher achieving) more than others (i.e., lower achieving).

And, as a reminder of the core premise of the sociocultural model, all of these features (reader, text, and activity) are situated in that all-important sociocultural context. And that context both expands and constrains the ways in which the features are enacted in the process of reading for understanding, even in the 2025 NAEP Reading Assessment.

The NAEP Reading Assessment Construct

Having illustrated some major ways in which the key features of the sociocultural model are reflected in the assessment construct (moving from left to right across the columns in Exhibit 3.1) the description now shifts to the complementary journey from the top to the bottom of the table, as each successive row is detailed in order to illustrate how each component of the assessment construct reflects the features of the sociocultural model.

Activity Structure and Purposes

There has been a gradual shift from the assessment approach for 2009-2024 toward the assessment construct proposed in 2025 NAEP Reading Framework. The 2025 NAEP Reading Framework continues the journey begun and partially travelled in the last decade by making an even fuller commitment to the sociocultural model, the 2025 Reading Framework aims to do everything possible to situate the NAEP reader within a purposeful activity setting. The key action in making this shift is adopting what the Framework refers to as an activity structure for all NAEP reading blocks.

Overview of the Activity Structure of Blocks Within the 2025 NAEP Reading Assessment. It is important to understand the key elements (and associated terms) of this approach.

- The NAEP Reading assessment has been organized in 30 minute **blocks**. Readers have typically completed two reading blocks when they participate in NAEP. That practice can continue under the 2025 NAEP Reading Framework, but NAEP should also investigate the advantages of including blocks that last 45 or even 60 minutes.
- Each **block** will have one **activity**, which is largely defined by the broad Purpose (Reading to Develop Understanding or Reading to Solve a Problem) for which readers read and respond during the activity. However, the purpose that guides the block activity

is an activity-specific instantiation (e.g., read this set of texts to be prepared to participate in a discussion of the plot of this story) of that broad Purpose (Reading to Develop Understanding).

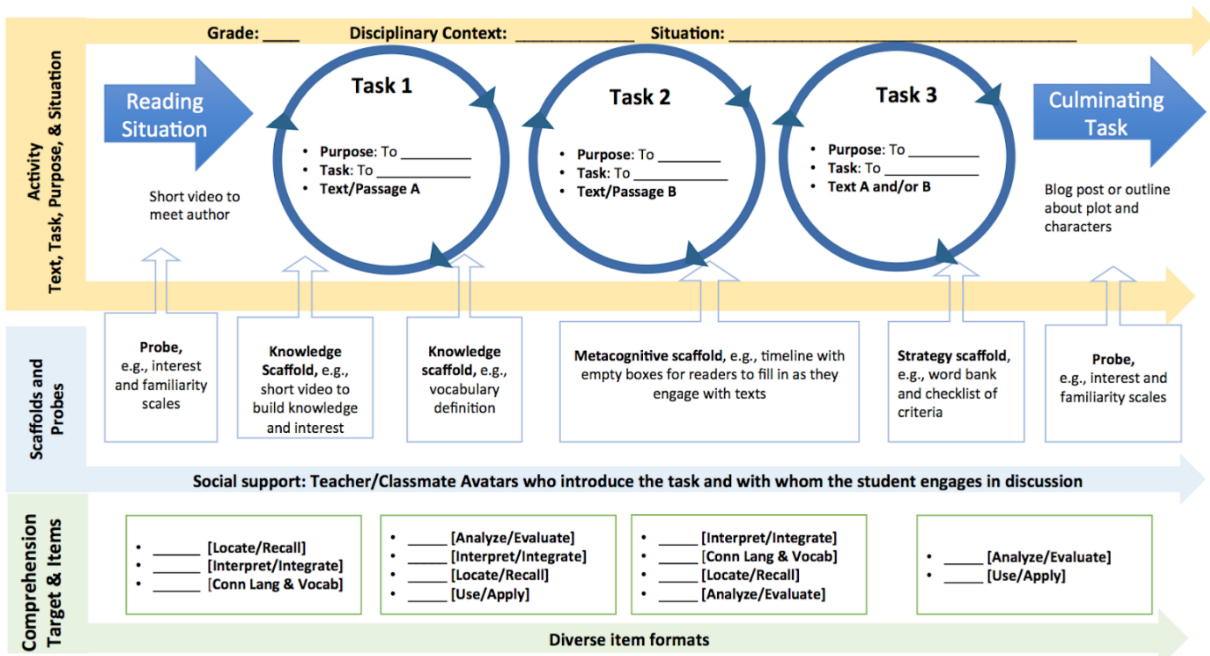
- Each **activity** is organized in an **activity structure**.
- An **activity structure** consists of two or more **tasks**. The **activity structure** is the grand total of everything a reader sees, hears, and does during a NAEP block; it also includes everything the test designers do to organize the flow of the activity within an activity structure.
- Each **task** includes a single **purpose**, one or more **texts**, between two and five comprehension **test items**, and an appropriate number of what NAEP refers to as **scaffolds**. Scaffolds include supports (e.g., a video about the topic) that enable students to perform at their optimal level, or probes (a self-assessment question about their knowledge of the topic).
- Some, but not all, blocks will end with a **culminating task**, which is an opportunity to draw from insights gathered from across the **tasks** completed throughout the **activity**.

The relationships among all of these elements are depicted graphically in Exhibit 3.2.

Getting Started: Situating the Reader. Right after receiving a tutorial about how to navigate the assessment on a digital tablet, readers are invited into the activity structure for the block by receiving a **purpose** for participating within the block and an indication of the **role** that they will play as the **tasks** within the **activity** unfold. This step is referred to as **situating the reader** within the **activity structure**. The principle behind this step is that when readers know why they are doing something (purpose) and what role they will play, their chances of succeeding are improved (O'Reilly et al, 2018).

Illustrating This Approach: An Example. To illustrate more concretely how this activity structure is enacted, consider the following example, accompanied by a visual model of the activity structure (Exhibit 3.2). At the beginning of a block activity, fourth-grade readers may be invited to participate in a book discussion group with three other 4th graders (represented by avatars in the assessment). Then, a teacher avatar explains the goal of the discussion and invites readers to engage with several passages from the book to prepare for that discussion. At that point the reader has been **situated within the activity** with both a purpose and a role. After responding to comprehension items for each of the three text segments (each of those would be categorized as a **task**), readers join the discussion with their avatar classmates and teacher to contribute their ideas about important plot events and how the character grew and changed over the course of the story. The ideas that they contribute to the discussion, perhaps in the form of an outline or a blog post, constitute a **culminating task**.

Exhibit 3.2. Elements of NAEP 2025 Assessment Activity



A design frame illustrates how all of the elements of a block activity are related to one another through the activity structure. It might also be used by test designers to plan the activity structure for activity blocks in the assessment.

Purposes Within the Activity Structure of Blocks. A major shift for the 2025 NAEP Reading Framework is the infusion of explicit purposes into the assessment to guide students' reading and responses as they complete a NAEP activity block. When students know why they are reading and responding to a text, particularly when the task(s) they complete is consistent with the established purpose, their work is more focused and comprehension generally is improved (Narvaez, van den Broek, & Ruiz, 1999; O'Reilly et al, 2018). Purpose gets realized in three ways in the planned NAEP assessment:

- The two **overarching Purposes** introduced earlier (Reading to Develop Understanding and Reading to Solve a Problem) help to frame the design of entire blocks in the **activity structure**,
- **Block/activity-specific purposes**, which represent a specific instantiation of one of the overarching Purposes, are established at the outset of a block activity as a way of framing the reading experience for students, and
- Highly **specific purposes** are established for a **task** (recall that a task consists of a text, some comprehension items, and some scaffolds) within the larger block activity.

For instance, in a Reading to Solve a Problem activity structure for a given block, 8th graders might read three texts about the mystery and history of the Egyptian pyramids, providing answers to questions about each text as well as across texts. For example, the three texts might include (a) a short History Channel video summarizing the important sites to visit on a trip to the Great Pyramids, (b) a travel brochure about sites to see in a tour of the great pyramids, and (c) a poem about the lingering mysteries of Egyptian pyramids. The activity-specific purpose for the entire block activity might be to:

1. Read these three texts to acquire information that you, along with your avatar classmates, might use to develop an itinerary for a trip to visit the pyramids of Egypt, taking care to give reasons for the specific sites you will visit. Your final product will be a letter to the Egyptian embassy, seeking funds for the trip you and your classmates plan to take.

But when they get to the second text, readers might be provided with an additional task-specific purpose, such as:

2. Read the brochure from the travel agency to see how it differs from the History Channel video you watched earlier in its list of "must-see" sites in the pyramid region of Egypt.

So, there may be, within a given NAEP block, both immediate and broader purposes provided to students. The list of possible purposes to guide reading is long, even for a given activity or task. These may include reading to lose oneself in a story, learn how to do something, develop a strategy for self-improvement, acquire information, or compare perspectives.

Scaffolds

Scaffolds are not entirely new to NAEP Reading; in fact, they have been increasingly included in NAEP since the introduction of the digital platform in 2017. Nonetheless, they represent perhaps the clearest example of just how much the underlying model of comprehension (the sociocultural model) has shifted from the cognitive models that guided earlier NAEP frameworks, and just how much of a shift this change has prompted in the reading construct being measured on NAEP. With the previous model, the confluence of the reading construct and assessment might be described as, "NAEP measures how well students understand what they read when they are equipped only with the knowledge and skills they bring to the assessment situation, and they are asked to read and answer questions about a text." This meant that tools like scaffolding, along with purpose setting, would be viewed as introducing construct-**irrelevant** variance (variations in performance unrelated to the assessment's underlying conceptualization of reading) into the testing situation. But in the sociocultural model, because reading is conceptualized as having both a social and cultural face, tools like scaffolding and purpose setting are now construct-**relevant**; in fact, without them, the 2025 NAEP Reading Assessment would not adequately represent the sociocultural model of reading. An expanded component in 2025 NAEP Reading Framework is greater use of scaffolds to support students in both navigating the assessment and performing to their potential.

Current NAEP Use of Scaffolds. In the 2019 NAEP assessment, four features qualify as scaffolds: the look-back button, a clear signal about when and how to read a text, a "resetting feature," and avatars as simulated partners. Look-back buttons are embedded links in a test item that refer readers directly back to the point in the passage at which information relevant to the item is situated; they are designed to save readers time by avoiding unnecessary scrolling or page turning on the tablet screen. These buttons are available for vocabulary items that ask a reader to examine the effect of particular words on the meaning of a paragraph or on comprehension items, for example, in which the reader must highlight the sentence(s) in a particular segment of the text that best support a claim made by an author.

Three additional scaffolds have been available on NAEP's scenario-based test blocks since 2019. First, unlike the situation for more conventional blocks, in scenario-based test blocks, right at the outset of the activity structure, students are told when and for what purpose to read a particular text; given what is known about the impact of purpose on reading, this can provide a powerful advantage. Second, resetting occurs between test items. For example, students respond to Item 2. After their response to Item 2 has been digitally recorded, but before they go to Item 3, students are told what the correct (or preferred) response to Item 2 is. The logic of resetting is that students should not carry over their misconceptions from one item to the next. Third, avatars are used in scenario-based tests to simulate a social dimension for assessment tasks. Students work with avatars, who represent classmates and often a teacher, to construct collaborative responses to particular test items.

Scaffolds for NAEP Reading in 2025. Informed by an asset orientation and scaffolding features implemented in the 2019 NAEP assessment, three categories of scaffolds will be systematically used in 2025 NAEP Reading Assessment to provide equitable opportunities to access and learn from information: knowledge scaffolds, metacognitive/strategy scaffolds, and motivational/social scaffolds. Each type of scaffold is designed to "level the playing field," increasing opportunities for readers to demonstrate comprehension and to sustain their willingness to engage and persist with NAEP assessment activities.

Knowledge Scaffolds. Students come to the assessment with a wide range of life experiences, prior learning, and perceptions of reality. Digital resources embedded into a reading experience can mediate a reader's vocabulary, prior knowledge, conceptual understanding, and problem-solving skills (Dalton & Proctor, 2008; Sparks & Deane, 2014). In turn, these resources increase opportunities for readers to demonstrate what they know and can do. In the NAEP Reading Assessment, knowledge scaffolds may be used to invoke or provide several types of knowledge: topical knowledge, vocabulary knowledge, or even discursive knowledge (knowledge about the norms and conventions for communicating ideas in a particular context or community).

Knowledge scaffolds may take the form of short videos, images, texts, graphic organizers, word banks, or a preview of key concepts addressed in the text. Other digital media (e.g., dynamic animations, glossary hyperlinks to related images—with or without language translation, simulations of interesting or challenging phenomena) can provide visual and multimedia cues to support readers' understanding of unfamiliar vocabulary or challenging concepts. For example, fourth-grade readers might have the option to access definitions or images when they encounter words or phenomena (e.g., entomology or okapi) that are likely to be both unfamiliar and necessary for comprehension. These just-in-time scaffolds enable all students, and especially English learners, to more deeply engage with critical text-based concepts while maximizing access to key ideas (Lotherinton & Janson, 2011; Zwiers, 2008). Discursive knowledge scaffolds can provide models of how different communities view a text or approach a problem (e.g., disciplinary reasoning, everyday problem solving, varied cultural practices). Twelfth-grade readers may, for example, be presented with a list of criteria used by literary critics to reason about narrative text or a reminder about what to look for in evaluating argumentative text. Importantly, knowledge scaffolds are not designed to provide answers to assessment items. Instead, they provide readers access to ideas and cultural practices that permit them to engage in the types of interpretive and evaluative processes required to demonstrate their comprehension of challenging text (Alexander & Jetton, 2000; Beuhl, 2017).

Metacognitive/Strategy Scaffolds. Embedded scaffolds can also be used to support metacognition and strategic, purposeful reading. At the beginning of an activity, strategy scaffolds (e.g., task-specific purpose statements or a sequential set of directions) can be used to clearly communicate expectations for how and why readers should engage with a collection of texts; they can also help readers plan and monitor their work across multiple texts and tasks (de Jong, 2006). Metacognitive scaffolds or probes may invite readers to reflect on their understanding or their reading choices at various points in the activity. These embedded resources can elicit information about how readers engage in metacognitive processes and whether different journeys through the assessment can explain levels of comprehension.

Strategy scaffolds such as simulated student work samples or mentor texts can offer models of literacy practices or approaches to solve a problem before students must do something similar on their own (e.g., Sparks & Deane, 2014). Scaffolds such as digitally enhanced note-taking spaces or tools that automate memory-intensive tasks (e.g., look-back buttons or select-in-passage items) can reduce time spent on low-level activities (scrolling to find the location) while providing students more time to read, evaluate, and engage strategically with text content. Other kinds of representational tools (e.g., graphic organizers) can help readers frame their conception of a reading task, make more explicit relations between ideas (e.g., cause/effect, problem/solution), or monitor their progress in the task (Suthers, 2001; Veerman, Andreissen, & Kanselaar, 2002). Metacognitive strategy scaffolds maximize the likelihood that readers are able to cognitively engage with complex NAEP-designed reading experiences.

Motivational/Social Scaffolds. Scaffolds can also be intentionally embedded into reading activities to support readers' agency, self-efficacy, interest, and engagement (Dalton & Proctor, 2008). Embedded motivational/social scaffolds serve to actively immerse readers in the context of an authentic understanding and/or problem solving space while communicating explicit connections between the broader purpose for completing a block and the sub-tasks that need to be completed along the way. Scaffolds in the form of avatars may provide written and/or oral directions, or they may interact directly with readers as experts or peers to provide assistance or moral support. Avatars may also represent members of an authentic target audience with whom readers can represent and communicate new understandings about what they have read and learned (e.g., Use and Apply). Knowledge scaffolds, as described earlier, may also serve as motivation scaffolds because being situated in an optimal topical space is likely to increase readers' willingness to initially engage and continue to persist with reading, both for everyday assignments and assessment activities (Buehl, 2017; Moje, 2005; Tovani, 2004). To the extent that assigned purposes (and related texts, tasks and goals) are viewed as meaningful and relevant, readers are more likely to be motivated to bring a variety of resources, such as personal and cultural knowledge, to understand and emotionally engage with or react to the reading activity as a whole (Guthrie & Klauda, 2015; van den Broek, Bon-Gettler, Kendeou, & Carlson, 2011).

Variations on Scaffolding. One alternative, or perhaps a supplement, to knowledge scaffolding is measuring students' knowledge of the topic of a block. Recent advances (McCarthy, et al, 2018; O'Reilly, Wang, & Sabatini, 2019.) in using short and efficient probes as indicators of prior knowledge, suggest that it is feasible to consider them for NAEP. The advantage of including knowledge probes is that they provide NAEP with the option of reporting comprehension outcomes in relation to students' perceived levels of prior knowledge for topics on the assessment. It is also possible to deploy a short measure (a few probes) of prior knowledge followed by an information-rich scaffold such as a video. These short probes could also be used to index other reader attributes, such as interest in the topic, effort required to complete the task, perceived level of difficulty, and the like.

Universal Design for Learning in NAEP. 2025 NAEP Reading Framework is informed by the Universal Design for Learning heuristic, a framework for expanding access to learning and assessment for all students (in contrast to accommodations, which are targeted for particular groups of students, such as ELLs or students with IEPs). Universal Design, among other functions, embraces all of the aspects of scaffolding in both the 2019 and the 2025 NAEP Reading Frameworks. The Universal Design principles are reflected in NAEP's attention to multiple means of support, action, and expression (e.g., giving learners various ways of demonstrating what they know) and multiple means of engagement (e.g., tapping into learners' interests). For example, a common list of response options could be made available to all students:: (a) to construct a response in English or their home languages (with digital translation into English), (b) to use a speech-to-print option for composing constructed responses (including in their home language) or (c) to select or construct a response. In addition, all readers could be given choices that increase their engagement, including choices regarding digital partners (avatars) and how they navigate the various texts and tasks within a testing block.

The Nature and Characteristics of Texts in 2025 NAEP Reading Framework

The 2025 NAEP Reading Framework proposes that the Assessment draws samples from the large domain of texts that 4th, 8th, and 12th graders are likely to encounter in school subjects (including language arts/literature, history and social studies, and science) plus the texts of everyday experiences outside of school (e.g., magazines, newspapers, public documents and government forms, and on-line sites). In addition, 2025 NAEP Reading Framework reflects changes in what counts as text, mainly due to paradigm-shifting research in the realms of digital, multimedia, and multimodal literacies.

This section provides a description of the range of texts that will be represented in the NAEP assessment at grades 4, 8, and 12. These texts vary considerably in terms of important dimensions: (a) the reasons for which they are read, (b) the fields of human experience from which they originate, (c) how they are organized, both at the broadest level (genres and text types) and a narrower, highly specific level (discourse and language structures), and (d) the degree to which they incorporate characteristics that were not a part of the "world on paper" that preceded the current revolutionary digital age.

Purposes for Reading Texts

As documented in chapters 1 and 2 and earlier in this chapter, there are many reasons why readers read. The 2025 NAEP Reading Framework focuses on two overarching Purposes: Reading to Develop Understanding and Reading to Solve a Problem. The goal for the 2025 NAEP Reading Assessment is to ensure that both of these Purposes figure prominently in selecting texts for NAEP reading block activities.

A Disciplinary Account of Texts

In the 2025 NAEP Reading Framework the disciplinary grounding of reading figures prominently in many ways. Results will be reported for three scales—Reading to Engage in Literature, Reading to Engage in Science, and Reading to Engage in Social Studies. NAEP will also sample texts, both those students are likely to encounter in schools and those they are likely to encounter outside of school, within these broad disciplinary contexts. Each disciplinary context includes genres and text types and the specific discursive, rhetorical, and syntactic structural characteristics of texts within those disciplines. Moreover, these disciplines are sociocultural contexts in which particular social practices such as modes of inquiry and ways of making and sharing knowledge are valued. These sociocultural practices shape the forms and features of texts and the roles texts play in disciplinary activity. NAEP test designers employing 2025 NAEP Reading Framework will attend to these sociocultural practices in creating test items.

A Caveat. As noted in Chapter 2, even though text types, genres, and text features differ across disciplines, these are normative, not absolute differences. All texts possess internal structures and features that define both the overall organization (the global structure, which are often labeled as genres) and the relations among the ideas in sentences, paragraphs, and particular sections (the local structures that help to move along arguments, explanations, and narratives). And many of these structures travel across disciplines. For example, novels and short stories, even chapter books and picture books for younger readers, possess long stretches of a description of a setting that are organized quite like a description of a geological formation in earth science or a description of the coastline of a peninsula in geography. Rhetorical structures (those “mid-level” formats that organize paragraphs and sections of text such as cause-effect, problem-solution, and comparison-contrast) appear in both science and history texts, and for that matter in informational texts about almost any topic.

There are, nonetheless, useful family resemblances among structural characteristics within disciplines that, if known to students, permit them to exploit structure for both comprehending and constructing texts. Consistent with the sociocultural model of reading underlying the 2025 NAEP Reading Framework, these characteristics are not the exclusive property of texts; they are also a part of the communicative practices that communities engage in to conduct their daily lives in and out of school. In short, as the sociocultural model of reading would predict, these structural possibilities in text cannot emerge without readers, collaborators, and community practices to give them life.

Literature Context. Perhaps more than in any other disciplinary domain, texts are the center of literary study and enjoyment. Literary texts have been characterized by genre, including myths and legends, short stories, novels, dramas, and poetry, among other classic forms. In school, students learn about and through literature by engaging in the study of these genres. But literary texts play important roles in out-of-school contexts as well, providing opportunities for enjoyment, reflection, and connection around themes related to human experience. In the context of NAEP, both classical literary genres and more everyday genres, such as fan fiction, author interviews, and book reviews, are viewed as important examples of literary texts.

Genres and Text Types in Literature. Although literary fiction texts most often follow a narrative text structure, there is wide variation in how the particulars of this structure are employed across different literary genres. Narrative structures, such as characters and setting, take different forms in science fiction, fables, satires, myths, coming of age stories, and magical realism, allowing readers to anticipate future plot events and interpret the development and motivation of characters (Hillocks, 2016). For example, a children's fable focused on the vanity of its main character cues the reader to expect a global moral lesson on modesty. Plot and character structures are thus sociocultural in origin, derived from varied ways of artfully portraying and reflecting on the human experience. Across cultures, themes common to all human experience pervade works of literature—nature and humanity, struggle and survival, love and friendship, loss and betrayal, victory and defeat, mortality and meaningfulness. Literature texts also include non-narrative literary genres such as varied forms of poetry and literary criticism and thus other text structures (e.g. haiku, sonnet, blank verse, and argument).

Discourse and Language Structures in Literature. Beyond the long-standing genre distinctions are text structures and literary uses of language that distinguish works of literature. Authors create their fictional worlds by selecting, sequencing, and manipulating language and using structural elements to convey meaning. The figurative language of literature often includes imagery such as metaphor and simile to promote a more nuanced view of a character. Particularly as they move into adolescence, readers can also expect to encounter symbolism, irony, and satire that cue non-literal interpretations of events and characters. Literary works are further distinguished by reliable or unreliable narrators, narrative point of view, and word choices designed to foster a particular mood and tone. Language choices similarly place narratives or dramas in a time period, and position characters in social hierarchies and as occupying particular types (hero, villain, faithful companion) in metanarratives such as classic tragedy. Increasingly, literary writers make use of multimodal forms, spatial arrays, and visual elements; graphic novels offer a good example of this multimodality. Imagery is common, of course, in poetry, which can be written in rhyme, to a form such as haiku or sonnet, or in free verse. Further, literary texts may feature typography that helps communicate the author's message.

The Relationship Between Literary Text and Sociocultural Practice. Literature invites readers to examine language, rhetoric, and structure; connect to other texts and other authors; and situate the problems the text poses in a social and cultural world of meaning-making about the nature of human experience. Literary inquiry invites, then, broad connections to social and cultural ways of understanding human experiences. Authors assume their readers will share understandings about how language is used to signal particular meanings, including both literal meanings (the gavel is a wooden hammer used to open and close meetings) as well as symbolic or figurative meaning (the gavel is a symbol of power so it is her way of letting him know she is going to oust him from power) (Rabinowitz, 1987). Literary meanings are always constructed in relation to the particular set of experiences, social and cultural practices, and emotions that readers bring to the task, which can lead to varied interpretations of the same piece of literature even by members of a particular community (Hull & Rose, 1990; Lee & Spratley, 2010; Scholes, 1985; Smagorinsky, 2009).

Contemporary approaches to teaching literature aim to serve three goals, all involving deep engagement: reading for enjoyment, reflecting on one's own life and the lives of others, and analyzing and evaluating the artistry (the author's craft) through which authors construct their works. In the elementary grades, readers may begin their journeys of literary inquiry by forming connections between textual features and authorial purposes. For example, stories are organized around a set of recurring elements (e.g., plot, character, and theme), while literary texts intended to persuade (a book jacket blurb extolling the virtues of a new chapter book) are organized around a set of claims, evidence, and reasoning, often with an acknowledgement and refutation of counter arguments. More advanced readers of literature engage in particular modes of inquiry and interpretive processes to reflect not only on what a piece of literature conveys, but on the nuanced ways in which it does so—the author's craft (Lee, Goldman, Levine, & Magliano, 2016). Text-based discussions engage students in sharing their perspectives and interpretations grounded in evidence from the text. Students thus learn to make the case for their interpretations of a piece of literature, using examples and evidence from the literary works (Rex & MacEachen, 1984; new citation).

Sampling Literary Texts. When applying the principles of the 2025 NAEP Reading Framework to assessment development, test designers should carefully sample literary works for inclusion in literary testing blocks, with close attention to structural and linguistic elements. The range of possible texts includes the more traditional genre distinctions that have guided earlier NAEP Frameworks (e.g., the 2009-2019 Framework), plus the additional text types and discourse features outlined here. The texts should be sampled to support both the interpretive analysis and socioemotional connections to literature that are at the heart of literary inquiry. More complete guidelines, including the carryover from the 2019 Framework, appear in tabular form in Chapter 4, where design principles and guidelines are provided in more detail.

Science Context. In both its public and professional applications, science involves the use of varying text types for a range of purposes related to understanding and acting upon the natural world. For students, a common denominator is the textbook as purveyor of facts. But the texts used when scientists do science and when non-scientists read to learn about science and solve everyday problems reveal a broader palette of forms and content, and a more synergistic relationship between text and inquiry. Accordingly, NAEP will include a broad range of texts related to scientific issues and problems, including newspaper articles, public advisories and alerts, and blog posts, as well as texts written to convey science facts.

Genres and Text Types in Science. In their work, scientists encounter and employ a wide variety of text genres, including raw data, bench notes, journals, personal communications, refereed journal articles, and review articles (Goldman & Bisanz, 2002). Science texts written for the general public include such forms as press releases and news briefs. Websites and blogs convey science information and engage public dialog about science topics. Digital tools engage the public in collecting and sharing data about everyday phenomena that reflect scientific principles—ecology, climate change, and the spread of disease, for example, and they increasingly offer live portals into scientists' ongoing work, whether in Antarctic explorations, in oceanic expeditions, or in space. Elementary age readers often engage with explanatory ("all about") texts on scientific topics, biographies of scientists, and discovery narratives, as well as socio-scientific texts that apply science to understanding and solving contemporary problems.

They may collect and compare data about science topics (such as plant structure and growth) to those of real scientists. More advanced, adolescent readers commonly engage with more complex versions of these same kinds of texts and with a wider range of scientific texts that include raw data, reference, texts, and scientific products, such as journal articles. Science text structures include cause and effect, correlation, problem and solution, sequence, comparison, exemplification, descriptive classification, extended definition, and analogy.

Discourse and Language Structures in Science. Science texts make use of a variety of ways of structuring and representing scientific information. Representational forms common to science texts include visual and graphical elements such as tables, graphs, equations, diagrams, schematics (e.g., flowcharts), as well as extended description, exposition, and narrative (e.g., Cromley et al., 2010; Lemke, 1998; van den Broek, 2010). Although science writers employ narration and description rich in imagistic language, often invoking metaphors to infuse clarity and vividness to their explanations, precision rather than divergent uses of language prevails. Examined at a more granular level, science texts tend to contain nominalized forms of verbs (to digest becomes the process of digestion, for example), the use of passive and even agentless passive voice (a liter of hydrochloric acid is added to the solution), and technical and specialized meanings and expressions (note the specialized meanings of words such as heat, mass, and power in science compared to everyday usage of these words) (Fang & Schleppegrell, 2010; O'Hallaron, Palincsar & Schleppegrell, 2015). Further, science texts regularly communicate degrees of certainty, generalizability, or precision through specific lexical qualifying terms or phrases, such as "may," "it seems," "this suggests," "in this case," and the like.

The Relationship Between Science Text and Sociocultural Practice. Science provides a specific sociocultural context in which rationality, logic, and objectivity are highly valued; additionally, science puts a premium on the discourse practice of argumentation (Osborne, 2010). However, science is also a social activity in which values, motives, and interests play important roles (Cavagnetto, 2010; Latour & Woolgar, 1986). At the same time, varied cultures and social groups may differ in the ways they interpret scientific information in text. Science texts are involved in theory and model building, and argumentation from evidence. As well, texts are central to students engaging with a wide range of natural phenomena—elementary students coming to understand the life cycle of a butterfly, middle school students debating the nutritional value of school lunch, and high school students creating arguments for controlling air pollution. Modern science education engages students from the early grades and onward in collecting observations in varied forms (data), making claims about their observations, using data as evidence in support of their claims and/or to refute alternatives, and explicitly laying out the reasoning that connects this evidence to their claims as they evaluate the claims and reasoning of others (Bazerman, 1998; Berland & Reiser, 2009; Cavagnetto, 2010; Driver et al., 2000; Passmore & Svoboda, 2012).

Sampling Science Texts. The NAEP Reading Assessment should sample science texts that reflect the language and structural elements germane to pedagogical, public, and professional science discourse. Texts should come from school curricula and more public outlets such as magazines and newspapers (including those written for K-12 students), internet sites, and public brochures. In addition, NAEP science texts should reflect the range of science text formats, from traditional to digital and multimodal. Attention should also be given to the sociocultural framing of science text purposes (e.g., conveying information vs. supporting an opinion), values (e.g., preserving the planet vs. protecting individual liberties), and applications (e.g., solving a problem or constructing—or refuting—an argument) in developing NAEP blocks, simulating the inquiry practices of the discipline.

Social Studies Context. Social studies texts provide students with an intellectual context for studying how humans have interacted with each other and with the environment over time (College, Career, and Civic Life Framework for Social Studies; 2013). As such, social studies and historical texts often focus on how humans organize their societies and governments, how societies make use of resources available in particular locales, how cultures develop and change over time, and how they manufacture, exchange labor, distribute power and influence, ascribe roles related to gender and age, and govern themselves. The field of social studies includes such focus areas as cultural studies, geography, civics and government, and history, with less common forays into disciplines such as sociology and anthropology. These fields offer unique ways of thinking and organizing knowledge as well as evaluating the validity of explanations. Texts are key sources of information linked to these disciplinary lenses.

Contemporary approaches to teaching history and social studies increasingly involve interpretation and critical thinking about multiple perspectives, rather than solely acquisition of information. In the past two decades and in school contexts, the use of a single text to represent history has evolved into the use of sets of related primary source texts that revolve around historical events. This evolution in text type reflects related changes in the purposes for reading social studies texts, moving from an emphasis on memorizing and giving back historical facts to interrogating sources, assuming perspectives, and determining the trustworthiness of texts (Reisman, 2011). Similarly, social studies texts are considered places to both acquire information and to practice critical reading strategies, such as interpreting media focused on politics, pollution, and populations (Wineburg, et al., 2016).

Genres and Text Types in Social Studies. History and social studies texts provide documentation of human activity in varied cultures and societies and across time periods. Through this evidentiary record, texts describe the complexity and significance of human experience. Thus, the types of texts and genres relevant to the study of history and social studies is vast. In the study of history, primary source texts include newspaper articles, census data, diaries, letters, speeches, inventories and records of sale, advertisements, and official government documents. Primary sources can extend beyond written documents to include media such as photographs, cartoons, maps, artwork, music, as well as video and audio recordings. In history, texts also include the interpretive books and articles historians write about particular events, periods, or people (often labeled as secondary sources), and the broad survey textbooks that populate many history classrooms. Secondary sources include biographies, maps and data tables, monographs, journal articles, and editorials that communicate particular analyses and perspectives. Students read all these forms when studying history.

The landscape of social studies also includes texts with content on civics, government, politics, cultural studies, and geography. Social studies texts in these areas—specifically those read in school—tend to be the more traditional textbook. Such texts may follow temporal or causative narratives, with the straightforward and uncontested presentation of information intended to support readers' learning and memorization of content. In geography, students may study maps, land features, and resources that differently support human societies and their development. Students likely encounter varied lifeways as they study ancient and world cultures in social studies, perhaps engaging with cultural artifacts, photographs, and drawings as well as languages and forms of writing different from their own. Texts are also a primary vehicle through which individuals engage and participate in civic life. In studying civics, students may read laws and documents describing how governments are organized, newspaper accounts of current events, polls of public opinion, podcasts and speeches, blogs, and other forms of persuasion. Social studies texts also include public-facing genres such as debates, discussions, magazine articles, museum collections, documentaries, and historical blogs and podcasts designed to describe and debate important issues.

Discourse and Language Structures in Social Studies. Historians and social scientists draw on a broad range of texts and images to explain the social world in the past and present. The texts commonly organize ideas chronologically and/or thematically in terms of periods of history or structures of social organization, identify continuity and change, analyze cause and consequence, take varied social or historical perspectives to make valid interpretations of varying points of view, and consider the consequences of the past for the present or the current conditions for the future (Charap, 2015; Seixas, 2010; Seixas, et al., 2015; Schreiner, 2014). The ways in which language is used in the social studies and in the primary sources that document human enterprises include conventionalized linguistic expressions and word choices; linguistic frames for organizing arguments; rhetorical markers of persuasion; lexical expressions that mark chronology and the beginning and ending points of a historical story or argument; and linguistic signals of cause and effect, comparison and contrast. Historical texts and documents may also reflect the language of a time period long past and can therefore use archaic lexical and grammatical forms and expressions that have no modern counterpart. This requires readers to consider language and representations in their sociopolitical historical contexts. For example, understanding that the name a writer selects to talk about the mid 19th century war—the Civil War, or the War Between the States, or the War of Northern Aggression—lays bare their ideological assumptions.

The Relationship Between Social Studies Texts and Sociocultural Practice. Contemporary social studies engages students in reading and inquiry that is decidedly sociocultural in nature and spirit. In fact, the social studies are inherently meta-sociocultural, since they examine social structures and cultural systems through the lens of varied social structures and cultural systems. These tasks reflect the uses of texts to unearth, understand and critique multiple perspectives and points of view about events and their significance to human societies and cultures. Primary source texts in history and the social studies are understood to reflect the purposes and perspectives of their authors; examining these sources of information for the world views they embody (and those they omit) is therefore key to reading these texts. Reading historical and social studies texts therefore entails close attention to the specific forms of language their authors use to describe and chronicle events, make claims, or conjecture about the motives and states and perspectives of others to understand the authors' positions and perspectives (Schleppegrell, Achugar & Oteíza, 2004). In addition, texts are central to making comparisons to corroborate viewpoints and information among sources, to hypothesize cause and effect relationships, to investigate interactions among events and people, to examine the impact of competing forces, and to separate fact from fiction and opinion and perspective to evaluate the credibility and reliability of different sources (Seixas, 2010; Vasquez, 2017).

Sampling Social Studies Texts. NAEP should sample from the varied forms of texts in history and the social studies to include a wide array of text types and multimodal forms of representation as well as sources of information. In addition to the primary and secondary sources that are most commonly used in social studies instruction, NAEP should include public social studies texts listed above. Above all, the texts must authorize the kind of interpretive and critical analysis that is at the heart of social studies inquiry. Careful attention should be given to ensure that NAEP reflects and represents a broad range of human experiences, perspectives, and points of view.

Interdisciplinary Context. Blocks containing more than one text have been a part of the NAEP Reading Assessment portfolio at all grades since at least the early 1990s. Although not mandated in previous frameworks, some of these multiple text blocks have included texts and activities from different disciplines. The primary value of these interdisciplinary blocks has been to provide a lens on performance in a situation in which students have to integrate information and perspectives across disciplinary boundaries. These blocks present students with a situation that is not unlike what readers must do in community and workplace environments where problems that need solving do not present themselves in nicely compartmentalized disciplinary packages. The 2025 NAEP Reading Framework calls for ensuring that some proportion of block activities continue this practice by intentionally reflecting an interdisciplinary perspective. When interdisciplinary blocks are used, they will be assigned a primary disciplinary context. The item scores on these blocks can either be assigned entirely to the primary block or parsed out according to the disciplines of the texts to which particular items are attached (which is the current NAEP practice).

The Impact of the Digital Platform for NAEP Reading. The characteristics of text have changed since the 2009 NAEP Framework. In the earlier Framework, designers could assume all texts would appear as print on paper. Now that NAEP has moved to a digital platform, all text appears on screen. Granted, when conventional print texts appear on a computer screen, they are still fundamentally static texts that just happen to appear on a screen rather than on paper. But the digital platform changes the range of affordances present during the assessment, including an array of scaffolds (e.g., easily available word pronunciations and meanings) and media options (especially images and video) (Coiro, 2020; Fitzgerald, Higgs, & Palincsar, 2020).

The digital platform also enables NAEP to assess facets of reading that were not possible to assess with print on paper. The widespread presence of computers and smart devices in modern society has changed society's ideas about what counts as text. Increasingly students in school are not only allowed but required to read texts that are native to a digital environment, an environment in which the range of experiences that can be represented is dramatically different from the world of print on paper. New sets of characteristics are needed to define the portfolio of texts that could possibly be sampled in NAEP.

The widespread adoption of the Common Core State Standards (NGA-CCSSO, 2010) has accelerated the acceptance of this broader view of what counts as text for reading instruction and assessment, explicitly bringing new forms of text under the reading umbrella. State consortia (including SBAC and PARCC) broadened text types to be measured to include audio clips, podcasts, infographics, selections from graphic novels, and video segments, including video with and without sound. Even states that have moved away from the Common Core State Standards, have kept this expectation, and state testing systems designed to measure reading comprehension standards have similarly broadened the definition of text types to be measured. It would be hard for NAEP to provide a sampling of texts that are currently deployed in state curriculum and assessments without including this broader portfolio.

A major distinction in a digital world is that some texts move and others do not. A *static* (no movement) on-screen text (Barron, 2015) involves reading to make sense of ideas in a text with textual structures and features very much like those in a print-on-paper world. A static text can go well beyond alphabetic print; it can be multimedia (Mayer, 2001), including static words, numbers, and/or graphics, such as those in a still photograph, diagram, or table.

Dynamic texts, by contrast, do not stand still. And they disrupt, or at least discourage, a linear word-by-word, page-by-page, journey through the text. They require readers to follow movement across modes (e.g., between print and video or static image) or across locations (e.g., clicking a link that moves you to another section) in a text. A single text might be comprised of *multimodal* elements (words, moving images, animations, color, music, and sound) and require readers to move across two or more of these modes to construct meaning (Beach & Castek, 2016; Giroux & Moje, 2017; Kinzer & Leander, 2003; Kress, 2013; Manderino, 2012). A text that includes *nonlinear* textual elements (hypermedia or hyperlinks) introduces opportunities for readers to quickly move from one location or mode to another within the text (Burbules & Callister, 2000; Landow, 1994). A *dynamic image,* for example, may enable readers to hover over part of the image to reveal a pop-up box with a scaffold in the form of information (e.g., definition, example, or history, or even a short video with audio and animations) that elaborates ideas presented in the image. All this while still keeping the original image in view for readers to process (Moos, 2014; e.g., see www.thinglink.com). On-line newspapers and magazines are replete with moving parts in the form of graphs that allow readers to simulate different scenarios, showing what happens to an outcome, such as the spread of a disease , when one or another of the causal factors behind the disease is altered. Digital science texts for secondary (and even elementary) students include simulations that dynamically illustrate what happens to one human body system (the circulatory, for example, when variables in the other systems (the respiratory or the muscular) change.

When hyperlinks are used, they introduce another complexity—the need to navigate from one to another location within the text or even between texts (Goldman & Rakestraw, 2000; Salmeron, Canas, Kintsch, & Fajardo, 2005). In scrolling texts, hyperlinks move the reader up or down, and in paginated computer text, hyperlinks often take readers to the next screen of a passage to simulate page turning. Hyperlinks introduce another important dimension, *linearity*, that truly distinguishes print on paper from text on screen. Print on paper is decidedly *linear*; things come one after another, and it is only with great effort, and usually little effect, that writers refer readers to other sections or features within a text (a writer, for example, might tell a reader who is reading page 62 to look back to the figure on page 54 to note a salient detail). Dynamic texts are much more likely to encourage, if not require, a *non-linear* journey through a text, although expert readers of science text, and informational text more generally, often carve out their own non-linear pathways by reading the conclusions of a piece first or scanning for headings that will direct them toward their particular interests.

This dynamic character of some texts becomes even more important if an assessment requires a reader to navigate between multiple texts in order to note important similarities or differences between ideas, perspectives, or images (Barzilai & Zohar, 2012; Marttunen & Laurinen, 2006; Perfetti, Rout, & Britt, 1999). This dynamism can be engaged in both mundane tasks, such as determining how the major exports of Spain differ from those of Portugal, and in more critical tasks, such as comparing two first-hand accounts of a schoolyard scuffle to analyze how each assumes a different lens or set of assumptions about what happened and who is at fault.

Clearly, reading within and across multiple texts that contain both static and dynamic (multimodal, nonlinear) textual elements makes reading more complex. As well, texts may contain conflicting ideas and different stylistic features that further contribute to complexity. Readers must work actively within these text arrangements to construct meaning and create a situation model for a particular reading purpose. Hence the term *complex textual environments* provides an apt label in representing the set of texts, textual and stylistic elements, and the ideas constructed by readers as they move from one text to another, seeking relevant information and integration to complete a relevant assignment or examination task. As Coiro (2020) suggests:

> Multimodal internet texts, internet search tools, and animated digital advertisements alongside on-screen static texts create a dizzying array of possibilities for intertextual and multimodal connections...(Cope & Kalantzis, 2000b).The combination of multiple text types further complicates how to characterize text as part of digital reading in online spaces (Hartman et al., 2010; Hartman & Morsink, 2018).

These matters of designing specific assessments for varied textual environments are addressed in Chapter 4.

Text Complexity and Readability

In developing assessment blocks with previous frameworks, NAEP has taken a multifaceted approach to assessing the complexity and accessibility of texts (what has been referred to as readability for over a century) for students in each of the three grade levels assessed—4, 8, and 12. The 2025 NAEP Reading Framework continues this tradition but refines it in light of recent scholarship and policy developments. Consistent with the perspectives for both curriculum frameworks and assessments that have emerged in the last decade in response to new standards adopted by many states (cite a few states) and national consortia like the Common Core State Standards (NGA/CCSSO, 2010) and disciplinary groups (the Next Generation Science Standards, the National Council for the Social Studies), the 2025 NAEP Reading Framework adopts a three-pronged approach to evaluating the "appropriateness" (i.e., in terms of the conceptual and linguistic challenge) of texts under consideration for inclusion. NAEP will rely on a combination of quantitative, qualitative, and reader attributes (what were labeled reader-task connections in the CCSS).

Quantitative approaches rely on an algorithm to create either a single score or a small set of scores to estimate the difficulty readers might have understanding a particular text. The most common single scores are Lexiles (Stenner, 1996) or grade level designations, such as the popular Flesch-Kincaid (Kincaid, et al., 1975). Lexiles provide a point on a scale running from 0 for kindergarten text to 1000 for obscure scientific or legal documents that can be understood by only a handful of experts. Readability formulas like the Flesch-Kincaid usually convert their numerical scales to a grade level scale (from 1.0 to 20+, for example) to convey the idea of the typical student who would be able to understand a text that scaled at a particular grade level. Increasingly readability systems provide both an overall score and a small set of scores (e.g., Graesser, et al., 2014; Sheehan, et al., 2014).

For the 2025 NAEP Reading Assessment, NAEP will investigate the validity and utility of various quantitative indicators, including several of the more recent, more complex, and nuanced measures (see Hiebert & Pearson, 2014; Nelson, Perfetti, Liben, & Liben, 2012) indicators, such as TextEvaluator (Sheehan, Kostin, Napolitano, & Flor, 2014) and the Coh-Metrix Text Easability Assessor (Graesser, McNamara, Cai, Conley, Li, & Pennebaker, 2014), to select one (or more) that best fits the needs of NAEP—and that complement the approaches that NAEP uses to examine the qualitative facets of text complexity.

Similarly, NAEP will expand the range of qualitative tools currently in use (NAGB, 2009)—to include even more careful examination of the language used to render key concepts and the relationships among them accessible to readers. This is particularly important in light of greater emphasis in the 2025 NAEP Reading Framework on discipline-specific texts, settings in which language exerts substantial influence on the accessibility of texts for the general population of students as well as for specific groups, such as English learners and students with disabilities. The general approach employed in applying qualitative analyses of complexity is to train analysts to use specific criteria to unearth linguistic (largely vocabulary, syntax, or discourse) features that serve either as barriers or bridges to comprehension. Barriers can include rare words, obscure syntax (e.g., negative conditional clauses), or complex rhetorical frames for large sections of text (e.g., a conflict-resolution scenario). Bridges, by contrast, might include a diagram, an internal definition of a rare word, an explicit clue word like “unless” to signal the relationship among ideas, or explicit naming of the parts of a conflict-resolution frame. NAEP has employed qualitative analysis of text complexity since the early 1990s, when it adopted story grammars (Thorndyke, 1977) and concept mapping of the underlying propositional structure of non-fiction texts (Wixson & Peters, 1987) as primary qualitative analysis tools. According to the 2009 Item-specifications guidelines (NAEP, 2009), passage mapping is routinely conducted as a part of the passage selection process:

> Passage mapping procedures should continue to be used to identify appropriate passages for the 2009 NAEP Reading Assessments. Methods used in previous assessments have been expanded for the new assessment. Mapping procedures result in a graphic representation of a possible stimulus selection that clearly highlights the hierarchical structure and the interrelatedness of the components of the passages. Story mapping, for example, shows how the setting of a story is related to and contributes to the development of plot and theme. Concept mapping shows the structure of informational text, along with the concepts presented and

> the relational links among concepts. Organizing information hierarchically within a passage allows identifying the various levels of information within a text so that items can target the most important aspects of what students read. (p. 17)

For the 2025 NAEP Reading Framework, these successful practices from the past should be supplemented with more recent developments, particularly those deployed by PARCC and SBAC in developing their assessments (Hain & Piper, 2016).

Finally, NAEP will conduct analyses for what have been called reader-task considerations (NGA-CCSSO, 2010) or reader attributes or text-task scenarios (Valencia et al, 2014). All three of these approaches ask the question, "for whom, in what specific contexts, and with what levels of support are specific texts more or less accessible, i.e., harder or easier to comprehend?" These approaches situate text complexity within the sociocultural model outlined in Chapter 2 by noting that while factors inside the text may render it more or less complex, factors outside the text may render it more or less accessible to readers. In other words, there can be easy questions about difficult text, and difficult questions for easy texts. Furthermore, a hard question, for either an easy or difficult text, might become easier in the presence of a particular scaffold, pointing to the important role that scaffolds can and will play in the 2025 NAEP Reading Assessment. These issues are more design than definitional matters; hence they will be addressed and illustrated more completely in Chapter 4.

A completely new indicator, navigational complexity (Coiro, 2020), will be added to the complexity agenda for the 2025 NAEP Reading Assessment. It tracks the number and nature of moves a reader has to make within and across texts in order to consult information needed to complete a particular test item; it will be especially important for multiple-text blocks, and particularly for items that require comparison, contrast, or causal reasoning—all tasks that require readers to move back and forth between text sections both mentally and physically.

Taken as a whole, this set of complexity indicators calls for an extra measure of careful study on the part of NAEP as developers begin to use them to assess the appropriateness of passages for use in assessments built from the 2025 framework.

Comprehension Items: The Role of Comprehension Targets in NAEP Reading

Comprehension Targets are used in NAEP to generate test items—the questions that students respond to as they take the test. Test items provide the observable data (answers to questions) that NAEP uses to draw inferences about how well students engage in the comprehension processes detailed in Chapter 2—building a text base and a situation model. Across the Comprehension Targets, readers are asked to make different kinds of inferences (e.g., involving anaphora resolution or discerning causal relationships), using various textual and intertextual elements (e.g., information presented locally within texts, information drawn from across texts, or information presented in text and graphical features) and different types of prior knowledge (e.g., discrete factual knowledge or heuristics, such as those needed to understand or critique disciplinary texts). Although the targets tend to involve more challenging forms of inferential reasoning and tend to demand that readers bring more sophisticated forms of prior knowledge to bear across the targets, it is not accurate to think of the targets as a continuum of complexity.

Each target can involve a range of difficulty, depending on the particulars of the items in relation to the texts they are designed to probe.

Locate and Recall

The first set of Comprehension Targets are *Locate and Recall*. In order to comprehend, readers need to identify important information and form connections as they move through in a text. In addition, readers often need to locate information to fulfill a particular purpose, aid recall, and repair understanding. These kinds of processing help readers build a literal understanding of what the text "says"; in terms of the sociocultural model, it is the result of that processing— a text base (Kintsch, 1998).

Items assessing Locate and Recall targets typically focus on literal information positioned in a single location, such as a sentence, a paragraph, adjacent paragraphs, or a single graphic. Nevertheless, readers may need to navigate across different pages or documents to find the relevant information. Readers might be asked to respond to questions by relying on information that is explicitly presented in the text or graphics. For example, they might be asked to recall or locate discrete information about characters or settings in a story; to recount information from a segment of a video; or to locate a specific piece of information from a table in an expository text. Student responses are informed by what has been read and remembered, and by searching for information upon being prompted by the item. Locate and Recall targets become more complicated with the dynamic texts, which are often populated by hyperlinks and multimodal representations, that one finds in digital environments. Locating a text segment that is "right there" might be on a single page in a static text, but it might be on an altogether different page or distant part of a scroll in a dynamic text environment. These more complex design features are discussed again in Chapter 4.

Locate and Recall items can require students to make inferences across text segments that are near one another in the text. Specifically, readers may be asked to make "straightforward" local inferences that involve connecting ideas, for example, recognizing that the "she" in sentence 2 refers to Daniela in sentence 1 or inferring that two adjacent sentences are linked causally (B caused A) or chronologically (A happened before B). Finally, readers may be asked to infer the meanings of unfamiliar words by forming connections to information in the sentences immediately surrounding that word. In the language of Chapter 2, these local, straightforward (PIRLS, 20xx) inferences are a part of completing the text base.

Within the 2025 NAEP Reading Framework, Locate and Recall items may ask readers to:

- *Locate specific pieces of information by skimming, navigating, or searching one or more texts*
- *Find information in dynamic and multimodal texts*
- *Recall and record information*
- *Identify the nature of the relationship between pieces of information explicitly stated in text*
- *Discern the meaning of an unfamiliar word based on context*

Integrate and Interpret

The second set of Comprehension Targets describes what students do as they *Integrate and Interpret* information from one or more texts. Integration includes two venues: internal to the text and external to the text. The former involves making connections across many sentences, paragraphs, and even sections to synthesize ideas under a common theme (this story is about justice) or main idea (this text is all about how food goes from the farm to tables in people's houses). The latter form of integration involves readers in connecting those broad themes and ideas within and across texts to their knowledge, experiences, and perspectives—and sometimes with those of others. This kind of integrative processing helps readers retain information as they continually connect new textual information with prior knowledge. Through many iterations of this external integration (text to knowledge) cycle, readers develop global meanings, such as themes, lessons, patterns, and concepts. The result of these processes is the development of a coherent mental model of the text, or a situation model (Kintsch, 1998). Once the situation model for a text is built, readers may incorporate the new ideas in the model into their existing memory structures (learning in everyday parlance), where it will be available for future cycles of comprehension and learning from text.

In responding to Integrate and Interpret items, readers may be asked to form connections and establish chains of reasoning across different sections of a text or across multiple texts, including across different modalities. For example, they may be asked to form connections across hyperlinked documents, or between prose and graphical or multimodal representations. In order to engage in these processes, readers may be required to navigate complex hyperlinks or multimodal elements, such as video or interactive graphics. Integrate and Interpret items might also ask readers to recognize how specific features of language signal relationships or viewpoints within a text. For example, readers might be asked to make judgments about characters based on the adjectives used to describe the theme or to rely on signal phrases (e.g., "to the contrary") to understand the connections among ideas.

In Integrate and Interpret items, readers are asked to discern implicit connections within and across texts, relying on their understanding of the ideas in the texts, their disciplinary knowledge, their knowledge of text genres, and even their knowledge of how language works to convey the conceptual complexity of the ideas represented. When readers Integrate and Interpret, they compare and contrast characters and settings, examine causal and chronological relations across aspects of text, or formulate explanations for events or information in texts. For example, items may call on readers to form predictive or explanatory inferences about a character's behavior by relying on multiple pieces of information about that character's history and dispositions, or they might describe how the setting of a story contributes to the theme. Items might also call on readers to construct interpretations based on graphical representations that use sophisticated features, such as symbols. Readers may be required to engage in a "double integration" in inferring the meaning of a word in context: using context clues from the text to make connections plus invoking morphological knowledge of one or parts (the root or an affix) of the target word to arrive at a plausible meaning. Also invoked in Integrate and Interpret items are processes related to the development of global meanings across a text or texts. These might include the identification of morals, themes, or concepts (e.g., what the author is trying to say about families; how human behavior is implicated in climate change, etc.)

In order to respond to Integrate and Interpret items, readers need to develop an elaborated and cohesive understanding of one or more texts. They accomplish this by relating text ideas to knowledge from other sources, such as their previous learning and related, accumulated prior knowledge (Alexander, 2012; Lee, 2011). Readers draw on their ideas about how texts are constructed (e.g., genre knowledge and knowledge of textual features) and how graphical and multimodal representations work to communicate meaning using a variety of organizational, symbolic and dynamic features. In all cases, readers' interpretations are informed by the purpose communicated in the item, contextual characteristics such as the situation presented in the text, and their own knowledge and experiences.

Intertwined with integration are interpretation processes that readers use to clarify, explain, and elaborate insights gleaned from reading. As readers actively build and incorporate ideas within and across texts, they monitor and revise their thinking about ideas from earlier text in light of new information.

> *Within the 2025 NAEP Reading Framework, Integrate and Interpret items may ask readers to:*
> - *Create summaries*
> - *Form explanations and generalizations*
> - *Draw conclusions*
> - *Offer interpretations*
> - *Formulate questions*
> - *Make predictions*
> - *Provide evidence that supports particular interpretations or conclusions*
> - *Generate alternatives (e.g., alternative interpretations of an implied message or moral or alternative actions that a character might have taken)*

Analyze and Evaluate

The third set of comprehension behaviors, *Analyze and Evaluate*, describe the processes associated with examining and assessing ideas during and after reading of one or more texts. Thus, Analyze and Evaluate items require that readers work with the situation model of text(s) that they have constructed—and then add a dimension of interrogation. As readers Analyze and Evaluate, the focus remains on the text or texts, but the reader's purpose is to scrutinize the text, using different analytic tools and perspectives. The reader may *analyze* by closely examining the choices an author makes about content and form and how those choices affect meaning, e.g., asking a reader to demonstrate how the author's choice of words (the major *strode* into the room, he did not walk or creep in) shape readers' interpretive options. The reader may then use those analyses to *evaluate* by judging various elements of text based on a conception of quality as situated within, and varied by, a sociocultural context, e.g., was the author's use of language effective in shaping how you viewed the major, or did the author provide sufficient evidence to support a claim (Meola, 2004; Ostenson, 2014; Wineberg & McGrew, 2017).

Items designed to assess readers' Analyze and Evaluate processes might have readers make judgments about the coherence, credibility, and quality of one or more texts. For example, readers may be asked to make judgments about the effectiveness of an author's use of evidence or figurative language, or they may compare accounts that represent conflicting ideas or perspectives. In future versions of NAEP Reading, readers might be asked to use a set of digitally connected resources, using information from one text (that they might have to locate on their own) to analyze or evaluate ideas in a second text (that might be provided); for example, readers might analyze a particular use of language or author's craft in light of an author's biography, broader body of work, or affiliations. They might be asked to read with an eye to the author's underlying beliefs and goals that are communicated only implicitly within a text.

Analyze and Evaluate items might ask students to focus on small segments or dimensions of a text—for example, considering specific language choices. Alternatively, they may invite readers to apply analytic, evaluative, and reflective tools across entire texts or sets of texts. In either case, Analyze and Evaluate items may invite readers to revisit information to deepen their understanding or to seek out new information in complex digital sources.

In order to engage in Analyze and Evaluate processes, readers must view texts in relation to knowledge from other sources. These may include their existing knowledge base (Alexander, 2012; Lee, 2011) or tools and criteria such as those for literary analysis, historical reasoning, or scientific argumentation (Lee & Sprately, 2010; Greenleaf et al., 2016; van Drie & van Boxtel, 2008). Readers also draw on their personal knowledge about and preferences for particular rhetorical strategies, such as the use of language, organization of text, or articulation of claims and evidence. Readers leverage critical reasoning skills in applying these tools, or types of knowledge, to one or more new texts.

Another way to think about items generated from the lens of the Analyze and Evaluate comprehension target is that they invite different forms of critical reading or textual interrogation. An item might ask readers to critique factors related to internal markers of quality and conformity with genre expectations. For example, readers might evaluate the logic within a scientific or historical text, including the degree to which the evidence provided by the author to support the claims made is relevant and trustworthy, and linked to the claim with clear reasoning. Readers may also be asked to apply more external critical lenses that invite interrogation of sociopolitical and ideological dimensions of text. For example, readers of historical texts may be asked to examine an author's assumptions or agenda, or they may be asked to analyze implicit matters of power differential (who is privileged, who is marginalized, and who is absent from this text?). Within the 2025 NAEP Reading Framework, both the internal and external lenses are promoted by items generated by applying the Analyze and Evaluate comprehension target.

> *Within the 2025 NAEP Reading Framework, Analyze and Evaluate items may ask readers to:*
> - *Revise understanding in light of a new piece of information*
> - *Reconcile inconsistencies across a text or texts*
> - *Judge the plausibility of an event within a text*
> - *Evaluate how the author uses language, including vocabulary, and features of text to achieve a particular purpose*
> - *Determine significance*
> - *Argue for or against particular interpretations*
> - *Evaluate the relevance and strength of evidence to support claims.*

Use and Apply

The final set of Comprehension Targets, Use and Apply, reflect the impact of comprehension, in which understandings acquired during reading are used in new situations or applied in the development of novel ideas and products (cite RfU folks like CCDD and READi plus our big review). This set of targets reflects contemporary understandings that comprehension is best viewed as a series of productive processes that culminate in forms of personal and social production and engagements that "transcend the original activity" (Calfee & Miller, 2005, p. 218). That is, readers read to *do* things in the world outside of the text. These targets also reflect the ongoing evolution of the NAEP Reading definition, which in the 2009 Framework noted that readers "use meaning as appropriate to the type of text, purpose, and situation." In 2025, the principle of applying knowledge is incorporated into the test item portfolio.

Items designed to assess Use and Apply processes might ask readers to use ideas they acquire through reading as sources of information to develop a different kind of text for a new audience. For example, having read a set of commentaries, readers might be asked to produce a blog-type message for a public audience that captures the most relevant information. Readers might also be asked to use one or more texts as a model for developing a new text or graphic representation. For example, they might use a data table from one text to represent numerical information from a second text. Readers might also be asked to produce critical arguments about important social issues, such as the potential benefits and harms of digital media, bringing together their own experiences in digital environments with information from conflicting sources. Within a literary activity block, readers might be asked to rewrite an aspect of a story with a particular goal. In order to respond to Use and Apply items, readers may need to navigate complex static and dynamic text environments involving multimodal texts and to make judgments of reliability and quality as they use sources to inspire or anchor their own products.

As they engage in Use and Apply processes, readers must consider how to reframe ideas from their reading and their prior knowledge and experience into a different product for a certain purpose and audience (Marzano, 1988). As readers reflect on how to respond, they take into account their purposes, the norms established by any genre and disciplinary conventions, as well as expectations about what is deemed appropriate and compelling to members of the target audience (Gee, 2001; Goldman et al, 2011; Moje, 2005). While engaged in critical thinking processes to turn text understandings into tangible products, readers may also reveal information about their "metastrategic competence" (Kuhn & Dean, 2005), that is, their ability to reflect on their knowledge of why and in what circumstances particular choices or strategies might be appropriate according to one's goals (Sparks & Deane, 2014). For example, they need to understand how differences in individuals' values might shape the ways they use media and how they interpret messages and, thus, what strategies might be most effective for communicating with a particular sociocultural group. In developing their responses, readers will be asked to make choices about content, language, and/or presentation.

Culminating Tasks. The 2025 NAEP Reading Framework continues a consistent feature for a small set of reading blocks in the 2019 administration of NAEP. In that assessment, Scenario Based Tasks had a culminating task, an end-of-activity-block exercise that the students completed with their avatar classmates. It resulted in a product (almost always a written piece) that was related directly to the activity-specific purpose-setting experience that initiated the block activity, such as,

> Read these three texts to acquire information that you might use, along with your avatar classmates, to develop an itinerary for a trip to visit the pyramids of Egypt, taking care to give reasons for the specific sites you will visit. Your final product will be a letter to the Egyptian embassy, seeking funds for the trip you and your classmates plan to take.

Culminating tasks are grouped under the Use and Apply label because they represent a kind of "elevated" Use and Apply target.

> *Within the 2025 NAEP Reading Framework, items generated for Use and Apply targets may ask readers to:*
> - *Produce a novel solution for a problem described in readings*
> - *Apply ideas acquired though reading to a new problem or context*
> - *Take on a different perspective (e.g., a scientist vs. an historian; a different character in a story)*
> - *Rewrite or re-represent information in a purpose or audience-driven way*
> - *Rewrite a description of a character, using different language and words, to convey a different interpretation of their character or actions.*
> - *Use an understanding of legal and ethical principles (e.g., intellectual property or intellectual freedom) to develop an opinion on a current or past issue —or to plan a form of social action*

Language Structures and Vocabulary

Attention to dimensions of language structures and vocabulary, including discourse, semantic, and morphological knowledge and skill, are embedded in item development considerations within the Comprehension Targets framework above. In the near term, developers should continue to distribute these items throughout the testing blocks (and to classify them according to which of the four targets they best reflect), with a goal of devoting about 15 percent of total items to language structures and vocabulary-focused knowledge and skill. In the longer term, NAEP should, through special studies, explore the possibility of a more dedicated approach to the assessment of language structures and vocabulary, given the significance of language to reading comprehension at every stage of reading development. One way to achieve this goal would be to engage in special assessment efforts on a cyclical basis (e.g., every five years). In these assessment efforts, more robust assessments of the discourse, semantic, and morphological dimensions of language could be administered and used to better understand U.S. students' achievement on this critical dimension of reading comprehension and to further explain NAEP assessment results (see the extended essay on Language Structures and Vocabulary in the Development Panel Resource Library).

Operationalizing Explanatory Variables

Another major shift in the 2025 NAEP Reading Framework is a greater emphasis on explanatory variables in the analysis and reporting of student performance on NAEP Reading (see Chapter 5). In the past, NAEP Reading focused more on reporting than explaining performance, with the greatest emphasis on monitoring achievement trends over time. NAEP has reported reading performance in relation to certain factors, what NAEP has called contextual variables, such as race, language status, socio-economic status, region of the country, and, for special NAEP initiatives, states and large cities. It has even issued special reports that break down performance by variables under the rubric of habits and attitudes (e.g., How much do students like school; how often do they read for pleasure at home, go to the library, and/or read or write on their digital devices?). But because of time constraints on the administration of survey items, NAEP has not had the capacity to include many of the potentially useful explanatory variables that could be easily derived from the sociocultural model.

Explanatory variables in the 2025 NAEP Reading Framework fall into two important categories. First, there are a set of **reader attributes** related to the knowledge, interest, motivation, engagement, habits, attitudes, language competence, and skills and strategies that individual students bring to the reading act. Second, there are a set of **environmental variables** related to contexts that influence individual student performance, some emanating from home and community settings (e.g, home language, socioeconomic status (SES), parent education, and participation in community activities) and others related to the school environment (e.g., opportunities to learn, school and classroom supports for learning, and peer relationships).

The research base for the reader attributes and the contextual variables is robust and extensive (Guthrie & Humenick, 2004; Guthrie & Klauda, 2015) and merits attention in helping to explain reading outcomes. However, the questions on many questionnaires are operationalized in such a simplistic manner that they do little but perpetuate stereotypes; for example, reporting performance as a function of race or ethnicity without disaggregating the category by SES or educational opportunity does little to point toward policy solutions.

NAEP has an opportunity to more fully describe the nature of reading achievement by taking advantage of recent advances in measuring explanatory variables and weighing their impact on students' reading comprehension performance. What is needed is a shift in emphasis in the types of data collected to assess students' perceptions of their experience with the NAEP Reading Assessment, coupled with a commitment on the part of NAEP to devote more attention to the explanatory face of NAEP assessment, especially in the plans for reporting NAEP results. There are three places in the NAEP assessment space in which NAEP could bolster its efforts to increase its emphasis on explanatory variables.

Survey Responses to Reading Related Issues

The 2025 NAEP Reading Assessment will continue to use survey items to gather both general and reading specific views from students about their in and out of school experiences. These items are part of the questionnaire that all students complete at the end of their participation in the assessment. To better explain students' comprehension performance, NAEP should include more items that focus on student support, recognition, and opportunity to learn in school settings. Just such a proposal is outlined in Chapter 5, which considers how to report—and explain—the results of the NAEP Reading Assessment.

Block-Specific Measures of Students' Perceptions of Their NAEP Experience

Recent research from within the NAEP research community (Educational Testing Service, 2019) has resulted in a new form of student surveys to assess an array of reader attributes related to performance on a particular activity block. Assessing the role of reader attributes on student performance as they complete a specific block activity can produce more explanatory information about student performance. For instance, students could be asked to rate on a scale of 1-5 how hard they tried to comprehend a specific text or answer specific questions, their interest in reading the text, and/or what they knew about key ideas before reading the text. This approach is fast and efficient. Even fourth-grade students can make 8 or 10 judgments per minute. It has been shown to be relevant to comprehension; scores on these metrics explained a reasonable proportion of variance in comprehension scores in a pilot study at Educational Testing Service (2019). For students, making a personal judgment about how much you do something "in general" (How well do you do on tests?) is much more abstract than making a highly specific, particular judgment (How well did you do on this test that you just finished?). NAEP should set the highest possible priority on a special study to evaluate the validity, reliability, and utility of this "task-specific" approach to gathering information about students' affective profile.

Process Variables

With the advent of the NAEP Reading Assessment digital platform, NAEP has the capacity to collect process data, such as keystroke trajectories, how students navigate through the passages and items, and the amount of time spent on particular pages, images, or other facets of the assessment. These data can be used to draw plausible inferences about motivation and engagement (how deeply are students attending and to what?) as well as metacognitive behaviors (patterns that imply monitoring, for example, or systematic versus random approaches to item response?). Data from these process observations can be aggregated and used to predict or explain student performance on comprehension items. Early evidence from NAEP Reading administrations in 2017 (Feng, 20sx) suggest that these data do differentiate between more and less successful readers; for example, readers above the 75th percentile tend to be systematic in shifting attention from text to items while those below the 25th percentile tend to exhibit random navigational patterns. As with the block-specific measurement of reader variables, supporting an active program of research on process variables is rich in possibilities, especially given that technology can collect the process data automatically.

Summarizing the Assessment Construct

At the outset of this chapter, Exhibit 3.1 provided a blank version of the matrix depicting the relationship between the features of the sociocultural model and the components of the Assessment Construct. Exhibits 3.3 and 3.4 represent a "filled out" version of the matrix. Exhibit 3.3 summarizes the mapping for what might be called the assessment design characteristics—those components one orchestrates in order to build an assessment, which, of course, is precisely what Chapter 4 provides. Basically, Exhibit 3.3 answers the question, How, and how well, is the sociocultural model represented in the Assessment Construct? The short answer is that the sociocultural model is well-represented. As evidenced by the low incidence of blank cells in Exhibit 3.3, key features of the sociocultural have been incorporated in all components of the Assessment Construct. This bodes well for the prospect that the assessment blocks developed from the NAEP Reading 2025 Framework will reflect the equity-aligned, asset-oriented, best-foot-forward ethos of the most current views of the nature and development of reading comprehension.

Exhibit 3.4 maps the explanatory variables—the block-specific mini-probes, the survey items in the culminating questionnaire, and the process variables (tracking individual journeys through the assessment) onto the features of the sociocultural model. As it turns out, the mapping for the explanatory variables serves a slightly different function than do the design components, for they foreshadow the indicators that might be used to explain students' comprehension performance on the three scales that will be reported.

Chapter 4 takes the next step by illustrating the use of key design principles and practices that will allow NAEP test developers to deliver this important promise to America's students.

Exhibit 3.3. A Complete Mapping of the Sociocultural Model onto the Assessment Construct

	Features of the Sociocultural Model			
Assessment Construct Components	Reader	Text	Activity	Sociocultural Context
Activity Structure & Purpose Setting	Introduce social elements, such as a "guide" at the opening of a testing block, digital (avatar-based) peers, and engagement with contemporary social issues.	Include varied texts that align with the contexts, activity structures, and purposes established for the testing block.	Establish authentic contexts and activity structures for each of the testing blocks and, often, for specific texts.	Invoke rich contexts (discipline-related and otherwise) as a way of situating testing blocks in social and cultural settings that involve reading.
Scaffolds	Develop knowledge scaffolds that provide readers with a baseline level of knowledge needed to engage in the testing block.	Provide scaffolds to increase broad access to texts, such as providing definitions of key words and offering lookback buttons.	Provide social scaffolding (avatars) in SBTs and scaffolds that clarify the task and expected responses.	Reflect the kinds of social supports that are commonly part of reading in school and community contexts.
Texts	Select texts and design activities that provide opportunities for readers with varied backgrounds to find connections to their experiences and identities.	Include texts from a wide range of genres, modalities, formats, and disciplinary traditions.	Include varied texts that align with the contexts and activity structures established for the testing block.	Include a variety of texts that represent the range of reading related to particular disciplinary traditions and reading purposes.
Comprehension Items	Address an array of skills and strategies related to comprehension, including literal, inferential, analytical, and critical responses.	Query different types of comprehension within and across texts and different aspects of the texts, including local and global features and meanings.	Provide opportunities for some degree of reader choice regarding the order of questions and the format of questions.	Reflect a view of the outcomes of reading as influenced by factors within and outside of the assessment block.

Exhibit 3.4. A Complete Mapping of the Sociocultural Model onto the Assessment Construct: Explanatory Variables

	Features of the Sociocultural Model			
Assessment Construct Components	Reader	Text	Activity	Sociocultural Context
Explanatory Items: Block mini probes	Gather data about students' perceptions of their level of prior knowledge, interest, effort, and confidence as they completed a block.	Gather data about students' perceptions of familiarity, interest, and difficulty with respect to texts.		Gather information about whether or how often they do tasks like this in their classrooms
Explanatory Items: Survey items	Gather information about demographics, home life, motivation, and in- and out-of school reading practices.	Gather information about the amount and kinds of texts that readers encounter in and out of school settings.		Gather information about the sociocultural contexts of readers' lives and experiences in and out of school.
Process variables: Log file data gathered as a matter of course.	Track each participant's navigational journey through the assessment—reading texts and responding to items.	Compare pathways through blocks employing different sorts of texts	Compare pathways for different sorts of items, both format and Comprehension Targets.	

Introduction

This chapter will explain how the elements of the 2025 NAEP Reading Assessment construct fit together in a coherent design to reflect the 2025 NAEP Reading Framework sociocultural model. The chapter describes elements of the new design in some detail with the assumption that assessment designers will continue to adhere to measurement best practices outlined by the National Research Council (2001) and used in previous NAEP Reading assessments. These practices include incrementally augmenting the current design with new features that are carefully tested and refined over time, as has been a hallmark of NAEP development practices since the inception of the assessment.

Addressing the various aspects of the 2025 NAEP Reading Framework sociocultural model means that the assessment will continue to evolve to reflect the changing nature of reading demands in today's society. Most important, these new assessments will create new ways to invite readers with various kinds of strengths to optimize their assessment performance. This positions the 2025 NAEP Reading Assessment to maximize readers' opportunities to demonstrate reading comprehension in valid ways that reflect the changing demands of our increasingly complex society (Mislevy, 2016; National Research Council, 2018).

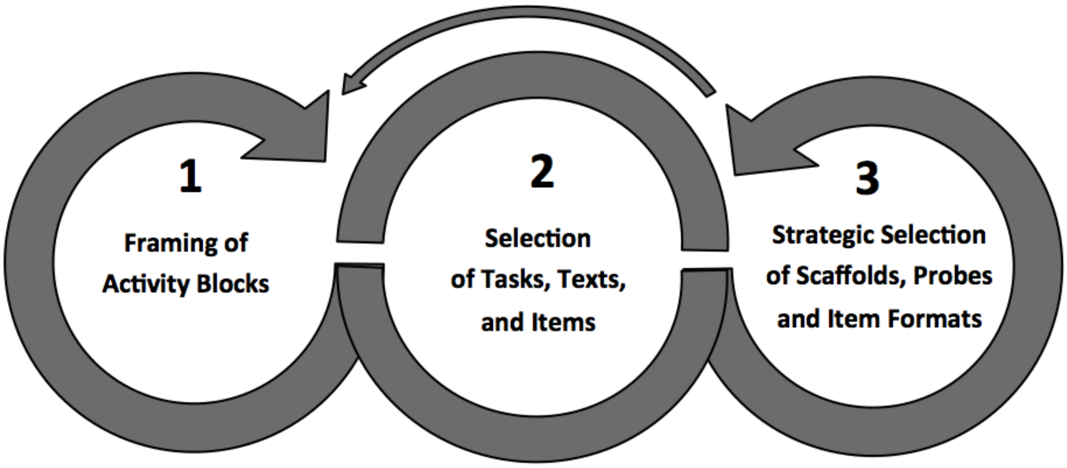
The remainder of this chapter is divided into three sections. The first describes the iterative nature of design cycles. The second provides a general overview of the development process, followed by more specific sets of guidelines to inform decisions made within and across phases of development of activity blocks, tasks, texts, items, and scaffolds. The aim is to illuminate the many complex design considerations required to articulate the assessment construct while also, to the extent possible, holding true to a sociocultural model of reading. The final section introduces two examples of how components of the assessment construct could play out in the design of activity blocks at different grade levels. These examples are featured to illustrate the range of possible design elements from which designers might sample.

Overview of Design Process

Iterative Design

Briefly, the iterative design process (depicted in Exhibit 4.1) includes but is not limited to: (a) situating each grade-appropriate reading activity in a primary disciplinary context, activity structure, and one or more socially situated purposes for engaging with texts; (b) iteratively selecting and revising purposeful tasks, texts, and items; and (c) selecting optimal locations in each activity block to support readers with varied scaffolds (see explanation of scaffolds in Chapter 3).

Exhibit 4.1. Iterative Cycles of the Design Process for Developing NAEP Activity Blocks



The design assumes that these cycles are interdependent, which means that changing elements in a single part of the cycle will likely mean changing elements in other parts of the cycle. Guidelines for selecting and balancing assessment components across grade-level blocks call attention to new considerations of the updated framework while maintaining valid and reliable ways of measuring comprehension in the years to come.

Designers necessarily face a dilemma when they begin the design process: Of the many things that must be done, which should be done first? Assuming that more than one sequence is possible, the 2025 NAEP Reading Framework suggests a plausible order, one that gives primacy to two important components of the Assessment Construct—the overarching purpose and the disciplinary context. The steps in the process are referred to as cycles, and they are considered to be iterative in the NAEP assessment development process. To say that they are iterative means that they are inherently provisional (designers rarely get everything right on the first pass) and recurring (it is almost inevitable that any given cycle will be repeated multiple times on the way to an acceptable outcome). Assessment developers should evaluate whether the suggested sequence in this chapter is viable or needs revision in response to the success or difficulty that item developers experience when they implement these development steps.

Cycle 1. Framing of Activity Blocks

Each activity block includes a single reading activity (and related sequence of tasks) that is situated in a particular disciplinary context. In this phase, decisions are made about grade level (4, 8, or 12), reading Context (Reading to Engage in Literature, Science, or Social Studies) and reading Purpose (Developing Understanding or Solving a Problem). At the outset of each activity, readers are immersed in a community-based situation in which their meaning-making from text is central to developing understanding or to solving a problem. Sometimes, readers may be assigned particular roles (student, classmate, community member, etc.), but in all cases, the scenario outlines expectations for how readers are expected to engage with particular tasks and to what end. For example, 4th graders engaged in reading literature may be asked to read several excerpts from a book and participate in a book discussion with simulated peers as they identify important events, analyze character traits, and express their understanding of how the main character changes over the course of the story. In a grade 8 assessment block focused on Reading to Engage in Social Studies, students may be asked to examine multiple texts to explore the history of urban communities, connect the history to the current event described in a news article, and use their new insights to craft an informed presentation for the general public that clarifies a historical understanding of the event.

Cycle 2. Selection of Purposeful Tasks, Texts, and Items

In this cycle, passage selection guidelines and the reading context are used as framing devices for selecting possible texts for readers as they work to accomplish task-specific purposes in the activity. Then, guidelines and principles are used to design items likely to elicit focal comprehension processes and reading outcomes (evidence of Comprehension Targets). Assessment tasks and related items are shaped to elicit a range of processes across the Comprehension Targets, while sampling across different disciplinary contexts, genres, text structures (narrative and informational), and text features (static, multimodal, and/or nonlinear). Designers must also consider that readers bring different funds of knowledge (or resources) to the task at hand. At multiple points in the development process, the framework guidelines (detailed in the next sections) are used to inform decisions about elements to include, exclude, or adapt to effectively align with activity contexts, the grade level of readers, Comprehension Targets, and tasks and items that are realistic to complete during a 30-minute reading experience in a constrained assessment space.

Cycle 3. Strategic Selection of Scaffolds, Probes, and Item Formats

In this cycle, additional sets of guidelines inform strategic decisions about scaffolds and probes selected to support and/or account for readers' knowledge, strategy use, and other reader attributes as well as considerations about which item response types are likely to elicit useful and valid information about comprehension processes. In line with the sociocultural model of reading, these efforts seek to ensure that readers have equitable opportunities to access and engage with the tasks (and related texts) and demonstrate what they can do with respect to measured comprehension processes.

Using Iterative Design Cycles to Develop Activity Blocks

As designers move through these cycles, multiple sets of guidelines inform their work (see Exhibit 4.2). Because the components of the assessment construct are so interrelated and interactive, designers need to be flexible in checking their progress as they move back and forth within and across the phases. The remainder of this chapter lays out detailed assessment design considerations, in accordance with the guidelines outlined in Exhibit 4.2.

Exhibit 4.2. Iterative Design Guidelines for the 2025 NAEP Reading Assessment

Cycle 1. Framing of Activity Blocks and Related Purposes

- **Design activity blocks** aligned with disciplinary context, situation, overarching purpose, one or more task-specific reading purposes, and anticipated comprehension outcomes relevant for the grade level

Cycle 2. Selection of Purpose-Driven Tasks, Texts, and Items

- **Select texts (and textual elements)** aligned with context, reading purposes, and comprehension processes
- **Design tasks and assessment items** that engage students with selected texts in ways that elicit anticipated comprehension processes and products

Cycle 3: Strategic Selection of Scaffolds, Probes, and Item Formats

- **Select elements that support and take into account reader attributes** (e.g., knowledge, language, experiences, motivations) to ensure that readers have equitable opportunities to access and engage with tasks (and related texts) and demonstrate what they can do with respect to measured comprehension processes

Cycle 1. Framing of Activity Blocks

Designating Primary Context. An activity block corresponds to a 20-40 minute assessment activity, and every block has a single overarching purpose. All grade-appropriate activity blocks may sample from a wide range of texts, including reading materials that students might use in their work and everyday lives (see for example, Creer, 2018 and Dobler & Azwel, 2007) representing one or more genres, modalities, or disciplines. However, one of the three Contexts (Reading to Engage in Literature, Science, or Social Studies) is identified as the primary focus of each activity block. In some cases, designers may sample from texts associated with more than one discipline; in these cases, the activity is characterized by both a primary reading context that shapes the overall reading Purpose (e.g., Reading to Engage in Literature) and a secondary context identified by one or more interdisciplinary or cross-disciplinary topics or genres. For example, an activity about different weather patterns may be designated primarily as Reading to Engage in Science for reporting purposes, and then assigned a secondary Context of Reading to Engage in Social Studies because one of the tasks involves reading to understand how weather has historically shaped life and ecology in the area.

The distribution of disciplinary Contexts by grade level should vary according to the approximate amount of time that students in the U.S. are engaged in the respective Context at that grade level. Exhibit 4.3 shows the proposed distribution of reading Contexts at each grade level.

Exhibit 4.3. Distribution of Disciplinary Contexts by Grade Level

	Reading to Engage in Literature	Reading to Engage in Science	Reading to Engage in Social Studies
Grade 4	50%	25%	25%
Grade 8	40%	30%	30%
Grade 12	34%	33%	33%

Designating Activity Purpose. After situating the activity block in one of these three reading Contexts, each block is then designated as having one of two overarching Purposes: Reading to Develop Understanding activity blocks are designed to measure what readers do when asked to deeply read—literally, inferentially, interpretively, and critically—in or across the disciplinary Contexts. Reading to Solve a Problem activity blocks are designed primarily to assess what readers do when asked to demonstrate understanding across multiple texts and perspectives while solving a problem. Reading to Solve a Problem activities entail developing understanding but in the service of using this understanding to take a specific action or create a product, often in the form of a text or a presentation. In the NAEP assessment, these culminating products reflect the original activity-specific purpose established at the outset of the activity block.

Both types of activity blocks specify particular situations that shape how and why readers are asked to engage with the set of tasks and texts in one of the three disciplinary Contexts. In either activity block, reading situations are designed to help readers prepare for and approach a particular reading activity as they read to develop understanding or solve a problem. In special studies designed to evaluate NAEP constructs, situating reading in purpose-driven tasks showed potential for promoting interest and engagement (Educational Testing Service, 2019). In either type of activity block, tasks and items are intentionally crafted to elicit responses aligned with specific comprehension processes and products relevant to the socially situated reading context of that particular activity. Guidelines for distributing primary Contexts across activity blocks are depicted in Exhibit 4.4, followed by further elaboration of what characterizes each kind of activity block.

Exhibit 4.4. Primary Context by Activity Block

	Reading to Engage in Literature	Reading to Engage in Science	Reading to Engage in Social Studies
Activity Block			
Reading to Develop Understanding	40%	30%	30%
Reading to Solve a Problem	30	35%	35%

Designing Components of NAEP Activity Blocks. After designating an activity block's Context, overarching Purpose, and situation to assess levels of reading comprehension in line with Comprehension Targets, developers follow additional guidelines to select from an array of features for each component of the activity (e.g., reader role, activity, task, text, and scaffolds). Ultimately, there will be an array of block designs in each of the three reporting categories (e.g., disciplinary context) that include relatively complex blocks (those with a greater number of texts, tasks, and text types that could require students to choose their own reading path by clicking on hyperlinks), as well as less complex blocks.

Exhibit 4.5. Continuum of Complexity Associated with Any Given Activity Block

Fewer parameters are specified; situation is less developed and complex and student may be less immersed within it	More parameters are specified; situation is more developed and complex and student is more immersed within it

Importantly, the goal for assessments built from the 2025 NAEP Reading Framework is for all block activities to be purpose-driven. Designers will select and tag elements in the activity block along a set of features aligned with each assessment component (e.g., reader role, activity, task, text, and scaffolds). Design options for each component range from features that are more constrained and conventional to features that are more complex, dynamic, interrelated and perhaps iterative. For each activity block, tagged elements are entered into a database to ensure a balanced or recommended distribution across the entire collection of 2025 NAEP Reading Framework blocks.

In the short run, the 2025 NAEP Reading Assessment design frame needs to be broad enough to include the traditional discrete blocks that have been part of NAEP for almost a half century; this will be necessary if for no other reason than to validate new blocks developed from 2025 NAEP

Reading Framework. The goal of the new Framework, however, is to move as expeditiously as possible toward tasks that seek to embed the reader in a situation in which both purpose and the reader's role are transparent. Overall, this continuum approach offers a broad framework within which NAEP can gather data from the design frame that reflects the nature and features of the activity blocks resulting from it. In Exhibit 4.6, details are provided for how several facets of the sociocultural model might vary along the continuum illustrated by Exhibit 4.5, in order to convey a sense of how activity blocks might differ from one another.

Exhibit 4.6. Range of Design Features for Each Assessment Component with which Students Might Engage in Any Purpose-Driven Activity Block

Features	**More constrained and conventional assessment features**		**More complex, dynamic, and iterative assessment features**
Role of readers	Reader is less immersed (assigned less of a role) in a social situation that contextualizes expectations for how to engage with provided texts and tasks.	⟺	Reader is more immersed in a social situation that contextualizes expectations (or provides choices) for how to engage with provided texts and tasks. Readers may be assigned (or choose to take on) particular roles, and their role may be more specified, particularly in relation to others' roles and the expected outcome.
Activity	Less involved initiating event (hook/"way in") that focuses students' attention on a theme, question, or problem to be explored during the block (e.g., consider how a character changes throughout a story), but not all tasks within the activity necessarily work directly toward this theme.	⟺	More involved initiating event (hook/ "way in") paired with an essential inquiry question or problem to be examined (e.g., using an author interview, nonfiction texts, and a fiction story based on real issues/current events, consider why an author includes characters with different perspectives despite the author's own perspective on the issue stated during the interview). All tasks within the activity will help readers work towards this theme, question, or problem.
Task	*Inter-relatedness:* Purpose-driven tasks are situated in line with context norms but tasks are more loosely structured with less probability of readers moving back and forth	⟺	*Inter-relatedness:* Purpose-driven tasks are situated in line with context norms but tasks are more tightly structured so that one task builds on the previous; more probability that tasks are dependent on one another; more need for resetting.

	across tasks; less need for resetting. *Culminating elements:* Less involved culminating task that loosely addresses the question/problem; not a major driver of the block.		*Culminating elements:* More involved culminating task at the end of an activity that directly addresses the question or problem; major driver of the block.
Text	*Number*: 1-3 related texts; excerpts from some texts may be included rather than in their entirety *Dynamism*: More static, minimal dynamic *Linearity:* Fewer nonlinear structures to navigate within or across texts; less variation in structures across texts *Features*: Texts include a narrower range of features and fewer types of media. *Perspectives:* Less variation in content, purposes, perspectives across texts.		*Number*: 2-4 interconnected texts (or excerpts from longer texts); readers may be asked to choose only some to engage with in line with task purposes *Dynamism*: Minimal static, more dynamic *Linearity:* More nonlinear structures to navigate within or across texts; more variation in structures across texts *Features*: Texts include a wider range of features and more types of media *Perspectives:* More variation in content and a wider range of purposes and perspectives across texts.
Scaffolds	Less complex reading situation that may involve scaffolds for knowledge or motivation/social interaction but lesser need for metacognitive/strategy scaffolds.		More complex and inter-related reading situation that may involve scaffolds for knowledge or motivation/social interaction but greater need for metacognitive/strategy scaffolds.

Moving forward, it will be important to recognize how each type of activity block and its collection of texts and contextual design features is likely to continue changing as new technologies and diverse reading purposes emerge (see Coiro, 2020). To inform future task and item development, formative research studies are needed to systematically explore variations in reading performance for different kinds of readers in relation to design features within and across the two types of assessment blocks.

After crafting the framing components of an activity block (e.g., grade level, reading context, reading purpose, and situation), designers iteratively cycle through efforts to select texts and design tasks and assessment items in line with those framing devices.

Cycle 2. Iterative Selection of Purpose-Driven Tasks, Texts, and Items

Guidelines for Selecting Texts. Passages selected as stimulus material for the 2025 NAEP Reading Assessment meet rigorous criteria. The most important criterion for including texts, regardless of the discipline in which a given block is situated, is *authenticity*, as described below. Several other criteria follow authenticity very closely; the full list includes:

- Authenticity: Do texts represent the types of texts that students encounter in their reading in and out of school?
- Engagement: Will texts spark and maintain student interest?
- Social and cultural diversity: Does the sample of texts represent the array of social circumstances and cultural traditions familiar to the diverse students in U.S. schools? \Do the authors of the texts parallel the population of students in their ethnic, cultural, racial, geographic and linguistic diversity?
- Developmental appropriateness: Are the texts appropriate to the experiences of the students assessed at grades 4, 8, and 12?
- Disciplinary appropriateness: Do the texts represent those common to the disciplinary Contexts of Literature, Science, or Social Studies? Within each discipline, do they represent the range of genres/text types and text features?
- Quality and coherence: Are the texts written in a considerate manner (meaning they are well organized in ways that promote comprehension and learning [Fisher, Frey, & Lapp, 2012]), with sufficient explanation/elaboration to ensure that students will be able to find links to their own experiences, understand the texts, respond to items accompanying the texts, and learn from texts?
- Complexity: Are the language features (vocabulary, syntax, discourse and rhetorical structures) appropriate to the grade and disciplinary context? Is there a range of complexity at each grade level such that both lower-achieving and higher-achieving students will have the opportunity to encounter texts at their level of competence?

Engaging Experts in Selecting Texts. To implement the 2025 NAEP Reading Framework, NCES and its contractors should ensure both disciplinary expertise and deep knowledge about the nature and structure of texts in the text selection process. This might take the form of one or more panels of teachers and curriculum experts who can represent issues related to diversity on three dimensions: development (primary, middle, and high school levels), disciplinary Contexts (Literature, Science and Social Studies), and cultural background (the range of identities in America's schools). These panels could also include students, parents, and librarians who study children's literature. Individuals in the groups should represent diverse cultures and languages in order to collect informed suggestions for the selection of high-quality and authentic texts that permit readers to see themselves, their communities, and their experiences in the assessment materials.

Original and Commissioned Texts. Most material included on the assessment will be presented in its entirety as students would encounter it in their own reading. However, some material may be excerpted, for example, from a novel or a long essay. Excerpted material will be carefully analyzed, and minimally altered if necessary, to ensure that it is coherent in structure. At the same time, texts will be selected to represent the range of comprehension practices introduced as readers encounter different combinations of static and dynamic texts in simpler and more complex reading environments. Exceptions may be made in cases when it is difficult if not impossible to find naturally occurring texts (e.g., a viable topic does not yield readily available fictional or informational texts; an infographic is needed to complement another text). For example, it may not be possible to find the appropriate number of authentic social studies texts at grade 4 for a block activity designed to address the Reading to Solve a Problem Purpose. Exceptions may also be made to simulate an authentic text created by a child or even an adult to show their learning or express their opinion (e.g., a student-created public service announcement or a short multimodal story about life in their community); these commissioned texts serve to represent the range of voices and creative digital compositions that readers may encounter in their everyday lives. In these cases, it may be necessary for NCES and its contractors to consider hiring authors to write texts to satisfy the needs of a particular block. Consistent with the sociocultural model of reading, practicing authors should be hired to write the necessary texts; that is, individuals who write similar types of texts for public and school use and who reflect the sociocultural practices that surround the publication, reading, critiquing, and use of such texts.

Passage Length. Material on the assessment will be of different lengths, as shown in Exhibit 4.7. Passages of these lengths are employed for several reasons. To gain the most valid information about students' reading, stimulus material should be as similar as possible to what students actually encounter in their in-school and out-of-school reading. Unlike many common reading tests that use short passages, the 2025 NAEP Reading Assessment will present more complete texts that challenge students to use their strategic reading skills in ways that reflect the kinds of reading they do in authentic in- and out-of-school situations (Paris, Wasik, and Turner 1991).

Exhibit 4.7. Passage Lengths for Grades 4, 8, and 12

Grade	Range of Passage Lengths (Number of Words)
4	200-800
8	400-1,000
12	500-1,500

It is expected that in some cases, two or more texts (with static and/or dynamic textual features) will be used together to assess students' ability to interrelate them in terms of their specific claims, themes and stylistic features. In these cases, the total length of the passages will fall in the recommended passage length range for each grade. As NAEP makes progress in its capacity to handle increasingly complex digital testing situations, it is possible that readers may be asked to solve a problem by applying different search strategies (generating search terms, skimming a set of search results, or skimming a collection of four texts, for example) to locate and engage with the one or two texts most relevant for a given purpose in the assessment activity.

Because videos may be used in NAEP assessments built from the 2025 Framework, video length standards must be developed. In order to control overall block length, the best approach for setting such standards is to use average reading fluency rates for students in grades 4, 8, and 12 to calibrate video duration. For example, the average 12th grade student reads 190 words per minute. Thus, at the lower boundary of grade 12 texts (500 words), it would take the average 12th grade student a little over 2.5 minutes to read a text; at the upper boundary of 1500 words, almost 8 minutes. Thus a sensible video length range for grade 12 might be 2.5-7.5 minutes. Average fluency rates for grade 4 and 8 are 120 and 150 words per minute, respectively (Hasbrouck & Tindal, 2006).

Disciplinary Appropriateness of Texts. The most important criterion for selecting texts is that they are representative of the discipline in both content and structure, reflecting the range of genres and discourse features detailed in Chapter 3.

Discourse Communities. One of the key understandings about texts as they present themselves in disciplinary Contexts is that reading comprehension is realized as much through the conversations about text (with other readers) as it is through the unique features of the texts themselves. When readers talk about text, they animate them—breathe life into them. And those conversations, and the academic language that emerges in those conversations, is as much a part of the understanding students take from the texts as the language in the text itself. In terms of assessment, the importance of this language can only be simulated by the conversations students experience with avatar teachers and/or students. Even so, it is an important way of scaffolding readers' encounters with key ideas in the texts.

Discipline-specific Considerations. Because reporting prompted by the 2025 NAEP Reading Framework will feature scales for the three disciplinary Contexts, it is also important to specify both the variability within and the central tendencies across each context. Based on the account provided in Chapter 3 of the range of text types, text structures, and text features, Exhibit 4.8 shows the important textual elements that characterize texts in each of the disciplinary Contexts, while acknowledging that many text features are common across disciplines (as described in the next section). The responsibility of test developers, as they build the portfolio of test blocks and activity structures at each grade level, is to ensure the entire array of text types and features is considered for inclusion in the blocks for each grade level.

Exhibit 4.8. Typical Text Elements Across Disciplinary Contexts

Context	Genres and Text Types	Discourse, Language Structures, and Text Elements
Literature	• Myths, legends, and fables • Short stories • Coming of age stories • Novels • Dramas • Poetic traditions • Science fiction • Satires • Magical realism • Biographies • Memoirs	• Plot and character structures • Figurative language (symbolism, imagery, simile, metaphor, personification) • Point of view • Dialogue • Diction and word choice • Repetition • Exaggeration • Theme and message • Flashback • Foreshadowing

	• Comic books • Graphic novels • Manga • Fanfiction • Literary analyses • Literature reviews and recommendations • Author profiles and biographies	• Mood, tone, irony, paradox, and sarcasm • Visual and graphical elements such as illustrations and photographs • Multimodal elements such as narrative soundscapes • Description, exposition and narrative elements and text structures
Science	• Reports • Press releases • News briefs • Discovery narratives, biographies, and first person accounts • Raw data • Bench notes • Journal articles • Personal communications	• Linguistic frames and signals for organizing arguments, comparisons, and/or causal chains • Abstraction and nominalization (e.g., technical terms like transpiration represent an explanation sequence) • Epistemological qualification of claims: may, probably, suggests, etc. • Visual and graphical elements such as tables, graphs, equations, diagrams, schematics • Multimodal elements such as simulations or animations
Social Studies	• Primary, secondary, and tertiary text traditions (mainly in history) • Primary: newspaper articles, census data, diaries, letters, speeches, inventories and records of sale, advertisements, archival documents, cultural artifacts • Secondary: interpretive explanations of historical, social, and cultural phenomena and trends.	• Linguistic frames and signals for organizing arguments, comparisons, and/or causal chains • Abstraction and nominalization (e.g., to develop a chain of reasonings across events and happenings, e.g., this stance of brinkmanship...) • Rhetorical markers of persuasion • Lexical expressions that mark chronology or argument • Historical and ideological markers of language • Visual and graphical elements such as maps, timelines, political cartoons, photographs • Multimodal elements such as digital stories, procedural

		texts, public service announcements • Event models (how historical events are described) • Spatial (place, location) and temporal indicators (era, time, sequence, and tense)

(Note: Many text types and elements are common across disciplines.)

Standards for Coherence and Complexity of Texts Within and Across Disciplines.

Quantitative and Qualitative Indicators. In addition to following discipline-specific guidelines for text selection outlined above, efforts should be made to promote the strategic balance and selection of texts across activity blocks. This process should be informed by general standards of quality, coherence, and complexity (e.g., conventional readability criteria, reader-text connections, language structures and vocabulary considerations) and reflect contemporary standards applied to digital texts and other contemporary media forms.

When selecting passages, conventional quantitative and qualitative indicators will be consulted as ways of characterizing the difficulty of text passages with widely used tools. These may include the application of quantitative algorithms (e.g., Lexile scales, [Stenner, 1996] or more structurally sensitive indices such as Text Evaluator [2014]; Reading Maturity Index [Landauer, Kireyev & Panaccione, 2011]; and Coh-metrix, [2014]) as well as the consideration of qualitative factors such as text structure, richness of detail, vocabulary density, author purpose, writing style, and the propositional structure (relationships among ideas) of the text (CCSSO, 2010, Kintsch & van Dijk, 1983). Generally, both global (e.g., genre) and more local (rhetorical structures such as cause-effect or problem-solution or conflict-resolution) text organization as well as a reader's awareness of this organization significantly impact reading comprehension (Armbruster, 1984; Bakken & Whedon, 2002; Pearson & Camparell, 1981). Because readers use specific knowledge to identify important information in different types of texts (Wixson & Peters, 1987), designers also attend to variations in organization and cohesion in line with common text structures that are found across disciplinary Contexts (see Exhibit 4.8).

Textual ideas in any disciplinary Context should be represented with appropriate vocabulary and, where needed, texts should have useful supplemental explanatory features such as definitions of technical terms or orthographic features (italics, bold print, headings) and connective signal words (e.g., first, next, because, however). When selecting texts that blend story and factual information, including biographies, memoires, and primary sources (famous letters or speeches), attention should be paid to features within texts cuing readers' attention to structure and influencing the recall of information.

Reader-text Connections. The extent to which a reader's prior knowledge, experience, and interest connect to a text and its topic must also be considered when evaluating a text's complexity (Fisher, Frey, & Lapp, 2012), suggesting that a text is not just complex "in the abstract" but more or less complex for particular groups of readers under specific circumstances (Valencia, et al., 2014). Just as the earlier guidelines for text selection suggested that texts should be selected with an eye toward diversity, so must there be multicultural guidelines that reflect varied language traditions and communication patterns (e.g., to share stories or information, recite poetry, in silence, orally, or in chorus). However, as the sociocultural model of reading suggests, designers should aim for a flexible and diverse representation of language and discursive structures across the activity blocks rather than making assumptions about what readers will bring to any particular reading task.

Digital Standards of Coherence and Complexity. Since the NAEP Reading Assessment is digitally-based, designers need to consider additional criteria for reading materials related to the variation in textual elements:

- Graphical displays of ideas: Does the sample of texts represent the array of graphical and structural representations (e.g., static, dynamic, multimodal, nonlinear) found in print and digital reading materials? Are the displays grade-appropriate in terms of structural complexity, topic, vocabulary, number of categories of information presented, and concepts (number, familiarity, abstractness?)
- Number of texts: Do the texts invite grade-appropriate opportunities for readers to engage with ideas within different sections of the same text as well as to process ideas across two or more texts?
- Digital arrangement of texts: Do some texts contain features that allow for navigating complex textual environments (e.g., search engines, hypertexts linked across documents) to reflect what readers do when they use the Internet? Do texts with digital features dynamically change to adjust for differences in screen size of the viewer?
- Interplay of ideas and media: Does the sample of texts present opportunities for readers to grapple with and reason about ideas from more than one perspective, mode of representation, or source? Are the ideas represented by more than one type of media (e.g., video and text) and do the ideas reflect what readers encounter in their academic and everyday lives?
- Digital scaffolds: Are scaffolds presented in a consistent manner across the activities that students complete? Is digital tool use or response format supportive of student comprehension?

Because all texts in 2025 NAEP Reading Assessment appear on the screen (rather than on paper), readers are presented with an array of textual distinctions that can be used to characterize variations in text structure and text features. These textual variations are likely to influence how readers make meaning as they engage with NAEP reading activities. For that reason, texts will also be tagged according to a variety of other textual structures and textual features (visual, multimodal, structural elements) that support, enhance, or potentially complicate meaning-making as readers move between single static (onscreen), single dynamic (nonlinear and/or multimodal), and complex textual environments.

As described in Chapter 3, *static texts* are those that contain no movement, *dynamic texts* contain movement, and *complex textual environments* are composed of two or more static or dynamic texts and the ideas constructed by readers as they move from one text to another. Text structures and features (see Exhibit 4.9) define the organization and elements within the text. *Text structures* refer to the linear and non-linear ways that ideas are connected to one another. *Text features* refer to the visual and multimedia elements used to represent these ideas. Both text structures and text features can influence the reader's ability to understand text. Consequently, tagging these digital elements creates important opportunities to investigate how readers use various structures and features to develop understanding and solve problems in the 2025 NAEP Reading Assessment.

Exhibit 4.9 presents a comparison of text structures and features that readers may encounter as they engage with single static texts, single dynamic texts, and complex textual environments. The exhibit itself, wherein static and dynamic texts are situated inside the boundaries of the figure, is designed to illustrate the idea that a complex textual environment contains one or more static and dynamic texts with varied distinctions in structures and features.

Exhibit 4.9. Text Structures and Features Within and Across Single Static and Dynamic Texts and Complex Textual Environments

SINGLE STATIC TEXT	SINGLE DYNAMIC TEXT
Textual structures are comparable to those in a printed format for texts designed to inform, entertain and/or persuade. **Textual features** may include visual media elements in a single text comparable to those in a printed format that convey meaning through primarily static words, numbers, and/or visual graphics, such as those in a still photograph, diagram, or table.	**Textual structures** include one or more **nonlinear** elements (e.g., hypermedia or hyperlinks) for readers to quickly move from one location or mode to another, but still *within the same text* (e.g., a navigational menu at the top of a document). **Textual features** include one or more **multimodal** elements (words, moving images, animations, color, music and sound) embedded into a single text or other media element

COMPLEX TEXTUAL ENVIRONMENT
Text structures may include one or more static or dynamic texts, with a strong likelihood of nonlinear elements both within a text (e.g., hypermedia or hyperlinks) that may lead to another text (e.g., another webpage within the same website or another webpage on a different website). **Text features** may include linked texts may contain either related or conflicting textual ideas. Multimodal elements (words, moving images, animations, color, music and sound) may appear in any or all texts.

Note: Ideas within each cell are likely to change and expand as new kinds of texts and technologies continue to emerge.

Illustrative examples of texts and other media belonging to each category are depicted in Exhibit 4.10. Importantly, these examples are likely to expand and change as new kinds of texts and technologies emerge and combine into new text forms. As technology becomes an increasingly integral part of reading and learning, designers are encouraged to sample from dynamic textual environments that include webpages, websites, search engines, digital media, simulations, augmented reality texts, digital creation tools, and other media forms likely to emerge to support the comprehension and production of digital text. Further, as new technologies and digital features continue to expand the ways readers engage with academic and everyday reading materials, the definition of *authenticity* may need to expand as well. A series of studies are now needed to better understand how variations in readability and text complexity relate to and/or influence meaning-making in the 2025 NAEP Reading Framework assessment space.

Exhibit 4.10. Illustrative Examples of Texts and Other Media Across Single Static and Dynamic Texts and Complex Textual Environments

SINGLE STATIC TEXT	SINGLE DYNAMIC TEXT
Examples of single static genres and forms of continuous prose, non-continuous prose, and everyday reading materials from which designers might sample as readers read to engage in literature, science, or social studies and history are found in Exhibit 4.9	*Nonlinear text* Single text with hyperlinks that only connect to ideas within the same document; may also contain one or more dynamic media elements *Dynamic media* ● Dynamic image ● Video ● Podcast ● Digital poster ● Infographic ● Interactive timeline ● Interactive chart or graph ● Data visualization ● Blog ● Simulation

COMPLEX TEXTUAL ENVIRONMENT	
● Augmented reality text	● Google document or Google folder
● Blog	● Role play simulation
● Database	● Search engine
● Digital creation/composition tool	● Social media (e.g., Facebook, Instagram, Twitter)
● Dynamic simulation	● Threaded discussion
● Email	● Webpage or website
● Interactive model	

Designing Tasks. Recall that tasks are structures that (a) reside within an activity, and (b) require the presence of at least one text, at least two comprehension items, and some unspecified number of scaffolds and probes. Thus the first criterion for task design is that it reflects both the purpose and the disciplinary context within which the entire activity is situated. Moreover the task as a whole, and more specifically the comprehension test items students complete, must be consistent with (a) the sociocultural practices of the disciplinary Context (detailed in the text section of Chapter 3) and (b) the Comprehension Targets (also in Chapter 3).

In addition, designers should also consider the *navigational complexity of text as it interacts with the reading task and the specific demands of the comprehension items attached to the text(s) within tasks* (see Coiro, 2020). Comprehension items may, for example, vary in difficulty according to the nature of associated comprehension processes (e.g., locating a topically relevant idea is likely easier than inferring the tone of a particular passage or analyzing the impact of an author's word choice on a particular audience).

Further, comprehension items may vary in difficulty due to the nature of inferences readers are asked (or required) to make across the comprehension processes; that is, the *type* of inference (a local, straightforward inference vs. a global inference across ideas in a text) combined with the *number* (one or multiple) and the *distance* of these inferences (within one text, across two texts, or beyond the text) introduce variations in task and item demands that impact the difficulty of a particular comprehension item on the reading assessment. Thus, designers will follow item specification guidelines to estimate levels of navigational complexity across an activity block as shaped by the number, levels, and types of inferences as well as the nature of texts, tasks, items, and response types included. In turn, estimated difficulty levels can be used to inform special studies that further explore how reader attributes interact with various task demands to influence comprehension performance.

Designing Items. After selected texts and tasks are situated in one of the three Contexts (Reading to Engage in Literature, Science, or Social Studies) and the chosen Purpose (Reading to Develop Understanding or to Reading to Solve a Problem), the next design task is to develop appropriate comprehension items based on the Comprehension Targets described in Chapter 3.

Design Principles. As with the selection of texts, item development is guided by a set of design principles in order to guarantee that readers are asked to respond to important aspects of the text and to utilize a range of processes that result in successful comprehension. These design principles include:

- **Importance**. Items should focus on central textual and intertextual concepts or themes or, on occasion, more specific information related to these themes and concepts. For example, a fact that provides evidence to support a claim or a detail that supports a main idea may be queried.
- **Balance**. The Comprehension Targets should be proportionally distributed across dimensions of the activity block
  - across grade levels.
  - across the disciplinary Contexts of Literature, Science, and Social Studies.
  - across activity purposes.

  While the percentage of Comprehension Targets may vary across these dimensions, items representing all Comprehension Targets should be represented at all levels of these dimensions.

- **Varied knowledge sources**. Items should invoke a variety of prior knowledge and textual knowledge sources. Across items and in accordance with the Comprehension Targets, readers should be called upon to **employ different kinds of prior knowledge** (e.g., topical knowledge, vocabulary knowledge, discursive knowledge) and to **draw information from different sources in the texts** (including information at various levels of abstraction [e.g. directly stated in prose, embedded in a visual representation, or implied through symbolism] and across different locations in the text).
- **Alignment with an array of skills of navigation and inference.** Across items and in accordance with the focus of the Comprehension Targets, items should call upon readers to locate information in different textual environments (e.g., static and dynamic) and to make different kinds of inferences, from local bridging inferences to more complex intertextual inferences and extratextual applications of knowledge to a new situation (Use and Apply). (See Chapter 3 for examples.)
- **Clarity and transparency.** Items should be accessible and transparent. Readers should know what they are being asked to do, the processes that will guide them toward an appropriate response, and how their responses will be evaluated.

Planning the Distribution and Characteristics of Comprehension Items. The four Comprehension Targets do not represent a hierarchy of skills. The difficulty of any particular item, regardless of which comprehension target it is designed to elicit, should be shaped by the content (the ideas themselves), the language and structure of the text (the language and relations among ideas) and the cognitive demands of the Comprehension Target. As a consequence, there can be relatively difficult items representing Locate and Recall Comprehension Targets and relatively easy items representing either Integrate and Interpret or Analyze and Evaluate targets. Below are example item starters from Chapter 3 that might prompt thoughts about high-quality items for each comprehension target—always keeping in mind that the single most important standard that 2025 NAEP Reading Assessment must meet is asking questions about matters of substance in the texts. Some prompts to trigger good item development for each comprehension target include the following:

Locate and Recall

- Locate specific pieces of information by skimming, navigating, or searching one or more texts
- Find information in dynamic and multimodal texts
- Recall and record information
- Identify the nature of the relationship between pieces of information explicitly stated in text
- Discern the meaning of an unfamiliar word based on context

Integrate and Interpret

- Create summaries
- Form explanations and generalizations
- Draw conclusions
- Offer interpretations
- Formulate questions
- Make predictions
- Provide evidence that supports particular interpretations or conclusions

- Generate alternatives (e.g., alternative interpretations of an implied message or moral or alternative actions that a character might have taken)

Analyze and Evaluate

- Revise understanding in light of a new piece of information
- Reconcile inconsistencies across a text or texts
- Judge the plausibility of an event or idea within a text
- Evaluate how the author uses language, including vocabulary, and features of text to achieve a particular purpose
- Determine significance
- Argue for or against particular interpretations
- Evaluate the relevance and strength of evidence to support claims made
- Identify which perspectives are missing and which are emphasized..

Use and Apply

- Produce a novel solution for a problem described in readings
- Apply ideas acquired though reading to a new problem or context
- Take on a different perspective (e.g., a scientists vs. an historian; a different character in a story)
- Rewrite or re-represent information in a purpose or audience-driven way
- Rewrite a description of a character, using different language and words, to convey a different interpretation of their character or actions.
- Use an understanding of legal and ethical principles (e.g., intellectual property or intellectual freedom) to develop an opinion on a social debate—or to plan a form of social action

Exhibit 4.11 presents guidelines for distributing items mapped to Comprehension Targets across grade level and activity blocks. These flexible distributions allow for the possibility of varying the number of items for each target depending on block type; that is, activities involving Reading to Develop Understanding may have more items designed to assess Integrate and Interpret or Analyze and Evaluate ideas while those involving how readers apply understanding to solve a problem may have more items designed to assess Use and Apply and fewer items designed to assess Locate and Recall processes. And the distribution targets should never outweigh the other principles in the bulleted list. In other words, for a given text, it is better to fall one item short in the number of items for a target than it is to include one that fails the importance or the clarity standard just for the sake of meeting the distribution goal.

Exhibit 4.11. Distribution of Cognitive Comprehension Targets Across Grade Level and Activity Blocks

Grade Level and Comprehension Targets	Reading to Develop Understanding	Reading to Solve a Problem
Grade 4		
Locate and Recall	30 - 40%	10 - 20%
Integrate and Interpret	30 - 40%	30 - 45%
Analyze and Evaluate	10 - 20%	10 - 20%
Use and Apply	10 - 20%	20 - 30%
Grade 8		
Locate and Recall	10 - 20%	10 - 20%
Integrate and Interpret	30 - 40%	20 - 30%
Analyze and Evaluate	30 - 40%	20 - 30%
Use and Apply	10 - 20%	20 - 30%
Grade 12		
Locate and Recall	10 - 20%	10 - 20%
Integrate and Interpret	30 - 40%	20 - 30%
Analyze and Evaluate	30 - 40%	20 - 30%
Use and Apply	10 - 20%	20 - 30%

Selecting Language Structures and Vocabulary to be Tested. The 2025 NAEP Reading Framework Reading Framework pays particular attention to language knowledge as an important contributor to reading performance. Moving beyond measuring vocabulary knowledge exclusively, **language structures and vocabulary** in the 2025 NAEP Reading Framework refers to the application of the reader's understanding of the meanings of language structures and words to text comprehension. This application of language structure and vocabulary knowledge to text comprehension encompasses three dimensions:

- *Discourse dimension* refers to textual relations across words and phrases.
- *Semantic dimension* refers to the meaning of individual words or expressions.
- *Morphological dimension* refers to parts of words.

These dimensions refer to relations across words and phrases, individual word meanings, and word parts that are important to understand the text.

As depicted in the left column of Exhibit 4.12, the 2025 NAEP Reading Assessment will assess language that is highly useful for grade-level reading across the curriculum, including words' meanings but also language structures beyond the level of words. To that end, items will exclude infrequent words or discipline-specific language; idiomatic expressions that characterize readers' participation in cultural and idiosyncratic discourse practices; syntactically intricate structures that obscure rather than clarify meanings; and language known to be part of students' everyday speaking vocabulary at a specific grade level. Thus, Language Structure and Vocabulary items will evaluate readers' application of their knowledge of useful grade-level words and language structures to their understanding of a text or a set of texts.

Exhibit 4.12. Inclusion and Exclusion Criteria for Connected Language and Vocabulary

Language Structures and Vocabulary to be Tested	● Words and language structures that appear across numerous texts, either across literary texts (e.g., *despise, benevolent*) or across social studies and natural sciences texts (e.g., *resolution, commit*) ● Words or phrases necessary for understanding at least a local part of the context and linked to central ideas such that unfamiliarity may disrupt comprehension ● Words and language structures found in grade-level texts ● Words that label generally familiar and broadly understood concepts, even though the words themselves may not be familiar to younger learners (e.g., *timid*). ● Words that include word parts (roots and affixes) useful to acquire and figure out the meaning of unfamiliar words (e.g, *disregard, counterargument*). ● Language that expresses logical relations between ideas (e.g., phrases that include connecting words such as *although, in contrast*) ● Expressions that refer to characters, events, or ideas previously introduced in the passage (e.g., *those alliances, this phenomenon*)
Language Structures and Vocabulary to be Excluded from Testing	● Narrowly defined and not widely useful words, such as those related to specific content domains (e.g., *photosynthesis, fiduciary*) or words with limited application (e.g., *deserter, hamlet*). ● Idiomatic expressions (e.g., in a nutshell, up in the air) ● Unnecessarily syntactically complex and long constructions (e.g., while ____, then ____). ● Words and language structures that are already likely to be part of students' everyday speaking vocabulary at grade level.

Due to the intricate relation between language knowledge and text comprehension, instead of including items that exclusively measure language knowledge, items should be doubly coded for Comprehension Target (e.g., Locate and Recall; or Integrate & Interpret) and Language Structures and Vocabulary. A total of 30 percent of items in any activity block will assess passage-relevant Language Structures and Vocabulary knowledge while concurrently measuring a specific comprehension process. This proportion of items dedicated to Language Structures and Vocabulary will allow for the examination of the contribution of students' language knowledge to reading performance. For an item to be doubly coded as Language Structures and Vocabulary, the reader's knowledge of the meaning of the words or expressions being tested in the item should be both (a) necessary and (b) sufficient to answer the item correctly. For example:

- **A sample item that measures *Locate and Recall* & Language Structures and Vocabulary:** *Which word in the first paragraph helps you understand the main reasons for the president's decision?* [answer is found in its exact form in the text, e.g., *due to...]*
- ***A* sample item that measures *Integrate and Interpret* & Language Structures and Vocabulary:** *What is the protagonist suggesting about her feelings when she uses words and phrases such as "uncertain" and "at a loss"?* [answer integrates information from the text and paraphrases the meaning of these expressions, e.g., the protagonist does not know for sure what to do or think]

In these sample items, the reader's understanding of the meanings of the underlined language expressions in the passage (e.g., *due to, uncertain, at a loss)* is both necessary (i.e., the reader will not be able to answer the item correctly without this understanding), and sufficient (i.e., all the reader needs is to understand the meaning of these expressions in the context of the passage) to answer the item correctly. Given these two conditions of necessity and sufficiency, most Language Structures and Vocabulary items are likely to be *Locate and Recall* or *Integrate and Interpret* items, in addition perhaps to some language-focused *Analyze and Evaluate* items. Some items that measure more complex processes, while certainly supported by the reader's language knowledge, involve many additional higher-order cognitive processes. Thus, correctly answering items that measure higher-order Comprehension Targets typically requires an integration of skills that go beyond the application of language knowledge as the main contributing factor.

One additional suggestion is to conduct special studies to examine the language produced by students in their open-ended responses to items across all Comprehension Targets. These responses could be automatically analyzed or be human-coded for language, which would be insightful. Relatedly, scoring rubrics and training for scorers need to be language-conscious in careful ways so that students are not erroneously penalized for language features irrelevant to the comprehension processes being assessed.

Cycle 3. Strategic Selection of Scaffolds, Probes, and Item Formats

Once a first vision of a NAEP reading activity is created, steps are taken to consider what is expected of readers (assumed knowledge, skills, and dispositions) and to anticipate possible locations where scaffolds, probes, and item formats might be used to support readers. The explosion of digital media resources has introduced myriad opportunities to integrate scaffolds in service of allowing diverse learners to show more optimal levels of performance on comprehension tests (Chen & Chen, 2014; Dalton & Practor, 2008; Sparks & Deane, 2014; Sabatini & O'Reilly, 2014). Informed by these research findings, Cycle 3 of the design process involves sampling from scaffolds intended to support readers' knowledge, motivation, and/or metacognitive strategy use (as described in Chapter 3) to maximize the possibility that they will perform to their potential on the assessment.

In this phase, designers consider the complexity of texts (with respect to content, genre, textual features, and organizational structures) along with task and item demands for the reading activity envisioned in Cycle 2. Then, designers select specific scaffolds related to knowledge, motivations, and/or metacognitive strategy use in line with targeted processes for readers of different grade levels at different points in the text or reading task. Consideration is also given to probes, process data, and item formats, as described in the following section.

Selecting Scaffolds to Support Comprehension. While acknowledging that all scaffolds should be purposeful, intuitive, and beneficial for comprehension as situated in a particular text, context, and activity block, designers are encouraged to follow broad guidelines about where and why scaffolds may be used in 2025 NAEP Reading Assessment activities. As explained in Chapter 3, *knowledge scaffolds* may be embedded into a reading activity to provide readers access to prior knowledge assumed by a text's author with respect to content, vocabulary, and/or discursive norms of a particular context or community (e.g., a short video or pop-up support for word-level or knowledge scaffolds, criteria guides for analyzing text from a particular lens or perspective). *Metacognitive/strategy scaffolds* such as task-specific purpose statements or a sequential set of directions may be used to communicate expectations for how and why readers should engage with texts while also helping them to plan and monitor their work across multiple texts and tasks. For example, at the outset of activity structures, designers outline expectations for how readers might engage with the selected texts and tasks; these directions serve as a strategy scaffold to aid readers in breaking down a 30-minute reading activity into two or three manageable parts aligned with task-specific purposes. Also, for Reading to Solve a Problem activity blocks, particularly those structured as Scenario-Based Tasks, a major scaffold is provided; namely the computer interface controls the presentation of tasks, texts, and even specific items, thus eliminating the need for students to determine when to read or answer comprehension items.

Strategy scaffolds may also take the form of digitally enhanced, easy-to-manage tools with which readers retain and retrieve information attended to at various points in the activity, or they may offer models of literacy practices or approaches to solve a problem before readers must do something on their own. Sometimes, knowledge and strategy scaffolds may serve overlapping roles as *motivational/social scaffolds* designed to actively engage readers' interest at the beginning of one or more tasks. That is, a short video about an abstract scientific phenomenon not only serves to facilitate background knowledge, but may also facilitate readers' interest or sustain their willingness to engage with abstract ideas and persist in the assessment activity. When Reading to Engage in Literature or Social Studies, a brief video interview with the author of the text may be included to remind readers of the personal and socially constructed nature of ideas with which they are asked to engage; in turn, this video may also serve to motivate readers to engage in the task.

In this third cycle, designers also decide on appropriate locations in which to insert scaffolds. Because some 2025 NAEP Reading Assessment activities involve metacognitive complexities in response to handling multiple tasks and texts, readers may be asked to check and reflect on their reading progress in the activity and allocate their attention accordingly. Intuitively designed transitions between each task, such as avatar guides, visual flow charts, or simple written statements, may be used to help readers overview the activity sequence and structure.

As the 2025 NAEP Reading Assessment evolves and develops new activity blocks, assessment developers should be encouraged to include scaffolds that are proven (either from previous versions of NAEP or in other well-documented research efforts) and to undertake special studies to evaluate the efficacy and utility of other scaffolds that show great promise for use in an asset-oriented assessment such as the 2025 NAEP Reading Assessment. A list of scaffolds, both those ready for use and those that show promise, appear in Exhibit 4.13. A major question for block designers is how to decide when to employ and when to forego the deployment of a specific scaffold as the potential for added support is weighed against the potential for increased cognitive burden on the reader.

Exhibit 4.13. A Scaffolding Plan for the 2025 NAEP Reading Framework

Type of Scaffold	Ready for Operational NAEP	Ready for Special Study
Knowledge Scaffolds	● Pre-reading vignettes (text, image, or video) ● Resetting (providing the answer to item X before going on to item X+1) ● Providing word meanings for key vocabulary excluded from testing	● Linguistic or discursive suggestions for previewing, defining, or illustrating ideas in a text or a text feature ● Anticipation guide to capture evidence of knowledge about key statements and insights gleaned after reading
Metacognitive and Strategy Scaffolds	● Visual organizers ● Look back buttons for vocabulary and connected language ● Graphic organizers for recording key information	● A planning tool ● Criteria to use in analyzing or evaluating a text or a text feature ● Criteria for selecting a response ● Criteria or word/phrase bank to aid in constructing a response ● Mentor texts (a text to illustrate what a good response looks like) ● Mentor strategies (reminder of expected routines)

		• Avatars to assist with organizing, analyzing, or communicating information
Motivational or Social Scaffolds	• Pre-reading vignettes (text, image, or video) designed to facilitate interest or real-world connections • Resetting (providing the answer to item X before going on to item X+1) designed to sustain willingness to persist • Avatars designed to assist with organizing information	• Avatars to assist with organizing, analyzing, or communicating information • Mentor texts or mentor strategies designed to promote self-efficacy and perceptions of agency

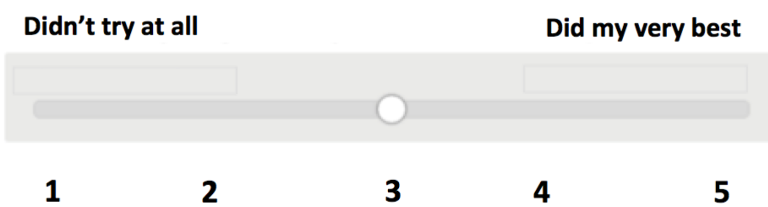
Selecting Probes and Process Data to Help Account for Reader Attributes. In Cycle 3, designers also anticipate the role of reader interest and perceptions of self and task in the reading activity. They decide which reader attributes should be captured inside the assessment using block-specific probes and/or process data or if it makes more sense to capture these attributes with survey items that ask students about classroom and community contexts outside the reading activity. Inside the activity, block-specific measures (or probes) in the form of open-ended prompts or slider scales can be used to gather information about readers' interest, familiarity, or effort, for example. Probes may be embedded into only one location in the activity (e.g., at the beginning of a block), or they may be embedded into multiple locations to gather information about how readers' interest and/or perceptions change over the course of an activity (e.g., pre-post scales or slider items).

Exhibit 4.14. Example of Slider Items Within

Thinking of a scale from 1-5, when you read the texts about pyramids and answered the questions,

- How hard did you try?
- How interesting were the texts?
- How well do you think you did?
- How much did you already know about the topic?
- How much did you learn?
- How often do you get assignments like this in your language arts classes?
- How much do you know about these topics (caravans, The Sphinx, pyramids)??

Each item is presented on a screen as a slider from low to high. So for the first item, the slider would read



Because 2025 NAEP Reading Assessment activities are situated in a fully digital environment, process data involving reader actions (e.g., number of mouse clicks, pathways through a task or hypertext, transcribed voice responses, length of time spent engaged with reading material or responding to an item) can be easily collected in digital log files stored in a database. These types of data hold potential power to measure levels of engagement in purpose-driven reading activities (e.g., capturing frequency, density, and intensity of engagement or identifying and comparing novice to expert level of practice). Process data from log files can be aggregated and interpreted to characterize how reader attributes or other explanatory variables explain comprehension performance at one or more locations in the NAEP assessment space. Examples of elements designers might sample from to account for reader variations are provided in Exhibit 4.15.

Exhibit 4.15. Examples of Probes and Process Data to Account for Variations in Knowledge, Metacognitive Strategy Use, and Personal Attributes

	Knowledge	Metacognition/ Strategy Use	Reader Attributes
Probes	• Multiple choice (2-3 items) • Rating scale • Slider bar • Knowledge prompts with open-ended response boxes	• Multiple choice (2-3 items) • Rating scale • Slider bar • Reflection or process prompts with open-ended response boxes	• Multiple choice (2-3 items) • Rating scale • Slider bar • Reflection or reaction prompts with open-ended response boxes
Process Data	• Clickstream data (e.g., page hits, link selections, reading sequences) • Activity logs (e.g., amount of time spent for text reading, item response, and other assessment features such as instructions, transitions, and avatar guide) • Search logs (e.g., search term choices) • Transcribed voice responses		

Overall, the strategic sampling and tagging of scaffolds and probes within activities enables the 2025 NAEP Reading Assessment to engage readers in asset-driven activities while also generating information to better account for and explain the reading performance of 4th-, 8th-, and 12th-grade students. Throughout the three cycles, designers organically work and rework options for scaffolds and probes to populate the grade-appropriate assessment space while recognizing that readers need to accomplish the expected outcomes in 30 minutes or less. Ultimately, decisions about what to include should be specific to the assessment block as designers consider what is needed to fulfill the goal of making the assessment asset-driven. As knowledge about the use of scaffolds and probes becomes more robust and precise, more of these features should be operationalized in the NAEP Reading Assessment in the years ahead.

Selecting Item Formats. Central to the development of 2025 NAEP Reading Assessment is the careful selection of ways that students respond to items. Since 1992, items on the NAEP Reading Assessment were limited to two formats: multiple choice and constructed response (write the response with a pen or pencil). 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. The 2025 NAEP Reading Framework retains selected and constructed response options that include a variety of formats.

Selected Response Options.

- **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 stem of the item.
- **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 ideas 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.
- **Select in passage**: Students select one or more ideas in the passage and drag them into the target fields.

As the 2025 NAEP Math Framework proposed, one forward-thinking area to consider is the use of discourse and collaboration responses as part of the 2025 NAEP Reading Assessment. These types of items align most directly to social and collaborative reading practices briefly discussed in Chapter 3. What might these look like? Current examples piloted with pairs of high school students (Coiro et. al., 2019) enable readers to ask text-based questions of simulated characters (avatars) and choose (e.g., through multiple-choice, limited option selections) from given conversational responses to move the reading activity forward. Such a selected-response choice then provides some information about the level of collaborative reasoning the learner exhibits. This leads to a selected response type that expands the selected response types listed above to include:

- **Discourse/collaboration limited option responses** – Students respond by selecting from two or more choices of conversational responses as part of a discourse-based or collaborative task.

Constructed Response Options.

- **Short constructed response** – Students respond by entering a short text in a response box that consists of a phrase or a sentence or two
- **Extended constructed response** – Students respond by entering an extended text in a response box that consists of multiple lines (a paragraph or two).
- **Hybrid constructed response.** Students respond by selecting two or more choices that meet the condition stated in the stem of the item. Then they write a short explanation about their choices.
- **Fill in the blank:** Students respond by entering a short word or phrase in a response box.

Flexible distributions of item responses type across grade level are presented in Exhibit 4.16.

Exhibit 4.16. Flexible Distributions of Item Response Types Across Grade Level

	Selected Response Items	Short Constructed Response Items	Extended Constructed Response Items
Grade 4	40-50%	40-45%	10-15%
Grade 8	40-50%	40-45%	10-15%
Grade 12	40-50%	40-45%	10-15%

With respect to Language Structures and Vocabulary, designers should also adhere to recommendations (see Exhibit 4.17) about which item formats and response choices are best suited to test Language Structures and Vocabulary items and which formats should be avoided. These guidelines call for innovative technologically enhanced items that can assess language knowledge through multiple item formats (e.g., drag-and-drop; text highlighting) and multiple modalities (e.g., enabling graphic-based or audio-recorded responses). Second, these guidelines include offering supports (e.g., a definition, an illustration) for discipline-specific language or idiomatic expressions that might be central to comprehending a passage but are not part of the Language Structures and Vocabulary construct or comprehension item. Finally, these guidelines call for specific attention to the use of accessible language throughout the 2025 NAEP Reading Framework, including assessment instructions, questions, and student surveys. Together, these guidelines for the selection of *what to test* and considerations for *how to test* intentionally and emphatically bring to light issues of equity in assessment. Given the intercultural negotiation of meaning through language involved in reading and reading comprehension, these guidelines seek to apply both the sociocultural lens and the equity lens to what language to test and how to use language to test comprehension.

Exhibit 4.17. Inclusion and Exclusion Criteria for Item Formats and Response Choices for Language Structures and Vocabulary

Item formats to be included	Item response choices to be included	Item response choices to be excluded
Innovative item formats that test language comprehension through multiple modalities (drag-and-drop items, highlighting of text sections, graphically-enhanced or audio-enabled responses)	*Response choices that:* ● Are expressed in highly accessible language ● Present a different common meaning of the target vocabulary word or phrase, which must be ignored in favor of the meaning in context. ● May present correct information or content from the text that is not what is meant by the target word. ● May be an alternative interpretation of the context in which the target word occurs.	*Response choices that:* ● May be the meaning of another word that looks or sounds similar to the target word. ● May present a common but inaccurate association with the target word.

Dynamic Response Items. NAEP is currently exploring the use of dynamic response items to assess comprehension such as graphic organizers and drop down menus, to name a few; it should continue this trend in the years ahead by further exploring the use of other interactive or dynamic response formats made possible with emerging digital tools. Useful frameworks (Scalise & Gifford, 2006) and guidelines (Measured Progress/ETS Collaborative, 2012) introduce a wide variety of innovative item types that should be considered by NAEP in implementing digitally-based facets of the 2025 NAEP Reading Assessment. For example, dynamic item formats introduce opportunities to assess how readers:

- Search and locate ideas (e.g., dynamic search engines);
- Select and identify ideas (e.g., multiple choice items with new media distractors);
- Reorder or rearrange ideas (e.g., ranking, categorizing, and sequencing items);
- Substitute or correct ideas (e.g, multiple drop down menus offering word choices embedded within lines; limited graphical elements that are adjusted or corrected to accurately represent ideas in the passage);
- Categorize or classify ideas (e.g., tiling, select and order); or
- Construct relationships among ideas (e.g., dynamic concept maps, multimodal representations).

When selecting the format of any particular item, designers should be mindful of the cognitive and logistical demands of varied formats and how these may interact with reader familiarity and the time constraints of each activity. While there is great potential in this arena, more research is needed to better understand the complex interplay among text, activity, reader attributes, and features of digitally enhanced items, scaffolds, and probes in digital environments. With the innovative digital assessment spaces proposed by this framework, NAEP will now be in a position to conduct a series of special studies with the goals of making more informed decisions for how and when to embed specific digital tools, resources, scaffolds and item formats into different activity arrangements to engage students in rich and asset-oriented assessment activities.

Addressing Universal Design for Learning in Assessment Design

To ensure that readers experience opportunities to maximize performance on the assessment, the 2025 NAEP Reading Framework is aligned with the Universal Design for Learning framework (Rose & Meyer, 2002). As noted in Chapter 3, Universal Design principles are reflected in NAEP's attention to giving readers multiple, varied experiences to engage with reading and demonstrate what they know. An important distinction between principles of universal design and accommodations is that the supports emanating from universal design are provided to all students, whereas accommodations are targeted supports available only to particular populations.

Anatomy of Two Different Activity Blocks

This last section presents two examples of activity blocks. The first outlines the assessment components in a Reading to Develop Understanding activity block in which fourth graders read to Engage in Literature. In this activity, fourth-grade readers engage with static excerpts and illustrations from a story and a short video as they read to develop a richer understanding of the main characters, key events, and author's craft and apply their insights to predict events beyond the story. The second example illustrates what eighth graders might encounter in a Reading to Solve a Problem activity block as they read to Engage in Social Studies. In this activity block, students are guided to engage in more complex kinds of reading tasks that might include two to four more dynamic texts and involve greater intertextuality and integration across items, all of which contribute to a generative culminating event, such as developing a solution to a problem. While both activity blocks include tasks, texts, items, and scaffolds, differences in what readers experience illustrate just a sampling of the range of possible design features from which designers might sample across the two purpose-driven activities in the creation of any single activity block.

Grade 4 Reading to Develop Understanding in Literature Activity Block

Context. This activity block is designed to assess how 4th-grade readers develop understanding within a single text and then apply that understanding to a simple culminating task, which involves answering a question in writing. As the culminating task, this particular example asks students to form an interpretation about the kind of person the main character, Hana, is and to then use that characterization to consider what might happen next in the story. While many of the items give students opportunities to develop their thinking across the story, the items and texts (video and story) are relatively independent of one another. The activity also includes opportunities to develop understanding around other aspects of the story that may, or may not, contribute to that characterization. At the beginning of the assessment, readers are provided motivational and strategy scaffolds. Throughout the activity, readers are asked to activate and employ their personal, cultural, and literary knowledge and resources by drawing on textual evidence to make thoughtful interpretations of the text.

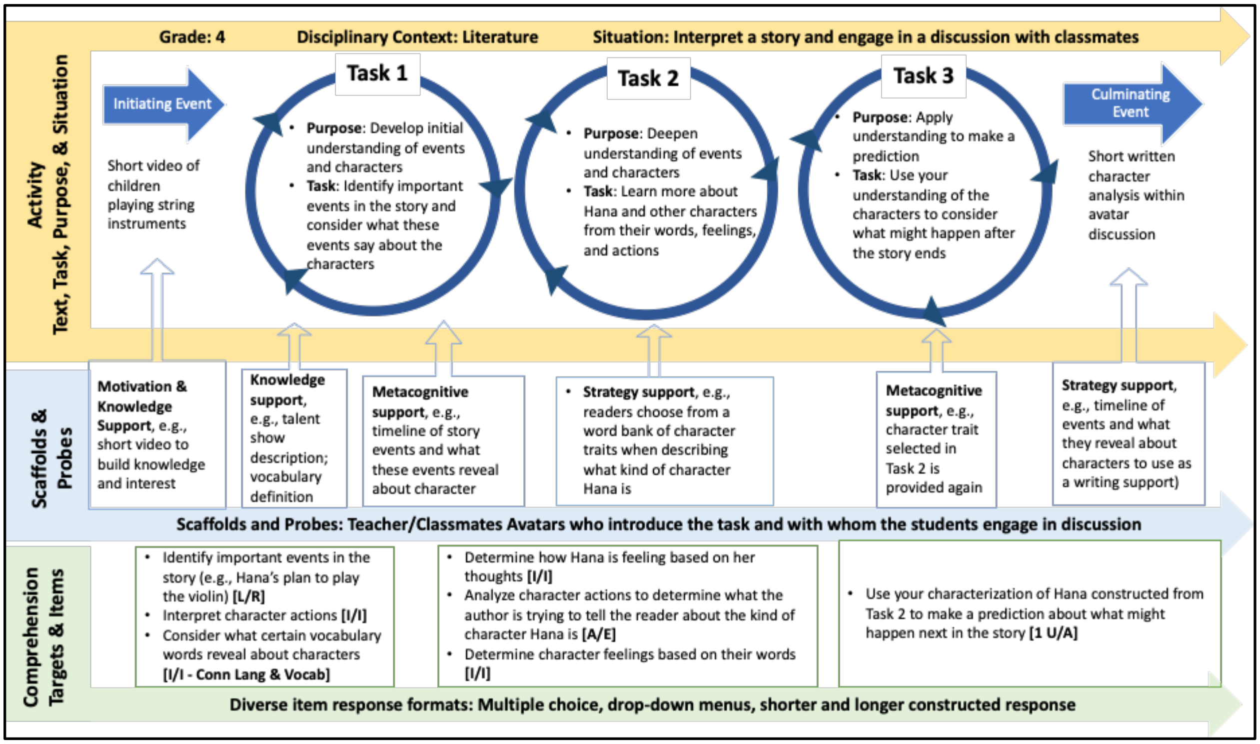
As illustrated in the concept sketch in Exhibit 4.18, the situation for this activity involves a group of fourth-grade students who are reading the text, *Hana Hashimoto, Sixth Violin*, by Chieri Uegaki. In this book, a young girl named Hana signs up to play the violin in her school's talent show after having had only three lessons. At the outset of the assessment, test-takers are first asked to rate their knowledge of violins (with a slider scale probe) and then they are invited to view a short video of young children playing in an orchestra - https://www.wonderopolis.org/wonder/why-do-orchestras-need-so-many-people
This initiating event introduces readers to the sounds and emotions one might experience when playing in an orchestra, while providing background knowledge to students who may be less familiar with stringed instruments such as the violin.

Tasks. After these efforts to socially situate and engage readers in the activity, readers are invited to participate in a book discussion group about the story *Hana Hashimoto, Sixth Violin* with three other 4th graders (represented by avatars in the assessment). A teacher avatar joins them to explain the goal of the discussion, which focuses on understanding how Hana grows and changes over the course of the story as a result of events involving her classmates and her family. To prepare for the book discussion, students are told they will read parts of the story and respond to items situated in three tasks to: 1) identify important events in the story and consider what these events say about the characters; 2) learn more about Hana and other important characters from their words, feelings, and actions in the story; and 3) apply their understanding of the characters to consider what might happen after the story ends.

Texts and Items. After viewing the video and learning about the three tasks in this literary reading activity, readers engage with several passages from the book that contain important information about Hana and other minor characters. Through these passages, readers learn that Hana's desire to take lessons was inspired by a recent visit to Japan to see her Ojiichan, or grandfather, who plays the violin. They also learn that despite much teasing and doubting from her brothers, Hana practices and practices for the talent show, inviting everyone she can to be her audience. Sample questions at this point may, for example, include multiple choice items to assess readers' ability to Locate and Recall important details (e.g., What does Hana plan to do in the talent show?) as well as short constructed-response items that ask readers to interpret and integrate character traits into their understanding of the story (e.g., At this point in the story, do you think Hana's brothers supported her decision to play in the talent show? Please explain your thinking.).

Toward the end of the story, readers learn that when Hana is on stage, she first becomes nervous and doubts herself, but then imagines her Ojiichan telling her to do her best. Hana decides to play what she knows — the sound of a crow, her neighbor's cat, and rain on a paper umbrella. Her family loves her performance so much that later that evening, they ask her to play them some more notes around the dinner table. The story ends with Hana recalling her Ojiichan's numerous songs and imagining what she might play in next year's talent show. As readers engage with excerpts from this part of the story, they are invited to join the discussion group with three avatar fourth graders and contribute their ideas. After listening to one of the student avatars orally describe how Hana reacted to her brothers' behavior earlier in the story, readers might be asked: "Based on what you learned about Hana as a person, what do you think might happen if she plays her violin again in next year's talent show? What makes you say this? Support your answer with a detail from the text." This longer constructed response item is designed to assess readers' ability to Use and Apply their understanding to a new situation beyond the story itself.

Exhibit 4.18. Concept Sketch of a Reading to Developing Understanding Activity Block: An Activity Block Example Hana Hashimoto, Sixth Violin (by Chieri Uegaki).



Scaffolds. The block design (as shown in Exhibit 4.18) includes a range of scaffolds as students read texts, respond to items, and reflect on their performance. The activity block is initiated with a short video of young children playing in an orchestra to pique students' interest and provide background knowledge with which to engage in the reading and the task. Then, teacher and student avatars are used to introduce the activity at the outset of the scenario. Within the task, vocabulary scaffolds might include pop-up boxes describing what a talent show is or providing a hint about the meaning of certain vocabulary words, such as "fled." Students might also be given a word bank of character traits from which to select when asked to describe the kind of person Hana is. Finally, students could also be given a timeline of story events (character actions, words, and feelings) on which to drag and drop their responses, or designers might insert a slider bar probe somewhere toward the end of the activity that invites readers to react to Hana's character (how much would they like to meet Hana or how brave they think she is) or rate their interest in playing the violin after engaging with this activity. As students respond to items about what these events say about the characters, this information could be added to the timeline so that students can use this informational graphic as a writing support when answering the last item. Additionally, item response types would vary from simple multiple choice to short answer constructed response to give readers different kinds of opportunities to respond.

Performance Evidence and Indicators. When interpreting reading achievement from performance on the 2025 NAEP Reading Assessment, multiple indicators can be used to situate and explain what students are able to do. As indicated earlier in this chapter, each activity block will be tagged according to a primary disciplinary context, grade level, and overarching purpose. The *Hana Hashimoto, Sixth Violin* activity, then, can demonstrate what Grade 4 students can do in a Reading to Engage in Literature activity block designed to measure their ability to develop their understanding within a single text and then apply that understanding to a simple culminating event (in this case, making a prediction, based on the story, about what will happen after the story ends). Items within each activity block will be tagged according to comprehension processes they are designed to measure; these tags enable NCES and its contractors to link performance to specific Comprehension Targets (e.g., Locate and Recall, interpret and integrate, Analyze and Evaluate, Use and Apply) as well as to one or more reader attributes (e.g., knowledge or interest) while noting any supports (knowledge, strategy, and/or motivational) embedded into the activity.

Using this kind of tagging system makes it possible to connect these indicators to 4th graders' performance on activity blocks that involve Reading to Engage in Literature. Ultimately, NAEP can compile a description of what 4th graders (or sub-groups of 4th graders) can do in each disciplinary context as they engage with texts and test items, while also being encouraged to draw from and use the knowledge, skills, and experiences they bring to that reading context.

Grade 8 Reading to Solve a Problem in Social Studies Activity Block

Context. This 8th grade activity block focuses on a group of students who are motivated to learn about a current civic project deeply rooted in the history of the city: The City of Pittsburgh has recently announced an ambitious plan for the construction of an overpass park that reconnects the Hill District and Downtown. Park designers at a landscape architecture firm have created a proposed park design. Students are asked to learn about this project and to consider the role of a key aspect of the proposed park design: the inclusion of a 13-year-old African-American girl named Keisha on illustrated signs throughout the park. The park designers have proposed including signs of Keisha in many locations in the park to provide details about the African-American community's history in the Hill.

At the beginning of the assessment activity, students learn that the city has recently unveiled the park plan to the public on its website, and the plan is now open for public comment. Some city residents have posted comments asking why the park is needed and why it should be constructed as an overpass. Other commenters are confused about Keisha's repeated presence throughout the park's plan; they wonder who Keisha is and what role she plays in the planners' vision. This situation inspires the question that guides the readers' inquiry: Why does Keisha matter to the city park project? Through a sequence of four tasks, readers are asked to explore the background history of the Hill District, demonstrate an understanding of the texts they encounter, and craft an historically informed presentation for the general public that clarifies and illustrates Keisha's role in the park (e.g., representing and celebrating the history of the Hill).

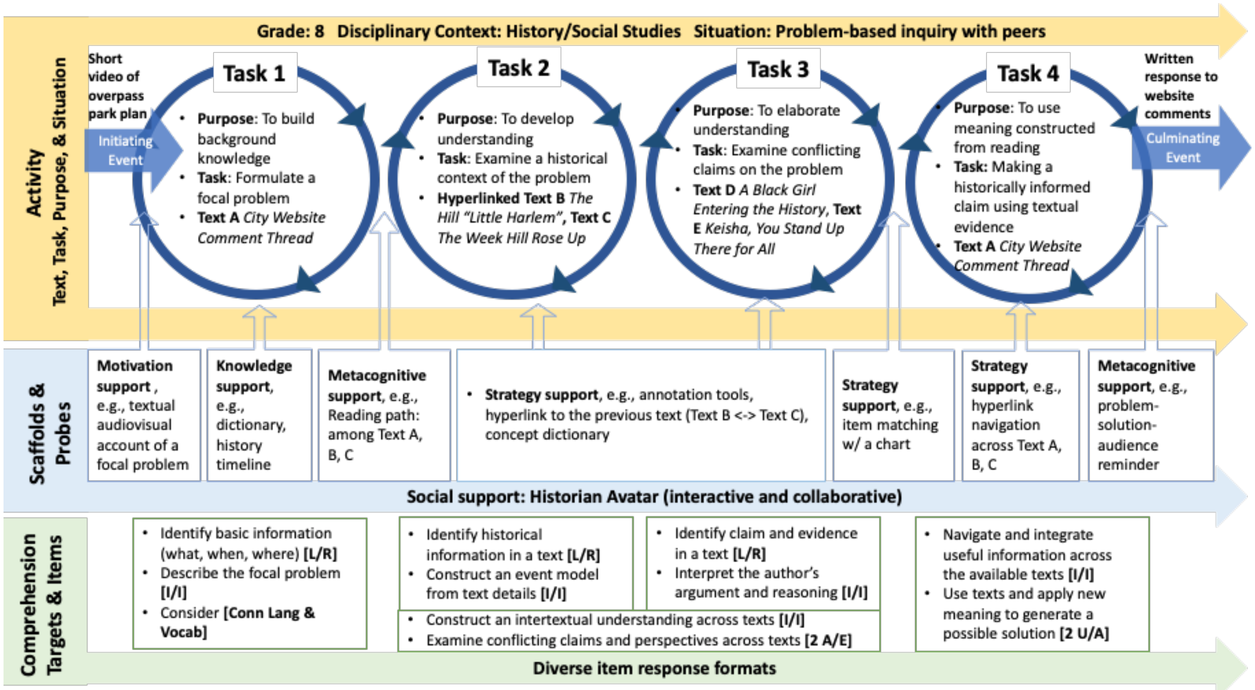
Tasks. Readers are asked to engage in the activity across multiple stages of reading (see Exhibit 4.18) to make sense of a focal *problem*, the historical *context* in which the problem is rooted, different *perspectives* on the problem, and the potential *action* in response to the problem. In the initial stage, students have opportunities to build background knowledge about the problem (i.e., People lack understanding of the Hill District and why Keisha matters). In the following stage, students will be guided through opportunities to learn about the history of the Hill District, which helps them explicitly understand ideas that might initially be confusing to park visitors. Topics are selected to help students build knowledge about various aspects of the Hill (e.g., vibrant cultures, thriving community businesses, discrimination, and segregation) to understand what it was like in the past and what has happened to the Hill from the 1940s to the present (e.g., urban renewal, demolition of the Hill, civil protests, civic arena and parking lot development). Students are supported in examining ideas from two different perspectives to help them to imagine a possible pathway to address the problem (e.g., how to clarify Keisha's role and why she is effectively positioned to fulfill the park planner's vision).

Texts. Readers are guided to comprehend and consult different forms of disciplinary texts and popular media texts. Historical texts may include both primary and secondary sources, such as historical photos and maps, archived black-and-white news articles, textbook-like written summaries, or visual timeline charts. Students may also be asked to read some online multimodal texts when learning about the problem and people's diverse opinions through news articles and website comments.

Readers also are supported in carrying out a series of historical reading tasks with specific purposes as they demonstrate the range of comprehension processes, such as those involved in close reading of a historical text, synthesizing within and across multiple texts, analyzing historical arguments using textual evidence, employing historical frameworks such as social structures or historical patterns, evaluating historical interpretations, and demonstrating historical perspectives. These tasks are also socially situated in that the purposes, processes, and consequences of reading are considered in relation to the challenges associated with urban development both locally, in Pittsburgh, and across the country.

Items. Comprehension processes are identified throughout the activity and linked to an appropriate balance of items among the intended targets (Locate and Recall, Integrate and Interpret, Analyze and Evaluate, Use and Apply). Given that this is a Reading to Solve a Problem activity block, more attention might also be given to Use and Apply items (with less focus on Locate and Recall items), so that readers have time to fully develop and express their solution to the problem in a 40-minute timeframe. Item difficulties might increase throughout the activity with variations in attention paid to unique text features and task demands as well as qualitative differences within each comprehension target category.

Exhibit 4.19. Concept Sketch of a Reading to Solve a Problem Activity Block: *Keisha Reconnects the Hill with Downtown in the City of Pittsburgh*



Scaffolds. As shown in Exhibit 4.19, the block design includes a range of digitally enhanced scaffolds and supports as readers comprehend texts, respond to items, and reflect on their performance. In the initial stage, an avatar (a regional historian) presents the reader with a primary purpose for reading; then, the reader (alongside avatar classmates) is asked to decide how to conduct brief research to find out more about the history of Pittsburgh's Hill District and generate their claims and responses to the inquiry question.

Cognitive and motivational support may include an image-based timetable that sequentially displays important local and national histories designed in the form of a graphic banner with pop-up notes. A list of keywords and relevant information offers a built-in knowledge support in the form of a searchable resource compilation (e.g., historical terms, specific names and places, civil rights movement). In line with important cognitive and motivational supports, all texts may need to be modified with a consideration of the level of text complexity suitable for 8th graders completing the entire activity in 40 minutes (e.g., passage length, structures, vocabulary, knowledge demands, motivational features). It should be remembered that NAEP will undertake a special study to consider extending block lengths to 45 or even 60 minutes.🔍 Diverse but intuitive response formats can be selected to support reader engagement and reduce the cognitive burden involved in expressing responses to test items. Students are likely to benefit from embedded metacognitive strategy support provided by avatar guides and/or a graphical overview of the entire reading activity to help monitor where they are and where they should focus their attention next to work toward the culminating task. Ultimately, decisions about supports and scaffolds should be specific to the block as designers consider what is needed to fulfill the goal of making the assessment asset-driven.

Performance Evidence and Indicators. The Hill District activity block information reveals what Grade 8 students can do when Reading to Engage in Social Studies to solve a problem. Ultimately, NAEP produces descriptions of what 8th graders (or sub-groups of 8th graders) can do in each disciplinary reading context. Thus, participating in the Hill District activity block (and other Reading to Engage in Social Studies activity blocks on the assessment), it is possible to characterize how well eighth-grade students are able to comprehend and use multiple sources while engaging in social-studies inquiries involving a collection of complex texts and a range of digitally enhanced items and supports.

Conclusion

The opportunities presented by the use of these innovative design features come with a caveat. Pilot offerings of all design features, including the examples above, should be carefully studied, as was noted in the introduction to this chapter. Various reader populations should be sampled carefully in these studies. One reason for this is to ensure that design features yield their intended outcomes for as many students as possible. A second reason is to ensure that new design features do not unintentionally disadvantage some populations of students. In addition to describing how scores will be reported, Chapter 5 illustrates how these new design features allow the 2025 NAEP Reading Assessment to explain the reading accomplishments of the nation's children in newly revealing ways.

CHAPTER 5: REPORTING NAEP 2025 RESULTS AND EXPLANATORY FRAMEWORK

Introduction

The purpose of Chapter 5 is to describe how the results of the NAEP Reading Assessment will be communicated to the nation from the year 2025 onward. This description completes the framework cycle that started with the overview of the 2025 NAEP Framework (Chapter 1), extended to the underlying sociocultural model of reading upon which the framework is built (Chapter 2), continued with the conceptual grounding for the NAEP Reading Assessment Construct (Chapter 3), and moved on to the assessment design in chapter 4. In bringing that cycle to a close, Chapter 5 addresses the central communication responsibility of NAEP—to report scores in a manner that informs the public about current results and performance trends over time on NAEP Reading Assessment in what has become known as the Nation's Report Card.

In addition to describing how scores will be reported, Chapter 5 outlines how the 2025 NAEP Reading Assessment can better **explain** the results it reports. NAEP has always served as the gold standard for reading assessment, but this new initiative has the potential to position NAEP Reading as an even more useful resource for informing educational policy related to teaching reading and learning to read—and to serve as a model for what assessment systems at the state and district levels might attempt to do. The shifts required for NAEP to assume such an explanatory role include:

1. Reframing the reporting system for NAEP Reading Assessment within the sociocultural model guiding this entire framework.
2. Undertaking a major revision in the items included in the reading-specific and the general (i.e., core) part of the questionnaires administered to students, teachers, and administrators whose schools participate in the NAEP Reading Assessment—with a goal of increasing the emphasis on opportunity to learn and student support.
3. Disaggregating scores for demographic subgroups in greater detail to increase explanatory power and guard against inappropriate inferences. For example, NAEP could not only report by race/ethnicity and by socio-economic status (SES) but also by SES within race/ethnicity.
4. Expanding the number of categories for reporting the achievement of English learners (ELs) to better reflect the variability of English language proficiency within this population.
5. Adding questionnaire-like items to the assessment blocks students complete to gain more specific information about how they experienced each activity block of the assessment. The explicit goal is to obtain more precise and sensitive indices of reader variables such as motivation, interest, metacognition, and agency.
6. Transforming the navigational data (what are often referred to as process data, Ho, 2017) that are collected when students take the test (how students make their way through the texts and test items,) into measures that can be used to explain test performance, especially for indices of interest and metacognition.
7. Enhancing the visibility and capacity of NAEP Reading (including enhancements for the NAEP Data Explorer) to encourage educators, policy makers, and researchers to utilize the currently available databases to conduct more nuanced analyses of NAEP Reading performance.

8. Reconceptualizing reporting and explanatory variables as an integrated system designed to explain student performance in a way that better informs policy makers and educators.

NAEP Reporting Categories

The Opportunity

Grounded in a tradition of groundbreaking reading frameworks, the 2025 NAEP Reading Framework provides some serendipitous opportunities to address persistent inequities in reading achievement. Reading is now understood as a relational process that involves both cognitive skills and the sociocultural contexts that affect reading processes and reading comprehension. The reporting system suggested for the 2025 NAEP Reading Framework, therefore, seeks to enable the nation and communities to address equity issues by expanding on ways to analyze and report the performance of diverse student populations. Specifically, the new reporting system offers NAEP the opportunity to broaden the interpretations that might be drawn from NAEP Reading results by amplifying the demographic and descriptive student categories. The new reporting system could expand understanding of how SES and race/ethnicity intersect with opportunities in schools and communities (e.g., the availability of libraries or access to challenging curricula). To support productive interpretations of results, the reporting of achievement results for the NAEP Reading Assessment could benefit from disaggregated reporting of SES and of English learners. If adopted, this richer reporting could expand analyses that account for diversity to help explain how and why students do or do not become literate in and outside of schools.

The Tradition

Historically, NAEP Reading has reported data for the nation as a whole, for participating states, and for large urban school districts that volunteer to participate in the NAEP Trial Urban District Assessment—TUDA. The NAEP Data Explorer is a publicly accessible tool that allows all users, but especially state and local administrators, to customize reports and to investigate specific aspects of student reading achievement.

Furthermore, NAEP has reported results by subgroups, including by school characteristics (e.g., urban, suburban, rural) and by socio-demographic student characteristics, such as gender, race/ethnicity, SES, disability status (i.e., supported by an individualized educational program), and English learner status. Although NAEP Reading Assessment results have provided indispensable information on students' performance, traditional reporting variables have portrayed *where the nation's students are* by parsing the data into subgroups, but the design of NAEP has not offered explanations about *why* different groups may perform as they do on reading comprehension measures. Even so, by expanding reporting categories and adding more explanatory variables, NAEP could point the way to plausible hypotheses for policy makers to consider in crafting reforms.

Expanding Reporting Categories

The 2025 NAEP Reading Framework builds on the strengths of the prior NAEP reporting system, but based on the sociocultural model of reading, calls for several initiatives to enhance the reporting and explanatory capacity of NAEP, including:

1. Reporting assessment data by reading Contexts (i.e., Reading to Engage in Literature, Reading to Engage in Science, and Reading to Engage in Social Studies)

2. Attending to the variability within groups (e.g., disentangling race/ethnicity and SES); and
3. Expanding the number of categories for reporting the achievement of English learners.

Reporting by Disciplinary Contexts. Aligned with the sociocultural model of reading comprehension, the 2025 NAEP Reading Framework calls for capturing and reporting reading performance within and across three disciplinary Contexts: Reading to Engage in Literature, Reading to Engage in Science, and Reading to Engage in Social Studies. For each of these Contexts, outcomes can be reported in two ways: (a) as a point on a scale that has, in the past, stretched from 0-500, and (b) as the percentage of students who score within different achievement level bands: *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*. In addition to the three disciplinary scales, scores are aggregated across the scales to create an overall performance scale. In contrast to the 2025 NAEP Reading Framework, the 2009-2019 framework called for two reported subscales: reading for literary experience and reading for information.

Disaggregating Results Within Demographic Categories. NAEP will continue to report reading scores by selected student subgroups. Student subgroups are defined by the following characteristics: gender; race/ethnicity; family income, as measured by student eligibility for the National School Lunch Program; disability status; and English language proficiency. In addition, results are reported by school characteristics, such as public/private, urban/rural, and region of the country.

Because the 2025 NAEP Reading Framework seeks to capture the dynamic variability within student groups, NAEP should disaggregate student group data to show, at a minimum, SES differences within the student subgroups of gender, race/ethnicity, and English language proficiency. In NAEP Reading, as in other large-scale assessments, lower levels of achievement historically are correlated with poverty. It is important to note that on international assessments such as PIRLS (Mullis, Martin, Foy & Drucker, 2011) and PISA (OECD, 2011), SES does not predict achievement in reading comprehension as accurately as it does in the U.S. Put another way, poverty does not offer a direct explanation for reading performance. Rather, access to resources that support rich literacy opportunities (e.g. high quality teaching, rigorous curriculum, community-based institutional structures such as libraries) is the underlying driver of achievement. Findings such as these could be important on NAEP because they might suggest that what is, on the surface, a characteristic of students that is difficult if not impossible to change, when examined with a more nuanced lens, turns out to be a variable that is highly amenable to change—resource allocation. The reporting of this sort of disaggregated data at the national level should be helpful to policy makers. When the data are disaggregated by states and districts, it also should be helpful to these stakeholders for addressing the needs revealed by the assessment.

Expanding Reporting Categories for English Learners. The 2025 NAEP Reading Framework calls for disaggregating students' status as English learners into additional reporting categories. English learners are defined by NAEP as students "who are in the process of acquiring English language skills and knowledge" (NAEP Nation's Report Card, 2019). These students have not yet reached the state-established standards for grade-level English proficiency, and so are at the beginning phases of acquiring English. Since 1998, NAEP scores have been disaggregated by students' "status as English language learners" into two groups: students designated as *not English language learners* vs. those designated as *English language learners* at the time of the assessment. The 2025 NAEP Reading Assessment results will expand these reporting categories in order to present data about English learners in a way that is more nuanced, more attuned to the complex reality of today's student populations, and, thus, more informative for states and school communities (Durán, 2006; Hopkins, Thompson, Linquanti, August, & Hakuta, 2013; National Assessment Governing Board, 2014; Kieffer & Thompson, 2018).

In line with the most recent research and with the direction in which state-level reporting is moving, the reporting system for the 2025 NAEP Reading Assessment calls for disaggregating scores by more refined English proficiency categories available through school systems participating in NAEP:

1. Current English learners--students designated as English learners at the time of the assessment;
2. Former English learners—students who have reached grade level standards of English proficiency within the last two years prior to the assessment;
3. Never English learners—monolingual students (they speak only English) and bilingual students (they speak English and another language) who were never classified as English learners throughout their entire educational trajectories.

This new approach can illuminate the most salient challenges and strengths for particular groups of English learners. For example, English learners might encounter challenges with background knowledge because their extensive knowledge and rich experiences might not reflect the knowledge and experiences represented in all-English classrooms, curricula, and assessments. Additionally, vocabulary and the language of assessment, while a challenge for many monolingual students, tend to pose a considerable challenge, not surprisingly, for students who are still in the process of learning English as a new language (García, 1991). At the same time, English learners can leverage the resources already learned in their home languages in the service of meaning making around texts. For instance, English learners whose home languages share cognates with English (e.g., Spanish/English: *inteligente/intelligent)* or those literate and/or schooled in their home languages draw from the language and literacy resources learned in their home languages to support their English reading comprehension (García, Sacco, & Guerrero-Arias, 2020; Phillips, Galloway, Uccelli, Aguilar, & Barr, 2020).

States are increasingly able—and required—to capture this more nuanced information on English learners' history in their data systems. If this trend continues, it is likely that by 2025, the majority of states, if not all, might have these data available as a matter of course; moreover, they might be collected in a process that is more reliable than the current approach used in NAEP Reading. The decision to add two more reporting categories is key because current English learners are an ever-shifting group identified by their emerging proficiency in English, as defined by the states. In the prior NAEP reporting system, if former English learners who reached standards of English proficiency and academic achievement were part of the broader group of never-English learners, this conflation made the success and increases in English learners' achievement impossible to observe or track. As noted by Kieffer & Thompson (2018), "focusing exclusively on the current EL subgroup can obscure the progress that educational systems make in moving students toward English proficiency and higher levels of academic achievement" (p. 393). It is in the best interest of the students, the states, and the nation to be able to document this progress accurately. Expanding NAEP's capacity to report the progress of English learners more accurately and with greater explanatory power should be a high priority for NAEP.

NAEP Explanatory Variables

The sociocultural theory, with its emphasis on the cultural assets of individuals and the power of context to shape learning and development, leads naturally to the identification of explanatory variables for reading. What explanatory variables have in common is that they are used to predict, explain, or account for variance in an outcome of interest, such as reading comprehension scores on NAEP. The construct of cultural validity provides a useful lens to guide the development of instruments to capture the proposed explanatory variables. Cultural validity refers to "the effectiveness with which the assessment addresses the sociocultural influences that shape student thinking and the ways in which students make sense of [test] items and respond to them" (Solano-Flores & Nelson-Barber, 2001, p. 555). Attention to cultural validity in assessments means anticipating how students with different background experiences will interpret what is being asked of them. Cultural incongruity can emerge among students' assumptions about the assessment topics, texts, and what they are being asked to do on the assessment and the assumptions of the assessment designers. Key questions to ascertain cultural validity include: Has the measure been tested for validity using a sample of students from diverse backgrounds? Do diverse communities of students understand the construct in relatively the same way with shared understandings of what the construct entails? (Vaux, Phillips, Holly, Thompson, Williams & Steward, 1986).

In addition, attention to cultural validity considers both cultural homogeneity as well as heterogeneity within cultural communities indexed by race/ethnicity and other salient group characteristics. In order to tackle this complexity, relationships among the comprehension outcomes on the assessment and the explanatory variables will include *disaggregation* within racial/ethnic groups by SES and gender as well as *examination* of the explanatory relationship between responses to survey items (e.g., student access to classroom, school, and community resources and their perceptions of classroom and community conditions) and NAEP Reading comprehension scores.

The 2025 NAEP Reading Framework calls for a shift from prior years in the scope of explanatory variables collected in conjunction with the NAEP Reading Assessment. This shift is informed by sociocultural perspectives that view reading as social and cultural practices that influence how readers approach, engage with, and make meaning from texts (Pacheco, 2015, 2018). Readers' values, beliefs, experiences, and ways of communicating and thinking are all drawn from their everyday experiences (Lee 2007, 2016). Readers' histories of engagement with texts also affect how often they read, the types of texts they read, and their purposes for reading (Cazden, 2002; Heath, 1983, 2012; Lee 1993, 2005). To explain students' differential performance on the NAEP Reading Assessment, it is important to take into account their differential engagement with reading and their access to home and community resources such as libraries, tutoring, and out-of-school programs. Although these issues are addressed to some degree in the 2019 NAEP Reading questionnaire, they could be expanded to form more robust constructs capable of explaining students' reading achievement.

An Expanded Set of Explanatory Variables

A unique feature of the 2025 NAEP Reading Assessment will be an expanded and more robust set of explanatory variables. The explanatory variables have been selected on the basis of prior research (Guthrie & Klauda, 2015; Guthrie, Wigfield & Von Secker, 2000) that has demonstrated their potential to a) predict reading achievement, b) exhibit malleability (can be influenced by changes in curricular or pedagogical practices), and c) be measurable.

The set of explanatory variables includes both reader attributes (e.g., students' self-perceptions about engagement and motivation, knowledge, self-efficacy, agency, effort, and interest) and environmental variables (students' perceptions about facets of home, community, or school settings, including their perceptions about classrooms, sense of affiliation, and support).

Reader Attributes. With respect to reader attributes, NAEP should seek to describe the role of students' perception of the interest, difficulty, and familiarity of texts, tasks, and contexts (Pintrich and Schrauben 1992; Eccles, O'Neil et al. 2005; Valencia, Wixson et al. 2014). The sociocultural model and the assessment construct call for better understanding of the role of student self-efficacy in carrying out particular tasks (Bandura 1993; Pajares 1996) and the relevance of such tasks for students' motivation and engagement (Guthrie and Wigfield, 2000). If students do not see an assessment as meaningful or relevant, then it may not adequately capture what they are able to do. In other words, the conceptualization of reading comprehension as a sociocultural process raises the question of whether students might demonstrate different levels of competency if the social and cultural circumstances of the assessment were different (Valencia, Wixson, & Pearson, 2014).

Environmental Variables. Environmental variables are equally important in accounting for student performance. For example, students vary in their history of participation in cultural communities that may value and integrate reading into their lives for different purposes, such as sharing stories, reciting poetry, or reading silently or jointly (Skerrett, in press). These histories of engagement and participation constitute sociocultural resources readers accumulate across their lifetimes and that they bring to bear on reading tasks, including NAEP assessments. Furthermore, what it means to read has evolved over time as cultural communities and societies have employed texts for different purposes and goals. It is imperative to understand students' differential access to community resources that support literacy development, such as libraries, tutoring supports, and out-of-school programs across grade levels. As these contexts have shifted, so has the role of reading and texts in students' lives. As schools and communities offer more or less access to these out-of-school resources, students' opportunities to learn, including their own self-initiated learning, may vary considerably too.

Aims of the Expanded Variables. The goal of the expanded variables is to offer insights into pathways that might influence students' efforts and their demonstrated competence. For example, their reported sense of reading engagement and motivation (including the affective dimension of reading enjoyment) could be positively related to higher levels of NAEP Reading performance (Guthrie, Wigfield & You, 2012). Students' positive perceptions of their teachers' supports and classroom climate could also be associated with higher NAEP Reading performance (Pitzer & Skinner, 2017). If relations such as these emerge from NAEP, it would have meaningful implications for how attention to perceptions, identity, and affect might be addressed to support reading comprehension and achievement. Consideration of such factors is consistent with research on the importance of social and emotional wellbeing to learning (Damasio 1995; Markus and Kitayama 1991; Weiner 1985), the incorporation of social-emotional learning into the design of classroom and school climate (Farrington, Roderick et al. 2012), and approaches that build on and engage students' out-of-school identities and interests to make learning meaningful and relevant (Shin, Daly & Vera, 2007).

These new variables can also add deeper explanations for surface level findings. For example, girls are often higher achievers than boys, but this information is of limited utility for pedagogical or curricular improvement. Girls often exhibit higher motivation than boys, and they spend more time reading than do boys. When boys and girls are equated on reading time, the gender gap disappears. (Torppa, Eklund, Sulkunen, Niemi & Ahonen, 2018). Since both reading time and motivation can be impacted by interventions, the more nuanced explanation of the gender difference could inform educators about the need to reorganize instruction and improve support for reading opportunities for boys in schools.

Format and Location. Beyond expanding the coverage of explanatory variables, 2025 NAEP Reading Framework also introduces a shift in the format and placement of survey items. Some items will remain in the *questionnaires* that are routinely completed by students, teachers, and administrators from participating schools. Others will be integrated into the actual *activity blocks* containing texts and comprehension items, e.g., students' perceptions of their knowledge, interest, self-efficacy, and motivation around particular texts/tasks after reading (Educational Testing Service, 2019). Finally, information about some of the variables will be obtained from the process data (computer-generated records) collected automatically as students navigate their way through the assessment (Bergner & Davier, 2018). Exhibit 5.1 provides a list of variables, along with their location and/or source in the revised explanatory variable plan.

Exhibit 5.1. Explanatory Variables

Reader Attributes	Source			
Cognition and Metacognition	SQ	TQ	BA	PD
Cognitive and metacognitive strategies in reading comprehension	√			
In block rating of background knowledge for specific topics within the block activity			√	
Process data derived from computer monitoring of students' navigation within the tasks completed				√
Familiarity with texts read in school	√		√	
Effort	√		√	√
Confidence	√		√	
Engagement and Motivation	SQ	TQ	BA	PD
Amount of text reading for enjoyment	√			
Motivations for text reading in school and recreation	√		√	
Engagement and motivation for specific assessment passages			√	√
Environmental Variables				
Perceptions of School and Community Resources	SQ	TQ	BA	PD
School social support	√	√		
Belonging in school	√	√		
Participation in out-of-school reading/literacy activities	√			
Perceptions of Teacher, Instruction, and Classroom Supports	SQ	TQ	BA	PD
Teachers' support for reading engagement	√	√		
Teachers' support for motivation	√	√		
Reading/language teachers value student cultures and life experiences in reading instruction	√	√		

*SQ = Student Questionnaire, TQ = Teacher/Administrator Questionnaire, BA = probe in Block Activity, PD = derived from Process Data.

A Plan for Enhancing NAEP's Explanatory Reporting Capacity

This chapter was intended to provide evidence on which to evaluate progress and potential of transforming NAEP's reporting capacity from simply reporting to explaining results, the question is, how does NAEP move ahead, balancing both the wise counsel of due caution and the equally wise commitment to ambitious goals to move toward the goal of more revealing and useful reports? Any plan must be based on the eight shifts set out at the beginning of this Chapter.

1. **Reframing the Reporting System Within a Sociocultural Construct.** If the assessment is based on the sociocultural model, as this Framework states, then the reporting system should reflect the sociocultural model as well. Such a reframing will enhance NAEP's explanatory power and key role in promoting equity in American education.
2. **Revising Questionnaires**. To increase the capacity to examine the impact of variables like opportunity to learn and student support, NAEP should eliminate questions that perpetuate stereotypes. A thorough review of current surveys—both the reading-specific and core questionnaires for the three categories of participants (students, teachers, and administrators)—should be implemented to determine questions that ought to be revised,

replaced, or discarded. While ensuring the appropriateness and sensitivity of all NAEP questionnaire items, this review also enables development of questions based upon the constructs outlined for questionnaires in Exhibit 5.1.

3. **Disaggregating Scores to Achieve More Nuanced and Explanatory Reporting.** The case for disaggregation is strong. It is time for NAEP to pre-empt stereotypic characterizations of the all-too-persistent achievement gaps in American education. Adding nuance to the reporting of performance for the major demographic categories (e.g,, SES within race/ethnicity) is a good first step in disarming unwarranted inferences about students who fall into the reporting categories.
4. **Expanding Reporting Categories for English Learners**. Expanding the number of categories for reporting the achievement of English learners (ELs) will enable NAEP to track the progress of different subgroups. Most important is the added category of Former ELs. At present, they are grouped with the group of students not currently receiving English language instruction, a group that contains both Never ELs and Former ELs. In the proposed approach, the performance of Never ELs and Former ELs would be reported separately, making it possible to determine whether the two groups perform at similar levels on the NAEP Reading Assessment.
5. **Adding Mini-Probes to the Activity Blocks**. NAEP and its partners should work to find at maximum an extra three minutes within each activity block to add these probes to the assessment. An initial study of the efficacy of these mini-probes suggests that (a) students can complete them very rapidly (one every eight seconds on average) and (b) they explain respectable amounts of the variation in student performance on the comprehension items in the activity blocks (Educational Testing Service, 2019). It appears that students are better able to answer a question about how hard they try if they are doing the self-assessment in relation to a specific task they just completed than in relation to schoolwork in general. If these probes explain reading comprehension performance for cognitive and motivation variables, especially if they explain more variance than parallel reading survey items designed to measure the same constructs, they could free up space for more student questionnaire items devoted to other important constructs, such as opportunity to learn or student support.
6. **Mining Process Data for Evidence of Cognitive and Metacognitive Processing**. Initial forays evaluating the utility of the process (logfile) data for NAEP (Bergne & von Davier, 2018) and other digitally delivered assessments and instructional programs (Ho, 2017) suggest that there is substantial potential for using these navigational data as indirect indices of cognitive and metacognitive processes. These indices can be used, perhaps in triangulation with measures of the same variables from activity block-based mini-probes or even reading questionnaire responses, to understand comprehension performance more deeply. Simple bar graphs can be displayed in the Report Card, and data can be related to reading performance in the NAEP Data Explorer.
7. **Enhancing the Visibility and Utility of the NAEP Reporting Portfolio.** There is no question about the visibility of NAEP results within the educational policy arena, both at the national and state level. When the NAEP Report Card for Reading is issued every two years, policy makers and the public pay attention, particularly to trend data. NAEP results are used to argue for one or another type of reform or funding (including redistribution of funding) initiative, and are often expressed in stark phrases, such as, "40 percent of America's fourth-grade students read below the NAEP Basic level." NAEP results have long been subject to

misinterpretation (Linn and Dunbar 1992; Jaeger 2003; National Research Council 2017). NAEP reporting is subject to two almost contradictory problems. First, because the results are reported in such broad categories (Race by Grade or Language Status by School Setting--Urban/Rural), they are overly or inappropriately interpreted. Second, in the Report Card for Reading (NAEP, 2019), achievement is seldom reported as a function of malleable factors, either for reader variables (e.g. student motivation) or environmental variables (e.g., opportunity to learn indices). As a result, readers of the Report Card may rarely look beneath the surface, focusing instead on broad indicators such as whether the percentage of students scoring at NAEP Proficient or above rose or fell over the past decade. Readers of the Report Card may simply accept these broad indicators and draw their own conclusions about what the results mean for reform initiatives.

Four initiatives can help to ameliorate this situation. First, if NAEP implements the previous seven shifts, it will be well down the road to implementing the eighth shift—building the kind of reporting system that focuses on explaining the reported results, not leaving important policy inferences to the political or ideological predilections of the readers of the Nation's Report Card.

Second, NAEP needs to expand, energize, and advertise the untapped resources it already has. As an example, the NAEP Data Explorer permits any user to go on-line and generate much more nuanced analyses than typically appear in the Report Card, which, by its very nature, provides only foundational reporting. In the current version of NAEP Data Explorer for the 2019 Reading Assessment, a user could query the database to obtain a report which, for 4th graders in the nation, breaks down the performance of Low- versus High-SES students on the cognitive targets of Locate and Recall, Integrate and Interpret, and Critique and Evaluate when reading literary and informational text. For sound psychometric reasons, NAEP results are not reported separately for the Comprehension Targets, but that does not mean that NAEP data cannot be used to obtain statistically reliable reports that report those targets separately.

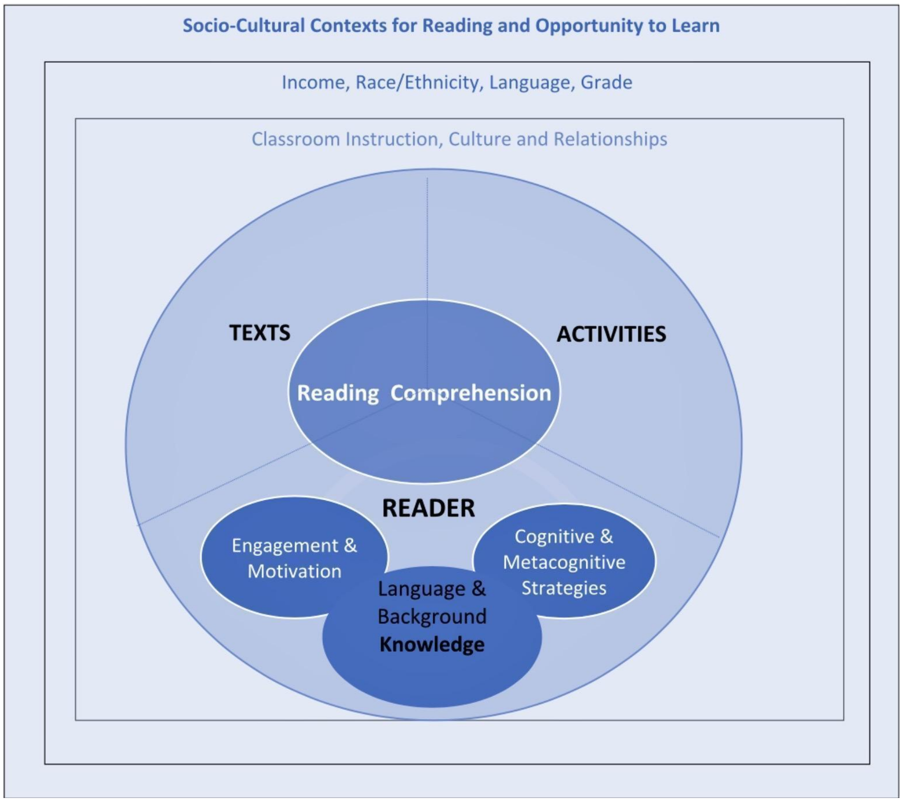
Third, NAEP has a long tradition of funding small grants for secondary analyses that permit scholars to answer, in a statistically robust design, the sorts of questions that anyone could query with the Data Explorer tool. Increasing the funding for these initiatives would dramatically increase the portfolio of the more nuanced explanatory analysis called for in this framework.

Fourth, NAEP should replicate the 1998 study conducted by the National Validity Studies Panel (Jaeger, 1998) regarding how NAEP results are used by policy makers and educational leaders, with a focus on whether the inferences that users draw from the NAEP Report represent valid interpretations of the evidence.

Shift #8: Reconceptualizing the Entire Collection of Reporting Categories and Explanatory Variables as an Integrated System Designed to Explain Student Performance. If NAEP implements these first seven shifts and takes these four recommendations, it would be enacting the eighth shift by creating an integrated system of explaining student performance. To this end, NAEP should undertake a systematic study of its reporting portfolio, taking into account reporting variables, explanatory variables, and the all important outcome variable of comprehension, to create and evaluate the efficacy and utility of just such a system.

NAEP should implement a special study to evaluate the costs, benefits, and feasibility of such a system. To that end, a proposal to conduct a project of this scope will be detailed in the *Assessment and Item Specifications*. 🔍 Perhaps the central attribute of a newly envisioned reporting system for the 2025 NAEP Reading Framework would be that it is based on the sociocultural model of reading detailed in Chapter 2. A visual representation of that model appears in Exhibit 5.2. The representation in 5.2 is similar to the visual model in Exhibit 2.1 of Chapter 2, but adapted, both in its visual framing and category labels, to represent the emphasis on explanatory variables of Chapter 5.

Exhibit 5.2. Sociocultural Model of Reading Comprehension



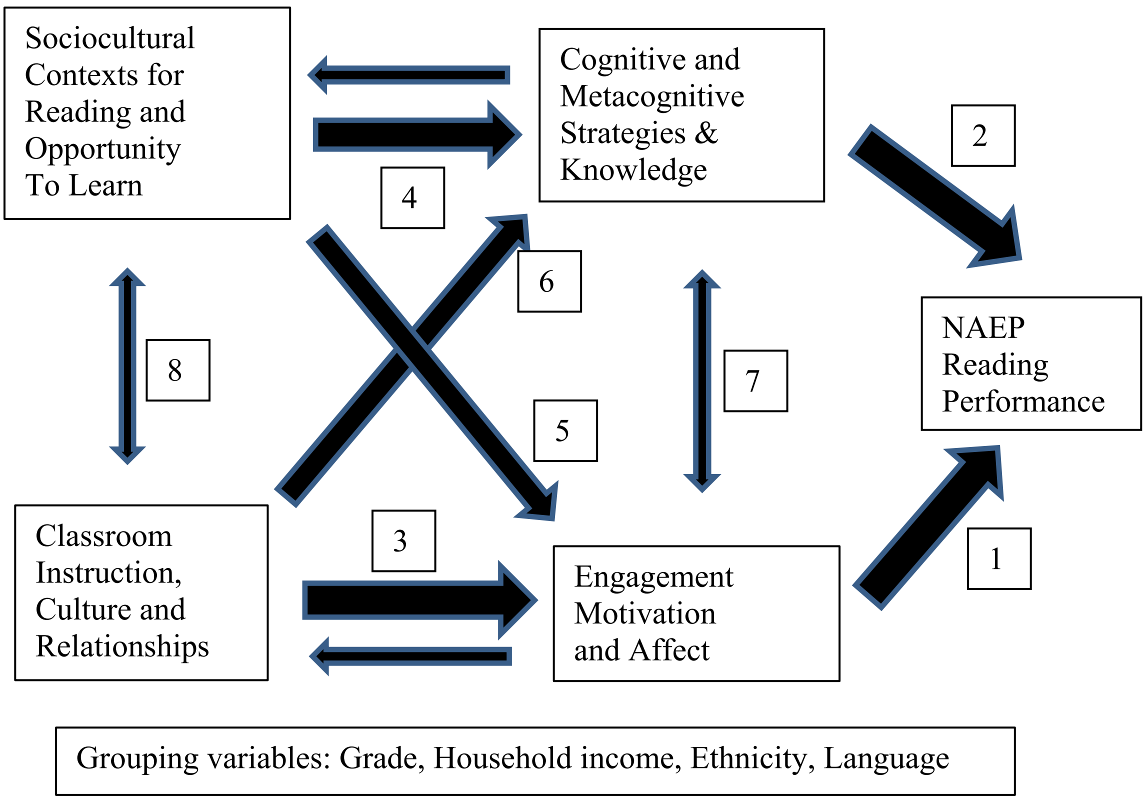
Important to note in Exhibit 5.2 is the positioning of reader attributes toward the center of the reader-text-activity triad and, by contrast, the inclusion of most of the environmental variables toward the periphery of the model. That does not imply that the environmental variables are peripheral, only that they surround, and hence always influence, the reader-text-activity triad. The proposal for such a system in the *Assessment and Item Specifications* document goes much further than locating all of these variables in the same space; it proposes highly specific explanatory prediction paths among these variables.

These hypothetical (but firmly research-grounded) paths are represented visually and spatially in Exhibit 5.3. Four features of this hypothetical model are noteworthy. First, even though it is a metaphor for a path model, some of the arrows are bidirectional. Bidirectional arrows imply synergy, with a change in either affecting the other. Second, not all variables have a direct connection to all other variables. For some, for example, the relationship of instruction to comprehension, is not direct but mediated, at least in design, by both cognition and engagement.

Third, the wider arrows depict relatively stronger relationships than narrower arrows. Fourth, this is a truly multivariate space in which many variables mutually influence one another in both modest and profound ways.

The path model in Exhibit 5.3 also reveals the organizational plan for explaining reading performance in 2025 and beyond. On the right side of the graphic is a box representing reading performance—either the scale score or the achievement level. The organizational plan plots various pathways in which variables influence, predict, and explain student performance. Exhibit 5.3 illustrates the hypothesized relationships among these explanatory variables and reading comprehension.

Exhibit 5.3. A Path Model for Reading Performance and Explanatory Variables



This brief representation of a systematic approach to conceptualizing reporting and explanatory variables does not do justice to the complexity of the relations among those variables. A more specific account of the relationships among these variables will be available in the *Assessment and Item Specifications*, complete with recommendations for the types of analyses and even the number of items required to construct a trustworthy scale for the entire set of explanatory variables.

Conclusion

NAEP 2025 recognizes that even with all the recommendations proposed, a one-time assessment can, at best, provide only an insightful proxy for the reading comprehension performances of the nation's students, and then only for the particular conditions that prevail in NAEP. Reading comprehension performances vary depending on the combination of individual and contextual factors at the time of the assessment. Thus, NAEP Reading scores provide only a snapshot of the nation's students' reading comprehension performance as displayed in a particular testing situation at a certain moment in time.

Recognizing these inherent limitations, the assessments derived from the 2025 NAEP Reading Framework nonetheless offer novel opportunities to explore the validity, efficacy, and utility of tapping into students' sociocultural assets to provide a more precise and accurate account of their reading comprehension. The NAEP Reading Assessment attempts to account for the role of background knowledge, readers' perceptions about the relevance and social utility of comprehension tasks, use of cognitive and metacognitive strategies, and socioemotional factors. This update of the NAEP Reading Framework provides opportunities to examine malleable explanatory variables that can help explain comprehension scores. The identification of malleable factors by the 2025 NAEP Reading Assessment reporting system also provides information that educators and policy makers can use to improve students' reading comprehension instruction and performance.

Moreover, the disaggregation of reporting that examines heterogeneity within groups (e.g. race/ethnicity, SES, gender, English learners) will also be important. Such disaggregation will help the field and the nation resist the persistent temptation to overgeneralize about groups. Efforts to disaggregate scores beyond what has been done in past iterations of the NAEP Reading Assessment provide opportunities for further explanatory power and greater utility for practice and research.

As NAEP embarks on a more ambitious and comprehensive reading assessment design, it has the potential, as have previous NAEP frameworks and assessments, to serve as a model for other assessment efforts at the state and district level. NAEP differs from almost all other assessments at the state, district, and classroom levels primarily because it does not report individual scores. Even so, framed by its influential sociocultural model of reading, the 2025 NAEP Reading Framework has the potential to point the way for other assessments to take the same equity-based steps to produce a new portfolio of assessments that recognize that students' cultural, social, and knowledge assets should be acknowledged, indeed leveraged, in evaluating student performance. And these parallel assessments could and should aspire to diagnostic goals that could help teachers plan classroom and individual instructional activities. Teachers will need access to diagnostic assessments that help them understand students' cultural assets, their cognitive and metacognitive processing skills, their motivations for reading, and their perceptions of the relationships and emotional supports in their classrooms and schools.

The new reporting system for NAEP, articulated above, would provide a wealth of new data sources for policymakers at state and district levels. Having access to reporting by states and districts, including networks of smaller rural districts, can inform state- and district-level initiatives around factors that not only predict performance but that are also malleable. Such state- and district-level reporting allows policymakers to re-examine policies intended to support students and teachers.

Finally, the new reporting system offers opportunities for researchers. In addition to the more expansive public reporting system on the Nation's Report Card, researchers will have access to a wider range of data for exploring foundational questions around the dynamic nature of reading comprehension. Ultimately, the sociocultural model of reading that informs NAEP 2025 can shape future investments in expanding student access to robust opportunities for reading and literacy engagement in and beyond schools.

Glossary terms placeholder

The NAEP Reading achievement level descriptions (ALDs) articulate specific expectations of student performance in reading at grades 4, 8 and 12. Like other subject-specific ALDs, the NAEP Reading ALDs presented in this appendix translate the generic NAEP policy definitions into grade- and subject-specific descriptions of performance.

NAEP Policy Definitions

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

Range ALDs

This Framework presents range ALDs for NAEP Reading. For each achievement level, the corresponding range ALD details observable evidence of student achievement. In many cases, range ALDs also illustrate "changes" in skills across achievement levels, portraying an increasingly sophisticated grasp of the material from one achievement level (and from one grade level) to the next. Achievement levels are also cumulative, meaning each ALD in each grade includes all the reading achievement expectations identified in all the lower achievement levels and grade levels.

Range ALDs should not be confused with reporting ALDs. The fundamental difference between the two is straightforward; range ALDs communicate **expectations**, and reporting ALDs convey **results**. In other words, range ALDs are **conceptually driven**, based on the model of reading and the Assessment Construct in the NAEP framework. They answer the question, given what we know about the development of reading, what should students be able to do at different grade and achievement levels when responding to different combinations of texts and tasks? By contrast, reporting ALDs are **empirically driven**, based on **actual** performance of students who have taken NAEP. They answer the question, given the distribution of NAEP performance, what can students at different grade and achievement levels do when responding to various combinations of texts and tasks?

The 2025 NAEP Reading Framework does not provide reporting ALDs; those are constructed using empirical data during a later stage in the NAEP cycle, i.e., a live administration of the NAEP Reading Assessment. Further detail about the development of the reporting ALDs for NAEP is provided in the Governing Board's policy statement on achievement level setting.

Multiple Disciplinary Contexts for Reading

The ALDs in this appendix are structured to mirror the presentation of the reading construct provided in the Framework narrative. The primary organizational structure in the Framework narrative is the disciplinary context. Whereas the prior (2009) NAEP Reading Framework identified two reading contexts (literary and informational) this 2025 Framework has identified three (science, social studies, and reading). In the ALDs below, all three disciplinary contexts are described within each performance level.

Connections to the Sociocultural Model of Reading

Comprehension Targets and Text Complexity

Over the course of the NAEP Reading Assessment, students will engage with texts of various discourse structures and an appropriate grade-level range of text complexity. While reading these texts, students will complete varied reading comprehension activities that include specific purposes, tasks, processes, and consequences. The reader, per his or her achievement level, will employ various knowledge types to accomplish the assessment's reading comprehension activities. In doing so, the reader will demonstrate achievement relative to four *comprehension targets:* (1) Locate and Recall; (2) Integrate and Interpret; (3) Analyze and Evaluate; and (4) Use and Apply. Students at each achievement level are expected to meet the demands of each comprehension target. However, as the complexity of texts increases on a given reading assessment, students, on average, are expected to demonstrate less competency with skills associated with higher-level comprehension targets, such as Use and Apply.

Purposes

According to the sociocultural model, reading activities are situated within not only a disciplinary context but also a purpose. This section describes the mapping of reading purposes to disciplinary contexts.

Literary Texts. People engage in reading literature for the following purposes:

- To understand human experience
- To entertain themselves and others
- To reflect on and solve personal and social dilemmas
- To appreciate and use authors' craft to develop interpretations

In school, students read, create, and discuss literary texts such as poems, short stories, chapter books, novels, and films. Outside of school, students participate in book clubs, create fan fiction and book reviews, follow and discuss authors, dramatize literary works with animation and music, and more. NAEP simulates these Contexts of Reading to Engage in Literature by providing test takers with activities to respond to literary and everyday texts like those read in and outside of school.

Science Texts. People engage in reading science for the following purposes:

- To understand natural and material phenomena
- To design solutions to problems
- To explore and discuss issues and ideas
- To consider impacts on themselves and society

In school, students read, create, and discuss science texts such as explanations, investigations, journal articles, trade books, and more. They design solutions to engineering challenges, use diagrams and flow charts, and follow step-by-step procedures to investigate scientific phenomena. Outside of school, students engage in reading science when participating in games, cooking, and crafts, and reading and viewing science and health news. NAEP simulates these Contexts of Reading to Engage in Science by providing test taskers with activities to respond to science and everyday texts like those read in and outside of school.

Social Studies Texts. People engage in reading social studies for the following purposes, among them these:

- To understand past events and how they may impact the present
- To explore and discuss issues and ideas
- To understand human motivation, perception, and ethics
- To advocate for change for themselves and society

In school, students read social studies texts such as primary and secondary source documents, historical narratives in textbooks, case studies, current events, court cases, and more. They read, create, and discuss memoirs, timelines, and biographies. Outside of school, people engage in reading history and social studies when participating in trivia games, crafts, civic activities, community discussions, self-help, and community service. NAEP simulates these contexts of reading to engage in social studies by providing test tasks with activities to respond to history/social studies and everyday texts like those read in and outside of school.

NAEP Reading Achievement Levels: Grade 4

NAEP Basic

Fourth-grade students performing at the *NAEP Basic* level should be able to locate specific pieces of information, identify relationships between explicitly stated pieces of information, make simple inferences and interpretations within and between texts, create summaries, and show understanding of vocabulary in the disciplinary contexts.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, fourth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to identify or determine literary elements such as character point of view, theme or central message, problem, and setting. Readers should be able to explain how a text's illustrations contribute to what is conveyed by the text, explain the differences between poems, drama, and prose, and show understanding of vocabulary and simple figurative language. Readers can produce a simple summary of a text and continue the narration of an incomplete story to a conclusion of their making.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including investigations), fourth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to determine the main idea and how it is supported by key details, determine and interpret an author's point of view or purpose, and distinguish between fact and opinion. Readers should be able to interpret and integrate information presented in a text visually, quantitatively, and orally, analyze specific results of a simple multistep procedure, and show understanding of academic and domain-specific vocabulary. Readers can apply simpler ideas acquired through reading to solve a new problem.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, fourth-grade readers performing at the *NAEP Basic* level should be able to determine the main idea and how it is supported by key details, determine and interpret an author's point of view or purpose, and distinguish between fact and opinion. Readers should be able to describe the overall structure of a text and compare and contrast explicit information found in a firsthand and secondhand account of the same event or topic. Readers can produce a simple summary of a text and integrate information from lower complexity sources to produce a new text of informational or argumentative purpose.

NAEP Proficient

Fourth-grade students performing at the *NAEP Proficient* level should be able to make more complex inferences and interpretations, reconcile inconsistencies across a text or texts, and explain how an author uses reasons and evidence to support particular points in a text.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, fourth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to describe in depth character, setting, and plot, and to explain how a theme or central message is conveyed through details in a text. Readers should be able to analyze how a printed version of a text relates to its multimedia version and show understanding of nuances in word meaning. Readers can produce a detailed summary of a text and rewrite a story from a different character's perspective.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including investigations), fourth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to explain events, procedures, ideas, and concepts based on specific information in and across texts. Readers should be able to make predictions and to interpret an author's point of view or purpose, including in reference to a procedure or experiment and in comparison to another text's author. Readers should be able to develop a new procedure or experiment based on knowledge acquired from information gained from reading texts.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, fourth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to explain events, procedures, ideas, and concepts based on specific information in and across texts. Readers should be able to explain how information presented in a text visually, quantitatively, and orally contributes to an understanding of a text. Readers should be able to produce a detailed summary of a text and adopt the persona of a historical figure when producing a new text of informational or argumentative purpose.

NAEP Advanced

Fourth-grade students performing at the *NAEP Advanced* level should be able to make complex inferences and to support their interpretations, conclusions, and their judgments based upon evidence within and across texts.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, fourth-grade readers performing at the *NAEP Advanced* level should be able to use textual evidence as support to explain character motivation and behavior and how characters interact with setting and plot. Readers should be able to evaluate how characters or themes resonate with society and their personal lives. Readers should be able to apply knowledge acquired about author's craft to produce a literary work evidencing their understanding.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including investigations), fourth-grade readers performing at the *NAEP Advanced* level should be able to determine the significance of information and arguments made in a text. Readers should be able to make predictions and to interpret an author's point of view or purpose and to argue for or against a particular interpretation.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, fourth-grade readers performing at the *NAEP Advanced* level should be able to determine the significance of information and arguments made in a text. Readers should be able to make predictions and to interpret an author's point of view or purpose and to argue for or against a particular interpretation. Readers should be able to use acquired knowledge about a topic, conduct brief research, and produce a historical document, such as a political cartoon or a personal bill of rights.

NAEP Reading Achievement Levels: Grade 8

NAEP Basic

Eighth-grade students performing at the *NAEP Basic* level should be able to find information in dynamic and multimodal texts, make simple inferences and interpretations within and between texts, make predictions, create objective summaries, analyze word choice, and show understanding of vocabulary in the disciplinary contexts.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, eighth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to determine theme or central idea and aspects of character, setting, and plot. They should be able to compare basic literary attributes of two or more texts and make judgments about how each author presents events. Readers show understanding of vocabulary and figurative language. They can develop a simple objective summary of a text and produce an argumentative text that prosecutes or defends the actions of a character by using evidence from the reading text.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), eighth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to determine the central ideas and conclusions of a text and explain how a text makes connections among and distinctions between individuals, ideas, and/or events. Readers should be able to integrate quantitative or technical information expressed in words in a text with a version of that information expressed visually (e.g., in a flowchart, diagram, model, graph, or table), show understanding of how to follow precisely a multistep procedure of an experiment, and show understanding of academic and domain-specific vocabulary, key terms, and symbols. Readers can apply simpler ideas acquired through reading to solve a new problem.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, eighth-grade readers performing at the *NAEP Basic* level should be able to determine the central ideas, determine and interpret an author's point of view or purpose, and distinguish between fact, opinion, and reasoned judgment in a text. Readers should be able to identify key steps in a text's description of a process related to social studies (e.g., how a bill becomes law). Readers can produce a simple objective summary of a text and integrate information from multiple sources to produce a new text of informational or argumentative purpose.

NAEP Proficient

Eighth-grade students performing at the *NAEP Proficient* level should be able to make more complex inferences and interpretations, form explanations and generalizations, generate alternatives, and apply new ideas acquired through reading to a new problem or context. Students should be able to use text-based evidence to support arguments and conclusions.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, eighth-grade readers performing at the *NAEP Proficient* level should be able analyze the development of the theme or central idea over the course of a text and how particular lines of dialogue or incidents in a text propel, the action, provoke a decision, or reveal aspects of character. Readers should be able to analyze how a printed version of a text relates to its multimedia version and how text structure contributes to meaning and style. They can analyze how word choice impacts a text's meaning and tone. Readers can develop a detailed objective summary of a text and produce an informational text that analyzes how different authors developed a similar theme or central idea.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), eighth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to analyze the specific results of a multistep procedure based on explanations in the text, analyze how the author acknowledges and responds to conflicting evidence and/or viewpoints, and analyze how two or more texts provide conflicting information on the same topic, identifying where the texts disagree on matters of fact or interpretation. Readers should be able to compare and contrast information gained from experiments, simulations, video, or multimedia sources with that gained from reading a text on the same topic. Readers should be able to generate an alternative procedure or experiment based on knowledge acquired from information gained from reading texts.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, eighth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to explain how a text makes connections among and distinctions between individuals, ideas, and/or events (e.g., through comparisons, analogies, or categories). Readers should be able to analyze the relationship between a primary and secondary source on the same topic and analyze how two or more texts provide conflicting information on the same topic, identifying where the texts disagree on matters of fact or interpretation. They should be able to analyze the structure an author uses to organize a text and develop a detailed objective summary of a text. Readers can produce an argumentative text that proposes a form of social action based on knowledge acquired and opinions formed from the reading texts.

NAEP Advanced

Eighth-grade students performing at the *NAEP Advanced* level should be able to make complex inferences and to support their interpretations, conclusions, and their judgments based upon evidence within and across texts. Students should be able to evaluate the relevance and strength of evidence to support an author's claims.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, eighth-grade readers performing at the *NAEP Advanced* level should be able to use textual evidence as support to analyze how multiple literary elements in a text relate to each other and to analyze points of view of and between character(s) and the reader/audience. Readers should be able to analyze how a modern text draws on themes, patterns of events, or character types from myths or traditional stories, and then evaluate how these elements resonate with society and their personal lives. Readers should be able to produce a literary text that adapts elements of a myth into a contemporary retelling based upon the reader's personal experience.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), eighth-grade readers performing at the *NAEP Advanced* level should be able to analyze the development of the central idea over the course of the text. They should be able to delineate and evaluate the argument, claims, and reasoning in a text, including whether the evidence is relevant and sufficient to support the claims. Readers can produce a new argumentative or informative text that synthesizes information from a range of sources to demonstrate a coherent understanding of a process, phenomenon, or concept.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, eighth-grade readers performing at the *NAEP Advanced* level should be able to analyze the development of the central idea over the course of the text and analyze how the author acknowledges and responds to conflicting evidence and/or viewpoints. Readers should be able to delineate and evaluate the argument, claims, and reasoning in a text, including whether the evidence is relevant and sufficient to support the claims. They can produce an informative text that traces and connects various factors (e.g., economic and societal) by incorporating acquired knowledge through reading multiple sources and conducting brief research.

NAEP Reading Achievement Levels: Grade 12

NAEP Basic

Twelfth-grade students performing at the *NAEP Basic* level should be able to find information in dynamic and multimodal texts, make inferences and interpretations within and between texts, make predictions, create objective summaries, analyze word choice, and show understanding of vocabulary in the disciplinary contexts.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, twelfth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to analyze the development of the theme or central idea over the course of a text and to analyze points of view of and between character(s) and the reader/audience. They should be able to compare literary attributes of two or more texts and make judgments about how each author presents events. Readers show understanding of vocabulary and figurative language. They can develop an objective summary of a text and produce an informational text that applies a common theme or central idea culled from multiple texts to a current societal issue.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), twelfth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to analyze the specific results of a multistep procedure based on explanations in the text, explain how specific individuals, ideas, and/or events interact and develop over the course of a text, and analyze how the text structures information or ideas into categories or hierarchies. Readers should be able to compare and contrast findings presented in a text to those from other sources and show understanding of general academic and domain-specific vocabulary, key terms, and symbols. Readers should be able to generate an alternative procedure or experiment based on knowledge acquired from information gained from reading texts.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, twelfth-grade readers performing at the *NAEP Basic* level should be able to explain how specific individuals, ideas, and/or events interact and develop over the course of a text, determine and interpret an author's point of view or purpose, and distinguish between fact, opinion, and reasoned judgment in a text. Readers should be able to show understanding of general academic and domain-specific vocabulary and of figurative language and be able to develop an objective summary of a text by paraphrasing its complex concepts and information. They can integrate information from multiple sources to produce a new text of informational or argumentative purpose.

NAEP Proficient

Twelfth-grade students performing at the *NAEP Proficient* level should be able to make more complex inferences and interpretations, form explanations and generalizations, generate alternatives, and apply new ideas acquired through reading to a new problem or context. Students should be able to use text-based evidence to support arguments and conclusions.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, twelfth-grade readers performing at the *NAEP Proficient* level should be able to analyze how two or more themes or central ideas interact and build on one another to produce a complex account over the course of the text. Readers should be able to analyze how text structure contributes to meaning and style. They can analyze how word choice impacts a text's meaning and tone. Readers can develop a detailed objective summary of a text and produce a new text of literary purpose based on an archetypal conflict discovered in the reading texts.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), twelfth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to analyze an author's point of view or purpose, including in providing an explanation, describing a procedure, or discussing an experiment, identifying important issues that remain unresolved. Readers should be able to integrate and evaluate multiple sources of information presented in diverse media or formats (visually or in words) in order to address a question or solve a problem. Readers can produce a new argumentative or informative text that synthesizes information from a range of sources to demonstrate a coherent understanding of a process, phenomenon, or concept.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, twelfth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to analyze how the central ideas interact and build on one another to produce a complex account. They should be able to analyze the themes, purposes, and rhetorical features of foundational U.S. documents and evaluate the effectiveness of the structure in the text's exposition or argument. They should be able to develop a detailed objective summary of a text. Readers can evaluate multiple sources of information presented in different media or formats (visually or in words) in order to produce an argumentative text with evidence to structure and support a judgment.

NAEP Advanced

Twelfth-grade students performing at the *NAEP Advanced* level should be able to make complex inferences and to support their interpretations, conclusions, and their judgments based upon evidence within and across texts. Students should be able to use an understanding of legal and ethical principles to develop a text or presentation on a matter of social debate.

When engaged in reading literary texts such as fiction, drama, film, poetry, and literary nonfiction, twelfth-grade readers performing at the *NAEP Advanced* level should be able to use textual evidence as support to analyze and evaluate multiple interpretations of text (e.g., multimedia versions of a text) to the source text. Readers can use acquired knowledge to produce an informational text analyzing how elements of an era's poetry (e.g., Romanticism's celebration of nature; rejection of industrialization) are evidenced in the work of one or more poets.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), twelfth-grade readers performing at the *NAEP Advanced* level should be able to delineate and evaluate the argument, claims, and reasoning in a text, and evaluate the hypotheses, data, analysis, and conclusions in a text. They should be able to explain how style and content contribute to the power, persuasiveness, or beauty of the text. Readers can produce a new argumentative or informative text that utilizes an understanding of legal and ethical principles to address a scientific matter of debate (e.g., uses of genetic databases).

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, twelfth-grade readers performing at the *NAEP Advanced* level should be able to delineate and evaluate argument, claims, and reasoning in a text. They should be able to explain how style and content contribute to the power, persuasiveness, or beauty of the text. Readers can produce a new argumentative or informative text that utilizes an understanding of legal and ethical principles to address a societal matter of debate (e.g., indigenous peoples' land rights).

(NCES, 2018).

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher, 33*(1), 4-14, doi.org/10.3102/0013189X033001004

Afflerbach, P., Biancarosa, G., Hurt, M., & Pearson, P. D. (2020).Teaching reading for understanding: Synthesis and reflections on the curriculum and instruction portfolio. In P. D. Pearson, A. S. Palincsar, G. Biancarosa, & A. Berman (Eds.), (2020). *Reaping the rewards of the Reading for Understanding Initiative* (pp. 215-250). Washington, DC: National Academy of Education.

Afflerbach, P., Hurt, M., & Cho, B. (2020). Reading comprehension strategy instruction. In D. L. Dinsmore, L. A. Fryer, & M. M. Parkinson (Eds.), *Handbook of strategies and strategic processing* (pp. 99-118). New York: Routledge.

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2007). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*(5), 364-373, doi.org/10.1598/RT.61.5.1

Alvermann, D. E. & Wilson, A. A. (2011). Comprehension strategy instruction for multimodal texts in science. *Theory Into Practice, 50*(2), 116-124.

Aguilar, G., Uccelli, P., & Galloway, E. P. (2020). Toward biliteracy: Unpacking the contributions of mid-adolescent dual language learners' Spanish and English academic language skills to English reading comprehension. *TESOL Quarterly* (early view)*,* DOI: 10.1002/tesq.570

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology, 32*(4), 665-683, doi.org/10.1111/j.1559-1816.2002.tb00236

Alexander, P. A., & Grossnickle, E. M. (2016). Positioning interest and curiosity within a model of academic development. In K. Wentzel & D. Miele (Eds.), *Handbook of motivation at school* (2nd ed., pp. 188–208). New York: Routledge.

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal, 31*(2), 313-337, doi.org/10.3102/00028312031002313

Ambruster, B. B. (1984). Learning from content area textbooks: The problem of "inconsiderate text." In G. Duffy, L. Roehler, & J. Mason (Eds.), Comprehension instruction (pp. 202-217). New York, NY: Longman.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, P. Mosenthal (Eds.), *Handbook of Reading Research* (Vol. I) (pp. 255-291). New York: Routledge.

Anderson, R. C. (2019). Role of the reader's schema in comprehension, learning and memory. In D.E. Alvermann, N.J. Unrau, & R.B. Ruddell (Eds.), *Theoretical models and processes of literacy* (pp. 136-145). New York: Routledge.

August, D., & Shanahan, T. (Eds.) (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 583-596). Mahwah, NJ: Lawrence Erlbaum Associates.

Bakken, J. P., & Whedon, C. K. (2002). Teaching text structure to improve reading comprehension. *Intervention in School and Clinic, 37*(4), 229-233.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Upper Saddle River, NJ: Prentiss Hall.

Bandura, A. (1993). "Perceived self-efficacy in cognitive development and functioning." Educational Psychologist 28(2): 117-148.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: Information Age Publishing.

Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction, 30*(1), 39-85, DOI: 10.1080/07370008.2011.635495

Barzilai, S., & Zohar, A. (2016). Epistemic (meta)cognition: Ways of thinking About knowledge and knowing. In J. A. Greene, W. A. Sandoval & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 409-424). New York: Routledge.

Beach, R. (2000). Critical issues: Reading and responding to literature at the level of activity. *Journal of Literacy Research, 32*(2), 237-251, doi.org/10862960009548075

Beach, R., & Castek, J. (2015). Use of apps and devices for fostering mobile learning of literacy practices. In B. Guzzetti & M. Lesley (Eds.), *Handbook of research on the societal impact of digital media* (pp. 343–370). Hershey, PA: IGI Global.

Bergner, Y. & von Davier, A. A. (2018). Process data in NAEP: Past, present, and future. Journal of Educational and Behavioral Statistics, 20(10), 1-27. DOI: 10.3102/1076998618784700

Bråten, I., Braasch, J.L.G., Salmerón, L. (2020). Reading multiple and non-traditional texts: New opportunities and new challenges. In E. B. Moje, P. Afflerbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of reading research* (Vol. V) (pp. 79-98). New York: Routledge.

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48-70, doi.org/10.1080/00461520.2011.53847

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.

Bunch, G. C., Walqui, A. & Pearson, P. D. (2014). Complex Text and New Common Standards in the United States: Pedagogical Implications for English Learners. *TESOL Quarterly, 48*(3), 533-559.

Business Roundtable. (2017). *Work in progress: How CEOs are helping close America's skills gaps.* Retrieved May 2018, from https://www.businessroundtable.org/skills

Calfee, R. C., & Miller, R. G. (2005). Comprehending through composing: Reflections on reading assessment strategies. In S. Paris & S. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 215-233). Mahwah, NJ: Lawrence Erlbaum Associates.

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2019). Malleability, plasticity, and individuality: How children learn and develop in context. *Applied Developmental Science, 23*(4), 307-337, DOI: 10.1080/10888691.2017.1398649

Carroll, J. M., & Fox, A. (2017). Reading self-efficacy predicts word reading but not comprehension in both girls and boys. *Frontiers in Psychology, 7*(25), 1-9, DOI: 10.3389/fpsyg.2016.02056

CAST. (2020). About Universal Design for Learning. *CAST.* http://www.cast.org/our-work/about-udl.html#.XrWphi-z3kI

Cazden , C.B. (2002). Classroom discourse: The language of teaching and learning (2nd ed.). Portsmouth, NH : Heinemann .

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann.

Cervetti, G. N. (2020). The nature and development of reading for understanding. In P. D. Pearson, A. S. Palincsar, G. Biancarosa, & A. Berman (Eds.), (2020). *Reaping the rewards of the Reading for Understanding Initiative* (pp. 39-64). Washington, DC: National Academy of Education.

Cervetti, G.N., & Wright, T.S. (2020). The role of knowledge in understanding and learning from text. In E. B. Moje, P. Afflerbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of reading research* (Vol. V) (pp. 237-260). New York: Routledge.

Chen, C-M, & Chen, F-Y. (2014). Enhancing digital reading performance with a collaborative reading annotation system. *Computers & Education, 77,* 67-81.

Cho, B. (2014). Competent adolescent readers' use of internet reading strategies: A think-aloud study. *Cognition and Instruction, 32*(3), 253-289, DOI: 10.1080/07370008.2014.918133

Cho, B.-Y., & Afflerbach, P. (2017). An evolving perspective of constructively responsive reading comprehension strategies in multilayered digital text environments. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (2nd ed., pp. 109–134). New York, NY: Guilford.

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. doi :10.1177/1086296X11421979

Coiro, J. (2020). Toward a multi-faceted heuristic of digital reading to inform assessment, research, practice, and policy. *Reading Research Quarterly* (early view), doi.10.1002/rrq.302

Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, 42(2), 214–257. doi:10.1598/RRQ.42.2.2

Coiro, J., Knobel, M., Lankshear, C., & Leu, D. J. (2014). Handbook of research on new literacies. Routledge.

Coiro, J., Sparks, J., Kiili, C., Castek, J., Lee, C., & Holland, B. (2019). Capturing dimensions of collaborative online inquiry and social deliberation with multiple-source inquiry tasks in face-to-face and remote contexts. *Literacy Research: Theory, Method, and Practice.* Online first version available at https://doi.org/10.1177/2381336919870285

Cole, M. (1996). *Cultural psychology: A once and future discipline.* Cambridge, MA: Harvard University Press.

Cote, N., Goldman, S.R. & Saul, E.U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes, 25*, 1-53, DOI: 10.1080/01638539809545019

Creer. A. (2018). Introducing everyday ‘digital literary practices’ into the classroom: An analysis of multi-layered media, modes and their affordances. *Journal of New Approaches in Educational Research, 7*(2), 131-139.

Dalton, B., & Proctor, C. P. (2008). The changing landscape of text and comprehension in the age of new literacies. In J. Coiro, M. Knobel, C. Lankshear, & D. Leu (Eds.), *Handbook of research on new literacies* (pp. 297-324). Mahwah, NJ: Lawrence Erlbaum.

Daniel, S. M., & Pacheco, M. B. (2016). Translanguaging practices and perspectives of four multilingual teens. *Journal of Adolescent & Adult Literacy, 59*(6), 653-663, doi.org/10.1002/jaal.500

Dinsmore, D., L., & Parkinson, M. M. (2013). What are confidence judgments made of? Students’ explanations for their confidence ratings and what that means for calibration. *Learning and Instruction, 24*, 4-14, doi.org/10.1016/jlearninstruc.2012.06.001

Dobler, E., & Azwell, T. (2007). Real world reading: Making sense of the texts that matter in our everyday lives. Kansas Career & Technical Education Resource Center. https://www.ksde.org/Portals/0/CSAS/CSAS%20Home/CTE%20Home/Instructor_Resources/FinalRealWorldReading.pdf

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading comprehension* (3rd ed.) (pp. 205-242). Newark, DE: International Reading Association.

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.

Eccles, J., et al. (2005). Ability self-perceptions and subjective task values in adolescents and children. Moore, K.

Educational Testing Service (2019). NAEP Reading Special Study: Scenario-Based Task and Discrete Task Report. Deliverable in Response to ID Task 9.2.1. Note: This report is not available to the general public because it contains secure information.

Educational Testing Service. (January 22, 2019). NAEP reading special study: Scenario-based task (SBT) and discrete (DI) task report. 2018 special study grades 4, 8, and 12.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process. *Educational Research Review, 1*(1), 3-14, DOI: 10.1016/jedurev.2005.11.001

Enciso, P., Volz, A., Price-Dennis, D., & Durriyah, T. (2010). Story club and configurations of literary insights among immigrant and non-immigrant youth. In R. Jiménez, D. Rowe, V. Risko, & M. Hundley (Eds.), *59th Yearbook of the National Reading Conference* (pp. 354–366). Oak Creek, WI: National Reading Conference.

Faircloth, B., Hamm, J. (2005) Sense of Belonging Among High School Students Representing 4 Ethnic Groups. *Journal of Youth and Adolescence*, Vol. 34, No. 4, 293–309 DOI: 10.1007/s10964-005-5752-7

Fang, Z., Shleppegrell, M., & Cox, B. (2006). Understanding the language demands of schooling: Nouns in academic registers. Journal of Literacy Research, 38(3), 247-273.

Farmer, T. W., Lines, M. M., & Hamm, J. V. (2011). Revealing the invisible hand: The role of teachers in children’s peer experiences. *Journal of Applied Developmental Psychology*, 32, 247–256. doi:10.1016/j.appdev.2011.04.006

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.

Fisher, D., Frey, N., & Lapp, D. (2012). *Text complexity: Raising rigor in reading*. Newark, DE: International Reading Association.

Fitzgerald, M., Higgs, J., & Palincsar, A. (2020). *New media, new literacies: Implications for reading for understanding.* Washington, DC: National Academy of Education.

Fletcher, J. (2015). Performing digital literature. *Caracteres: Estudios Culturales y Críticos de la Esfera Digital, 4*(2), 18-42.

Francis, D. J., Lesaux, N. K., & August, D. (2006). Language of instruction. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (p. 365–413). New York: Lawrence Erlbaum Associates.

Frankel, K. K., Becker, B. L. C., Rowe, M. W., & Pearson, P. D. (2016). From "What is reading?" to what is literacy? *The Journal of Education, 196*(3), 7-17, doi.org/10.1177/002205741619600303

García, G. E, & Godina, H. (2017). A window into bilingual reading: The bilingual reading practices of fourth-grade, Mexican-American children who are emergent bilinguals. *Journal of Literacy Research, 49*(2), 273–301, doi.org/10.1177/1086296X17703727

García, G.E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. Reading Research Quarterly, 26(4), 371-392

García, G.E., Saco, L.J., & Guerrero-Arias, B.E. (2020). Cognate instruction and bilingual students' improved literacy performance. *The Reading Teacher, 73*(5), 617-625.

García, O. (2009). *Bilingual education in the 221st century: A global perspective*. Malden, MA: Basil/Blackwell.

Gee, J. P. (2001). Reading as situated language: A sociocognitive perspective. *Journal of Adolescent and Adult Literacy, 44*(8), 714-725.

González, N., Moll, L. C., & Amanti, C. (Eds.). (2005). Funds of Knowledge: Theorizing practices in households, communities, and classrooms. Mahwah, NJ: Erlbaum.

Goldman, S. R., & Lee, C. D. (2014). Text complexity: State of the art and the conundrums it raises. *The Elementary School Journal, 115*(2), 290-300, doi/10.1086/678298

Goldman, S. R., et al. (2016). "Disciplinary Literacies and Learning to Read for Understanding: A Conceptual Framework for Disciplinary Literacy." Educational Psychologist: 1-28.

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.

Goldman, S.R., 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. DOI: 10.1080/00461520.2016.1168741

Greenleaf, C. Schoenbach, R., Cziko, C., and Mueller, F. (*2001*) Apprenticing Adolescent Readers to Academic Literacy. *Harvard Educational Review*: April 2001, Vol. 71, No. 1, pp. 79-130.

Gribben, M.; Patelis, T. & Schultz, S. (2020) *Review of the Impact of Student Choice in Assessment,* Report 2020:037. Alexandria, VA: HumRRO.

Guan, S. S. A., Nash, A., & Orellana, M. F. (2016). Cultural and social processes of language brokering among Arab, Asian, and Latin immigrants. *Journal of Multilingual and Multicultural Development*, *37*(2), 150-166.

Guthrie, J. T. and A. Wigfield (2000). Engagement and motivation in reading. Handbook of reading research. M. Kamil, P. Mosenthal, P. D. Pearson and R. Barr. Mahwah, NJ, Lawrence Erlbaum. III: 403-422.

Guthrie, J. T., & Klauda, S. (2014). Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Reading Research Quarterly, 49*(4), 387-416, DOI: 10.1002/rrq.81

Guthrie, J.T. & Klauda, S. (2015). Engagement and motivational processes in reading. In P. Afflerbach, Handbook of Individual Differences in Reading. (pp. 41-54). Routledge Publishers, New York.

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

Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading, 3*, 231–256, doi..org/10.1207/s1532799xssr0303_3

Guthrie, J. T., Wigfield, A., Von Secker, C. (2000) Effects of integrated instruction on motivation and strategy use in reading. Journal of Educational Psychology, 92(2), 331-341.

Gutiérrez, K. D., Morales, P. Z., & Martinez, D. C. (2009). Re-mediating literacy: Culture, difference, and learning for students from nondominant communities. *Review of Research in Education, 33,* 212-245, DOI: 10.3102/0091732X08328267

Hain, B.A. & Piper, C. (2016). PARCC as a case study in understanding the design of large-scale assessment in the era of Common Core State Standards. In H. Jiao & R.W. Lissetz (Eds). The next generation of testing: Common Core Standards, Smarter-Balanced, PARCC, and the nationwide testing movement (pp. 29-48). Information Age Publishing.

Hall, L. A. (2016). The role of identity in reading comprehension development. *Reading & Writing Quarterly*, *32*(1), 56-80.

Harris, Y. R., & Schroeder, V. M. (2013). Language deficits or differences: What we know about African American vernacular English in the 21st century. *International Education Studies, 6*(4), 194-204. DOI: 10.5539/ies.v6n4p194

Hasbrouck, J. & Tindal, G. A., (2006). Oral reading fluency norms:A valuable assessment tool for reading teachers. The Reading Teacher, 59(7), 636-644.

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge: Cambridge University Press.

Heath, S. B. (2012). “So what’s it about? Your book, I mean?” *Journal of Adolescent & Adult Literacy, 56*(4), 266-270, doi-org.libezproxy2.syr.edu/10.1002/JAAL.00137

Heath, S.B. (2012). Words at work and play: Three decades in family and community life. Cambridge, UK: Cambridge University Press.

Ho, A. (2017). Advancing Educational Research and Student Privacy in the “Big Data” Era. Washington, DC: National Academy of Education.

Hofer, P. J. (1997). Multicultural assessment of IQ. *Archives of Clinical Neuropsychology*, *12*(4), 335.

Hruby, G. G., & Goswami, U. (2019). Educational neuroscience for reading researchers. In D.E. Alvermann, N.J. Unrau, & R.B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed.) (pp. 252-278). New York: Routledge.

Hruby, G. G., Goswami, U., Frederiksen, C. H., & Perfetti, C. A. (2011). Neuroscience and reading: A review for reading education researchers. *Reading Research Quarterly, 46*(2), 156–172, dx.doi.org/10.1598/RRQ.46.2.4

Ilter, I. (2019). Efficacy of note-taking skills instruction supported by self-monitoring on the reading comprehension. *Education and Science, 44*(198), 229-253, DOI: 10.15390/EB.2019.7862

Immordino-Yang, M. H., & Gotlieb, R. (2017). Embodied brains, social minds, cultural meaning: Integrating neuroscientific and educational research on social-affective development. *American Educational Research Journal*, *54*(1_suppl), 344S-367S, doi.org/10.3102/0002831216669780

Jiménez, R. T., García, G. E., & Pearson, P. D. (1995). Three children, two languages, and strategic reading: Case studies in bilingual/monolingual reading. *American Educational Research Journal, 32*(1), 67–97, doi.org/10.2307/1163214

Jiménez, R. T., García, G. E., & Pearson, P. D. (1996). The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly, 31*(1), 90-112, doi.org/10.1598/RRQ.31.1.

Johnson, S. B., Riis, J. L., & Noble, K. G. (2016). State of the art review: Poverty and the developing brain. *Pediatrics, 137*(4), e20153075, DOI: 10.1542/peds.2015-3075

Kendeou, P., & O' Brien, E. J. (2016). Prior Knowledge: Acquisition and revision. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Text and context* (pp. 151-163). New York: Routledge.

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.

Kieffer, M. J., & Thompson, K. D. (2018). Hidden Progress of Multilingual Students on NAEP. *Educational Researcher, 47,* pp. 391–398. doi: 10.3102/0013189X18777740

Killi, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating internet sources: From versatile evaluators to uncritical readers. *Journal of Computing Research, 39*(1), 75-95, doi.org/10.2190/EC.39.1.e

Kim, Y.-S. G., Petscher, Y., Wanzek, J., & Al Otaiba, S. (2018). Relations between reading and writing: A longitudinal examination from Grades 3 to 6. *Reading and Writing: An Interdisciplinary Journal, 31*(7), 1591–1618, doi.org/10.1007/s11145-018-9855-4

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, Cambridge University Press.

Kintsch, W. (2019). Revisiting the construction-integration model of text comprehension and its implications for instruction. In D.E. Alvermann, N.J. Unrau, & R.B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed.) (pp. 178-203). New York: Routledge.

Kress, G. (2013). Recognizing learning. In I. deSaint-Georges & J. J. Weber (Eds.), Multilingualism and multimodality (pp. 119–140). Rotterdam, ND: Sense Publishers.

Kress, G., & van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication. London, UK: Arnold.

Kress, G., Jewitt, C., & Tsatsarelis, C. (2000). Knowledge, identity, pedagogy: pedagogic discourse and the representational environments of education in late modernity. *Linguistics and Education, 11*(1), 7-30, doi.org/10.1016/S0898-5898(99)00015-7

Kuhn, D., & Weinstock, M. (2002). What Is epistemological thinking and why does it matter? In B. Hofer, & P. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 121-144). New York: Routledge.

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.

Latini, N., Bråten, I., Anmarkrud, Ø., & Salmerón, L. (2019). Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple text context. *Contemporary Educational Psychology*, *59*, 101797, doi.org/10.1016/j.cedpsych.2019.101797.

Lee, C. D. (1993). *Signifying as a scaffold for literary interpretation: The pedagogical implications of an African American discourse genre*. Urbana, IL, National Council of Teachers of English.

Lee, C. D. (1997). Bridging home and school literacies: A model of culturally responsive teaching. In J. Flood, S. B. Heath, & D. Lapp (Eds.), *A handbook for literacy educators: Research on teaching the communicative and visual arts* (pp. 330–341). New York, NY: McMillan.

Lee, C. D. (2005). Culture and language: Bi-dialectical issues in literacy. In P. L. Anders & J. Flood (Eds.), *The literacy development of students in urban schools*. Newark, DE: International Reading Association.

Lee, C. D. (2006). Every good-bye ain't gone: Analyzing the cultural underpinnings of classroom talk. I*nternational Journal of Qualitative Studies in Education, 19*(3), 305–327.

Lee, C. D. (2007). *Culture, literacy and learning: Taking bloom in the midst of the whirlwind*. NY, Teachers College Press.

Lee, C. D. (2014). A voyeuristic view of possibilities and threats: Neurosciences and education. *Human Development*, *57*(1), 1-4, doi.org/10.1159/000360166.

Lee, C. D. (2016). Examining conceptions of how people learn over the decades Through AERA presidential addresses: Diversity and equity as persistent conundrums. *Educational Researcher, 45*(2), 73-82, doi.org/10.3102/0013189X16639045

Lee, C. D. (2016). Influences of the experience of race as a lens for understanding variation in displays of competence in reading comprehension. In P. Afflerbach, (Ed.), *Handbook of individual differences in reading: Reader, text, and context* . NY, Routledge: 286-304.

Lee, C. D. (2020). Social and cultural diversity as lens for understanding student learning and the development of reading. In E. B. Moje, P. Afflerbach, N. K. Lesaux, & P. Enciso (Eds.), *Handbook of reading research, Volume V* (37-56). New York: Routledge.

Lee, C. D. and A. Spratley (2009). *Reading in the disciplines and the challenges of adolescent literacy.* NY, Carnegie Foundation of New York.

Lee, C. D. (in press). Social and Cultural Diversity as Lens for Understanding Student Learning and the Development of Reading Comprehension. In Handbook of Reading Research,Volume V. Moje, E.B., Afflerbach, P., Encisco, P. & Lesau, N. NY, Routledge.

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.

Lippman, L. NY, Springer: 237-270.

List A., & Alexander, P. A. (2018). Cold and warm perspectives on the Cognitive Affective Engagement Model of Multiple Source Use. In J. L. Braasch, I. Bråten, & M. T., McCrudden (Eds.), *Handbook of multiple source use* (pp. 46-66). New York: Routledge.

Manderino, M. (2012). Disciplinary literacy in new literacies environments: Expanding the intersections of literate practice for adolescents. In P. Dunston, L. Gambrell, K. Headley, S. Fullerton, & P. Stecker (Eds.), *Sixty-first yearbook of the literacy research association* (pp. 69–83). Oak Creek, WI: Literacy Research Association.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development, 50*(3), 643-655, DOI: 10.2307/1128929

Marzano, R. J., Brandt, R. S., Hughes, C. S., Jones, B. F., Presseisen, B. Z., Rankin, S. C., & Suhor, C., (XXX). Dimensions of thinking: A framework for curriculum and instruction. Alexandria, VA: Association for Supervision and Curriculum Development.

Massey, D. D. (2009). Self-regulated comprehension. In S. E. Israel & G. D. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 389-399). New York: Routledge.

McNamara, D. S. and W. Kintsch (1996). "Learning from text: Effects of prior knowledge and text coherence." Discourse Processes 22: 247-287.

Measured Progress/ETS Collaborative. (2012). *Smarter Balanced Assessment Consortium: Technology-enhanced item guidelines*. https://www.measuredprogress.org/wp-content/uploads/2015/08/SBAC-Technology-Enhanced-Items-Guidelines.pdf

Meyer, B. J. F. (1975). Identification of the structure of prose and its implications for the study of reading and memory. *Journal of Reading Behavior, 7,* 7-47.

Michalsky, T., Mevarech, Z. R., Haibi, L. (2009). Elementary school children reading scientific texts: Effects of metacognitive instruction. *Journal of Educational Research, 102*(5), 363-376, DOI: 10.3200/JOER.102.5.363.376

Mislevy, R. J. (2016). How development in psychology and technology challenge validity argumentation. *Journal of Educational Measurement, 53(3),* 265-292.

Moje, E. B. (2015). Doing and teaching disciplinary literacy with adolescent learners: A social and cultural enterprise. *Harvard Educational Review, 85*(2), 254–278. doi:10.17763/0017-8055.85.2.254

Moje, E. B. (2009). A call for new research on new and multi-literacies. *Research in the Teaching of English, 43*(4), 348-362.

Moje, E. B. (2000). "To be part of the story": The literacy practices of gangsta adolescents. *Teachers College Record, 102*(3), 651–690.

Moll, L. C. (2014). L. S. *Vygotsky and education.* New York, NY: Routledge.

Moll, L. C., & Gonzalez, N. (1994). Lessons from research with language-minority children. *Journal Of Reading Behavior, 26*(4), 439–456.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice*, *31*(2), 132-141.

Mullis, I. V. S., & Martin, M. O. (2019). *Chapter 1 PIRLS 2021 Reading assessment framework*. Boston: TIMSS & PIRLS International Study Center.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2019). *PIRLS 2021 assessment frameworks.* Boston, MA: International Association for the Evaluation of Education Achievement, & TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. http://pirls2021.org/wp-content/uploads/sites/2/2019/04/P21_Frameworks.pdf

Mullis, I.V.S., Martin, M.O., Foy, P., & Drucker, K.T. (2012). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47* (1), 91–108. doi: 10.1002/RRQ.011

Nasir, N. S., & Hand, V. M. (2006). Exploring sociocultural perspectives on race, culture, and learning. *Review of Educational Research, 76*(4), 449-475, doi.org/10.3102/0034655430/6004449

National Academies of Sciences, Engineering, and Medicine. (2018). *How People Learn II: Learners, Contexts, and Cultures*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24783.

National Academy of Sciences. (2018). *The science of science communication III: Inspiring novel collaborations and building capacity: Proceedings of a colloquium.* Washington, DC: The National Academies Press, doi.org/10.17226/24958

National Assessment Governing Board. (2017). *Reading framework for the 2017 National Assessment of Educational Progress.* Washington, DC: National Assessment Governing Board and U.S. Department of Education. https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/reading/2017-reading-framework.pdf

National Assessment Governing Board. (2017). *Reading framework for the 2017 National Assessment of Educational Progress.* Washington, DC: National Assessment Governing Board and U.S. Department of Education. https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/reading/2017-reading-framework.pdf

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.

National Center for Education Statistics. (2018). Going digital: NAEP assessments for the future. *National Assessment of Educational Progress.* Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/nationsreportcard/subject/dba/pdfs/2018_dba_brochure.pdf

National Center for Education Statistics. (2020). Digitally based assessments. *National Assessment of Educational Progress.* Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/nationsreportcard/dba/

National Research Council. (2010). *Language Diversity, School Learning, and Closing Achievement Gaps: A Workshop Summary.* Washington, DC: The National Academies Press.
National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits.* Washington, DC: The National Academies Press.
New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*, 60–92.
Newell, G. E., Beach, R., Smith, J., VanDerHeide, J., Kuhn, D., & Andriessen, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly, 46*(3), 273–304.
O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. American Educational Research Journal, 44, 161–196.
O'Reilly, T., Wang, Z., & Sabatini, J. (2019). How much knowledge is too little? When a lack of knowledge becomes a barrier to comprehension. Psychological Science, 30(9), 1344–1351.
OECD. (2019). *PISA 2018 assessment and analytical framework.* Paris: OECD Publishing. https://doi.org/10.1787/b25efab8-en
Organisation for Economic Co-operation and Development. (2019). *Education at a glance 2019: OECD indicators. Paris, FR: OECD.*
Osher, D., Cantor, P., Berg, J., Steyer, L. & Rose, T. (2020). Drivers of human development: How relationships and context shape learning and development. *Applied Developmental Science, 24*(1), 6-36, DOI: 10.1080/10888691.2017.1398650
Pacheco, M. (2015). Bilingualism-as-participation: Examining adolescents' bi(multi)lingual literacies across out-of-school and online contexts. In D. Molle, E. Sato, T. Boals, & C. Hedgespeth (Eds.), *Multilingual learners and academic literacies: Sociocultural contexts of literacy development in adolescents* (p. 135- 165). New York: Routledge.
Pacheco, M. (2018). Learning and becoming writers: Meaning, identity, and epistemology in a newsroom community of practice. *Mind, Culture, and Activity*, 25(2), 105-124.
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.
Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578, doi.org/10.3102/0034654306600543
Paris, D. (2009). "They're in my culture, they speak the same way": African American Language in multiethnic high schools. *Harvard Educational Review, 79*(3), 428–448.
Paris, S. G., Wasik, B. A., & Turner, J. C. (1991). The Development of Strategies of Readers. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research* (Vol. 2, pp. 609-640). Mahwah, NJ: Lawrence Erlbaum Associates.
Pearson, P. D. and E. H. Hiebert (2014). "The State of the Field: Qualitative Analyses of Text Complexity." Elementary School Journal.
Pearson, P. D., & Camparell, K. (1981). Comprehension of text structures. In J. Guthrie (Ed.), *Comprehension and teaching* (pp. 27-54). Newark DE: International Reading Association.

Pearson, P. D., & Cervetti, G. N. (2015). Fifty years of reading comprehension theory and practice. In P. D. Pearson, E. H. Hiebert,& N. K. Duke (Eds.), *Research-based practices for teaching Common Core literacy* (pp. 1-24). New York: Teachers College Press.

Pearson, P. D., Hansen, J., & Gordon (1979). The effect of background knowledge on young children's comprehension of explicit and implicit information. Journal of Reading Behavior, 11, 201-210.

Pearson, P. D., & Hiebert, E. (2014). The state of the field. *The Elementary School Journal, 115*(2), 161-183.

Pearson, P. D., Palincsar, A. S., Biancarosa, G., & Berman, A. (2020). *Reaping the rewards of the Reading for Understanding Initiative.* Washington, DC: National Academy of Education.

Pearson, P. D., Valencia, S., & Wixson, K. (2014). Complicating the world of reading assessment: Better assessments for better teaching. *Theory into Practice, 53*(3), 236-246. DOI: 10.1080/00405841.2014.916958

Pérez, C. C. (2017). *The first rule of punk.* New York: Penguin Random House.

Peura, P., Viholainen, H. J. K., Aro, T. I., Raikkonen, E. M., Usher, E. L., Sorvo, R. M. A., Klassen, R. M. & Aro**,** M. T; (2019). *Specificity of reading self-efficacy among primary school children. Journal of Experimental Education, 87,* 496-516, doi.org/10.1080/00220973.2018.1527279

Phillips Galloway, E., Uccelli, P., Aguilar, G. & Barr, C. (2020) Exploring the cross-linguistic contribution of Spanish and English academic language skills to English text comprehension for middle-grade dual language learners. AERA Open, 1, 1-20. doi: 10.1177/2332858419892575

Pintrich, P. R. and B. Schrauben (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. Student perceptions in the classroom: Causes and consequences D. Schunk and J. Meece. Hillsdale, NJ, Erlbaum: 149-183.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Proctor, C. P., Silverman, R.D., Harring, J. R., Jones, R.L., & Hartranft, A. M. (2020) Teaching Bilingual Learners: Effects of a Language-Based Reading Intervention on Academic Language and Reading Comprehension in Grades 4 and 5 *Reading Research Quarterly 55* (1), 95-122 doi: 10.1002/rrq.258

Purcell-Gates, V., Duke, N.K., & Stouffer, J. (2016). Teaching literacy: Reading. In D.H. Gitomer & C.A. Bell (Eds.), *The AERA handbook of research on teaching* (5th ed.) (pp. 1217– 1267). Washington, DC: American Educational Research Association.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension.* Santa Monica, CA: RAND. https://www.rand.org/pubs/monograph_reports/MR1465.html

Rose, DH, & Meyer, A (2002) *Teaching every student in the digital age: Universal design for learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Roth, W. M., & Lee, Y.-J. (2007). "Vygotsky's neglected legacy": Cultural-historical ctivity theory. *Review of Educational Research*, *77*(2), 186–232, doi.org/10.3102/0034654306298273

Sabatini, J. P., et al. (2014). "Integrating scenario-based and component reading skill measures to understand the reading behavior of struggling readers." Learning Disabilities Research & Practice 29(1): 36-43.

Sabatini, J. P., O'Reilly, T., Halderman, L. K., & Bruce, K. (2014). Integrating scenario-based and component reading skill measures to understand the reading behavior of struggling readers. *Learning Disabilities Research & Practice, 29*(1), 36–43.
Sabatini, J., et al. (2018). Scenario-based assessment of multiple source use. Handbook of Multiple Source Use, Routledge: 447-465.
Sabatini, J., O'Reilly, T., Weeks, J., & Wang, Z. (2020). Engineering a twenty-first century reading comprehension assessment system utilizing scenario-based assessment techniques. *International Journal of Testing*, *20*(1), 1-23. https://doi.org/10.1080/15305058.2018.1551224
Sanchez, I. G., & Orellana, M. F. (2006). The construction of moral and social identity in immigrant children's narratives-in-translation. *Linguistics And Education: An International Research Journal, 17*(3), 209–239.
Schallert, D. L. (2002). Schema theory. In B. J. Guzzetti (Ed.), *Literacy in America: An encyclopedia of history, theory, and practice* (pp. 556-558). Santa Barbara, CA: ABC-Clio.
Schleppegrell, M.J. (2004). *The language of schooling: A functional linguistics perspective.* Mahwah, NJ: Erlbaum.
Schultz, K. (2002). Looking across space and time: Reconceptualizing literacy learning in and out of school. Research in the Teaching of English, 36(3), 356–390.
Scribner, S., & Cole, M. (1978). Literacy without schooling: Testing for intellectual effects. *Harvard Educational Review, 48*(4), 448–461.
Scribner, S. and Cole, M (1981). *The Psychology of Literacy*. Cambridge, Mass.: Harvard University Press.
Serafini, F. (2004). Images of reading and the reader. *The Reading Teacher, 57*(7), 610-617.
Shin, R. Q., Daly, B. P., & Vera, E. M. (2007). The relationships of peer norms, ethnic identity, and peer support to school engagement in urban youth. *Professional School Counseling*, 10, 379– 388.
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.
Singer, L. M., & Alexander, P. A. (2017). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research, 87*(6), 1007-1041, doi.org/10.3102/0034654317722961
Skerrett, A. (2011). English teachers' racial literacy knowledge and practice. *Race, Ethnicity, and Education*, 14(3), 313-330.
Skerrett, A. (2012). Social and Cultural Differences in Reading Development 36 Skerrett, A. "We hatched in this class": Repositioning of identity in and beyond a reading classroom. *The High School Journal*, 95(3), 62-75.
Skerrett, A. (2015). *Teaching transnational youth: Literacy and education in a changing world*. New York, NY: Teachers College Press.
Skerrett, A. (2020). Social and cultural differences in reading development: Instructional approaches, learning gains, and challenges. In E. B. Moje, P. Afflerbach, P. Enciso, and N. Lesaux. (Eds.). *Handbook of reading research* (Vol. V) (pp. 328-344). New York, NY: Routledge.
Skerrett, A. (in press) *Handbook of Reading Research*. Social and Cultural Differences in Reading Development: Instructional Processes, Learning Gains, and Challenges

Smagorinsky, P. (2001). If meaning is constructed, what is it made from?: Toward a cultural theory of reading. *Review of Educational Research, 71*(2), 133–169.
Snow, C. E., & Uccelli, P .(2009). The challenge of academic language. In Olson, D. R., & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp.112-133). Cambridge: Cambridge University Press.
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.
Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *38*(5), 553-573.
Sparks, J. R., & Deane, P. (2014). *Cognitively-based assessment of research and inquiry skills: Defining a key practice in the English language arts*. Manuscript in preparation.
Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407, doi.org/10.1598/RRQ.21.4.1
Stanovich, K. E. (2013). Why humans are (sometimes) less rational than other animals: Cognitive complexity and the axioms of rational choice. *Thinking and Reasoning, 19*(1), 1-26, DOI: 10.1080/13546783.2012.713178
Stenner, A.J. (1996). Measuring reading comprehension with the Lexile framework. Durham, NC: Metametrics Inc.
Street, B. V. (1984). *Literacy in theory and practice.* NY: Cambridge University Press.
Tatum, A. W. (1999). Reading and the African American male: Identity, equity, and power. *Journal of Adolescent & Adult Literacy, 43*(1), 62–64.
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.
The Council of Chief State School Officers and The National Governors Association Center for Best Practices (2010). *Common core state standards for English language arts and literacy in history/social studies & science*. http://www.corestandards.org/Standards/K12/
Tomasello, M. (2016). The ontogeny of cultural learning. *Current Opinion in Psychology*, *8*, 1–4, doi:10.1016/j.copsyc.2015.09.008
Uccelli, P., & Phillips Galloway, E. (2017). Academic language across content areas: Lessons from an innovative assessment and from students' reflections about language. *Journal of Adolescent & Adult Literacy, 60*(4), 395-404.doi:10.1002/jaal.553
Uccelli, P., Barr, C. D., Dobbs, C. L., Phillips Galloway, E., Meneses, A., & Sánchez, E. (2015). Core academic language skills (CALS): An expanded operational construct and a novel instrument to chart school-relevant language proficiency in pre-adolescent and adolescent learners. *Applied Psycholinguistics, 36,* 1077-1109. doi:10.1017/S014271641400006X
Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly, 50*(3), 337-356, doi.org/ 10.1002/rrq.104
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.

van den Broek, P., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (p. 123–139). IAP Information Age Publishing.

van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York, Academic Press.

Vaughn, S., Solis, M., Miciak, J., Taylor, W. P., & Fletcher, J. M. (2016). Effects from a randomized control trial comparing researcher and school-implemented treatments with fourth graders with significant reading difficulties. *Journal of Research on Educational Effectiveness, 9*(1), 23-44, doi.org/10.1080/19345747.2015.1126386

Vaux, A., Phillips, J., Holly, L., Thompson, B., Williams, D., & Steward, D. (1986). The social support appraisals scale: Studies of reliability and validity. *American Journal of Community Psychology,* 14, 195–219. doi:10.1007/BF00911821.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. Metacognition and Learning, 1, 3-4, doi.org/10.1007/s11409-006-6893-0

Vygostky, L. (1978). *Mind in society: The development of higher psychological processes.* Boston: Harvard University Press.

Wang, Z., Sabatini, J., O'Reilly, T., & Weeks, J. (2019). Decoding and reading comprehension: A test of the decoding threshold hypothesis. *Journal of Educational Psychology, 1*1(3), 387–401.

Wever & Kelemen, 1997 Weaver, C. A., & Kelemen, W. L. (1997). Judgments of Learning at Delays: Shifts in Response Patterns or Increased Metamemory Accuracy? *Psychological Science*, *8*(4), 318–321, doi.org/10.1111/j.1467-9280.1997.tb00445.x

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/

Wigfield, A., Gladstone, J., & Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives 10*(3), 190-195*,* DOI: 10.1111/cdep.12186

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Barber, A. M. T., Klauda, S., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in Schools, 45*(5), 432-445, DOI: 10.1002/pits.20307

Wixson, K. K., & Peters, C. W. (1987). Comprehension assessment: Implementing an interactive view of reading. *Educational Psychologist, 22*, 333-356.

Wu, JY (2013) Gender differences in online reading engagement, metacognitive strategies, navigation skills and reading *Journal of Computer Assisted Learning*

Yeari, M., & van den Broek, P. (2011). A cognitive account of discourse understanding and discourse interpretation: The Landscape Model of reading. *Discourse Studies, 13*(5), 635-643, 10.1177/1461445611412748

Action: 2025 NAEP Mathematics Assessment and Item Specifications

Framework development and update processes are closely monitored by the Assessment Development Committee (ADC). Based on an ADC review of the 2017 NAEP Mathematics Framework involving external experts and a Board-commissioned inventory of state standards, the Governing Board initiated an update of this framework (last updated in 2006). 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 to develop framework recommendations that balance necessary changes with the Board's desire for stable trend reporting, continued breadth of content coverage, and innovation. After identifying Board priorities relative to these recommendations and public comment, the Board adopted a responsive draft of the 2025 NAEP Mathematics Framework in November 2019.

In January 2020, the Board turned toward finalizing the draft specifications – a companion document intended for NCES assessment development. To support the NCES development timeline for a 2025 assessment that reflects the updated framework, the specifications must be finalized by early July 2020. At the May 2020 Board meeting, the full Governing Board took action to delegate authority to the ADC to approve the finalized specifications in accordance with this timeline.

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
Joint ADC/COSDAM Discussion on Specifications	January 2020 Webinar
WestEd Updated Specifications based on Joint ADC/COSDAM Feedback	February 2020
Full Board Discussion on Specifications	March 2020 Board Meeting
Board Action to Delegate Authority to the ADC to Approve Specifications Prior to August Board Meeting	May 2020 Board Meeting
WestEd Updated Specifications with Clarifications based on NCES Feedback	May – June 2020
ADC Action on Specifications	July 2020

2025 NAEP Mathematics Specifications Background

The Board adopted the 2025 NAEP Mathematics Framework in November 2019. The Assessment and Item 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

As the primary audience for the Assessment and Item Specifications, Board staff have collected feedback on the Specifications from NCES to ensure that the vision of the framework is sufficiently clear for NCES assessment and item developers. Below is a summary of how these clarification issues were addressed in the attached draft – shaded text in the attached draft reflects text carried over from the Board adopted framework, while unshaded text is text specific to the Specifications document only. WestEd edits made since the March 2020 Board meeting draft are noted in tracked changes.

Newly Resolved Issues

[1] Clarified the NAEP Mathematical Practice of Justifying and Proving

The brief description of this NAEP Mathematical Practice in the 2025 NAEP Mathematics Framework indicated that mathematical arguments can be proven, but it also said that mathematical arguments could be proofs. This led to some confusion that students would sometimes be asked to prove a proof. It was also not clear whether suppositions, claims, mathematical statements, conjectures, and assertions were being used interchangeably. Finally, exhibits and illustrations did not always reflect consistency in what was considered an adequate defense of an assertion. All of these issues have been addressed.

[2] Clarified the NAEP Mathematical Practice of Collaborative Mathematics

In the previous draft of the Specifications, sample items for Collaborative Mathematics sometimes had features that overlapped with sample items for the other NAEP Mathematical Practices, but this was not always acknowledged. Further, some of the sample items for Collaborative Mathematics were not situated in a collaborative context, which also raised some confusion. Accordingly, WestEd has provided several clarifying updates to sample items.

[3] Clarified Ancillary Tools Can Be Used During the Assessment

There were opportunities to clarify how digital tools and off-screen peripherals could be implemented to ensure fidelity to the intended construct. For example, digital on-screen tools are already incorporated into current NAEP Mathematics Assessments. When on-screen digital tools are to be included in the conception of tool-based (or object-based) responses for the 2025 NAEP

Mathematics Framework and Assessment, NCES required clarification about when current items in the item pool could be classified as matching the 2025 framework. This helps NCES determine how many new items need to be developed for the 2025 framework. WestEd has addressed these issues.

[4] Ensured Overall Accuracy of the Document

In several places, updates were needed for the labeling of illustrative items and the scoring criteria, as well as typographical errors. WestEd has thoroughly reviewed the draft to address inconsistencies, where they existed. Enhancements also include refining the glossary and providing additional rationale on the Panel's recommendations for mathematics-specific contextual variables.

Assessment and Item Specifications for the 2025 NAEP Mathematics Assessment

Draft 4.9
02 July 2020

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

July 2020

TABLE OF CONTENTS

What Is This Assessment and Item Specifications Document?

This document is a companion to the *Mathematics Framework for the 2025 National Assessment of Educational Progress* (NAEP). The 2025 NAEP Mathematics Framework informs NAEP assessment development, describing the subject matter to be assessed and the questions to be asked, as well as the assessment's design and administration. This Assessment and Item Specifications document 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 for the National Center for Education Statistics (NCES) and its 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, socioeconomic 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 *Developing Student Achievement Levels for the National Assessment of Educational Progress Policy Statement* (2018a). 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 A. Example items illustrating each achievement level for each grade level are provided in Appendix B. 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 *Framework Development Policy Statement* (2018b) 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 have been made for the 2025 assessment. These updates include revisions to the mathematics content objectives, descriptions 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 nonexamples, to assist in the development and implementation of updates for the NAEP Mathematics Assessment.

Introduction to the Assessment and Item Specifications

This Assessment and Item Specifications document includes five chapters and several appendices. Throughout this document, figures have been included to illustrate particular points of emphasis from the Framework. Exhibits that have been carried from the Framework into the Assessment and Item Specifications are labeled as "exhibits" and have the same numbering as in the Framework. Figures that are not included in the Framework are labeled as "illustrations." Illustrations in this document include nonexamples—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). In this document, illustrations are numbered consecutively and separately from exhibits.

Chapter 2 describes the content areas: 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 have been 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 examples and nonexamples—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. 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 what they can do. This will involve expanding the assessment 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 nonexamples, 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 A). 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 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 where the focus is 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' lives out of school. 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 can be understood as an additional resource in instruction and assessment. Therefore, the Framework's conception of opportunity to learn includes students' experiences, out-of-school learning, and funds of knowledge as an instructional resource.

Relevant opportunity to learn 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 and practices, instructional strategies, teacher factors, and instruction-relevant 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 opportunity to learn 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 variables are described in Chapter 5.

Major Changes in the 2025 NAEP Mathematics Assessment and Item Specifications

This Assessment and Item Specifications document reflects several major changes, both those made to the Framework and those made to support item development. The changes are summarized in the following sections and elaborated in Exhibit 1.2.

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] [NCES, 2019] and the Programme for International Student Assessment [PISA] [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., Bargagliotti et al., 2020; 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; and 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 (NRC, 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; and
- *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 reiterated 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 previously and as further explained in Chapter 4, technological innovation is relevant to NAEP because it allows for more authentic assessments and for a broader range of accommodations to meet students' needs.

Since 1992, the NAEP Mathematics Assessment has used two types of items (questions): multiple choice and 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 students' 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 were 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; and 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.

Exhibit 1.2. Summary of Changes (continued)

Topic	Change	Rationale
Mathematics Content (continued)	Illustrations containing items and associated text providing clarification for specific text from the Framework were included.	Illustrations containing example and nonexample items, as well as discussions of these items, were included to assist item writers in developing a richer understanding of what was (and was not) intended by the Framework.
Mathematical Complexity (2017 Framework)	This was a chapter that defined mathematical complexity as "the demands on thinking that an item expects" (Governing Board, 2017a, p. 37). The chapter was removed.	From 2009 to 2017, "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 previous 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.

Exhibit 1.2. Summary of Changes (continued)

Topic	Change	Rationale
NAEP Mathematical Practices (continued)	A distribution of items for each mathematical practice was 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 nonexamples 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, Chapter 4 – 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.

Exhibit 1.2. Summary of Changes (continued)

Topic	Change	Rationale
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 and beyond 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 nonexamples of best practices in development of items for the NAEP Mathematics Assessment were also included.	Illustrations containing example and nonexample 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.

Exhibit 1.2. Summary of Changes (continued)

Topic	Change	Rationale
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 acknowledgment 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 with the Framework and the 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 in these Assessment and Item Specifications. More specifically:

- The content of the assessment should be matched with the content of the Framework and the 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 the 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 the Assessment and Item Specifications.
- The mathematical practices reflected in items on the assessment should be matched to the NAEP Mathematical Practices in the Framework and the 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 the 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 the 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 the Assessment and Item Specifications, and should 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 NAEP Mathematical Practices covered by the Framework and the 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 and nonexamples in this document are marked with the designation that matches the item name or identified by the question ID from the NAEP Questions Tool website (NCES, n.d.).

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. For item development, the asterisk applies to the entire objective if it appears immediately after the objective letter, or to an immediately following word or phrase if it prefaces a word or phrase within the objective description:
        - In grade 12 Meas – 3.d, the entire objective is indicated: "d) * Interpret and use…"
        - In grade 12 Num – 1.d, only the "logarithms" aspect of the objective carries an asterisk: "d) Represent, interpret, or compare expressions for real numbers, including expressions using exponents and *logarithms."
        - In grade 12 Num – 2.b, of three aspects listed for the objective, only the third, "analyze the effect," carries the asterisk: "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."
    - Some objectives in grade 12 are marked with the number/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 that are not included in the Framework. These include examples and nonexamples 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."

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 rational 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 by mathematical induction of a simple numerical relationship, and (2) interpreting a proof by mathematical induction of a simple numerical relationship.

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" (OECD, 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 options for assessing 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 reflects not only the focus on fraction instruction in the early grades, but also the importance of understanding students' 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 illustrations come from a variety of sources, including released items from the NAEP Questions Tool (NCES, n.d.), suppliers of state assessments (e.g., SBAC, 2018; 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 most illustrations 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

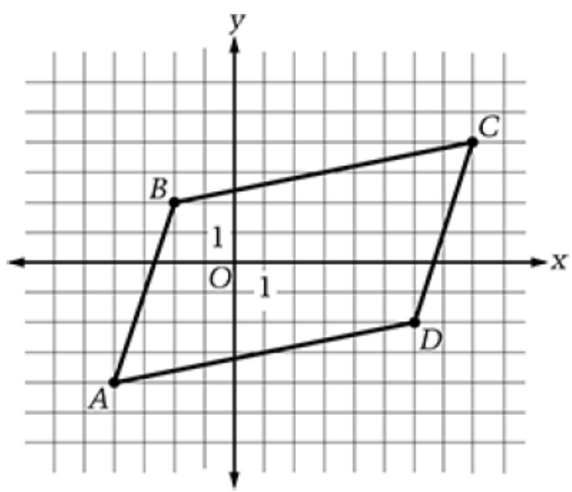
Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Geometry	Other	Geom – 3.h	SCR

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

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 frameworks (see Exhibit 2.4 on page 62 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]

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2009 grade 12 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 scoring information, see Illustration 4.17c in Chapter 4]

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2013 grade 4 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 Framework have been augmented in this document (Exhibits 2.2, 2.3, 2.4, 2.5, and 2.6). The presentation of these specifications includes frequently occurring phrases, the use of italics, and the inclusion of indicator symbols such as caret (^) and plus (+).

Italicized text provides clarifications or limitations to inform item development. All such text is from the 2009 Assessment and Item Specifications document (Governing Board, 2007), except for text that includes a leading symbol.

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

Specifications related to the wording of statements in 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 that 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 that 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 that 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 that include specific details related to human perception, activities, or relationships with the physical world. These specific details are necessary in order for students 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 (Governing Board, 2007), 2025 NAEP Mathematics Framework authors, public-facing information from current state and national assessments (e.g., state assessment websites; PARCC, 2015; SBAC, 2018 and SBAC-related *Progressions* documents https://www.math.arizona.edu/~ime/progressions/), mathematical modeling guidelines (Garfunkel & Montgomery, 2016), and preK–12 statistics guidelines (Bargagliotti et al., 2020).

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 or 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, number sense 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 fractions.*	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.* *^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.*

Exhibit 2.2. Number Properties and Operations (continued)

Num – 3. Number operations (continued)		
Grade 4	**Grade 8**	**Grade 12**
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.*		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.* • *use a real-world context.* • *require application of order of operations.*
c) Divide whole numbers: • Up to three digits by one digit with paper and pencil computation, or • Up to five digits by two digits with use of calculator. *Items written for calculator blocks should not have remainders.*		c) Perform arithmetic operations with expressions involving absolute value.

Exhibit 2.2. Number Properties and Operations (continued)

<table>
<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></td><td>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.</td><td>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 $2\sqrt{3}$ by 1/2.</td></tr>
<tr><td>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.</td><td>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.</td><td>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 $n(n+1)/2$.</td></tr>
</table>

* 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 rationals 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 noncalculator 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, is 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 in the General Guidelines for Measurement subsection of this chapter.

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.

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2013 grade 4 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

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2009 grade 12 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, 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.

Units. Units used in items are cumulative and listed below.

- Grade 4: nonstandard units, common customary units (inch, foot, mile, cup, quart, gallon, pound, hour, minute, day, year, degrees of measured angles, degrees Fahrenheit) and common 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.

Instruments. 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.

Conversions. 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.

Formulas. 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 $A = l \bullet w$.

Both grade 8 and grade 12 students should know formulas for the areas of a rectangle, a triangle, and a circle; the circumference of a circle; and the volumes of a cylinder and a rectangular solid. When other formulas are needed to complete an item, they should be given. See the General Guidelines for Geometry subsection of this chapter 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* • *that 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 that 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 that items should assume students know or can use.*
	i) Solve problems involving rates and ratios such as speed or 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* $\sin 20^\circ$ *and* $2 \sin 10^\circ$ *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

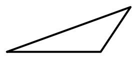
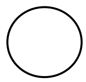
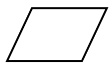
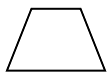
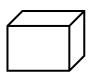
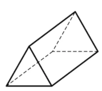
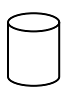
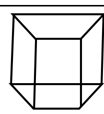
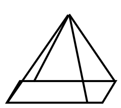
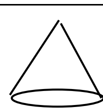
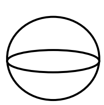
Geometry began thousands of years ago in many lands as sets of practical rules related to describing and predicting locations of astronomical objects, calculating land areas, and building structures. More than 2,200 years ago, the Greek mathematician Euclid organized the geometry known at that time into 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 noncongruent 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 if 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 diagram, paragraph, two-column).* *^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 problems. *+Emphasis should be on grade-level appropriate representations or figures.*	# c) Represent problem situations with geometric figures to solve 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, quadrilaterals, and other polygons; circles; 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 relationships among 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 a 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, 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	2017 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, of the least value, or of any clustering or grouping of data in the distribution.

The composite item in Illustration 2.8 shows two ways objective 2.b can be assessed in grade 4. The item is adapted from England's *Key Stage 2, Paper 3: Reasoning* (Standards and Testing Agency, 2019) 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	

Scoring Information

Part A 255; 410

Part B

Mass (grams)	Number of kittens
250 – 299	2
300 – 349	3
350 – 399	2
400 – 449	1

The item in this illustration is adapted from an England *Key Stage 2* item. The original version of this item appeared as Item 7 in the 2019 administration of *Paper 3: Reasoning*.

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 for each grade level are indicated in fourth row of Exhibit 2.5. Within an objective, a parenthetical list of representations indicates which of the grade-level appropriate representations is applicable for that 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 numbers 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 in the next row. For some objectives only a subset of the representations is applicable, indicated by a parenthetical list at the end of 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/bimodal). *+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, of 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) 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 a 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) 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 in which 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.*	

	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.*

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

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, or 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 highlight 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 nonlinear equations is not expected at grade 8, but is 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 only in items written for objectives identified with an asterisk (*).

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 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 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 Alg – 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.*	

* 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**
	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 nonlinear 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, or 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 noncalculator 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* $g(x) \bullet f(x) = 0$.
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, $C = \pi d$, 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.
* 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**
		# 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), *Guidelines for Assessment and Instruction in Statistics Education* (GAISE; Bargagliotti et al., 2020), *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 policies; 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 (Governing Board, 2017a) 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 and 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 et al., 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 previously noted, 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).

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 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.

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, or refuting mathematical claims 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, 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.

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 Mathematical 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 School Mathematics* (1989), *Principles and Standards for School Mathematics* (2000), and *Catalyzing Change in High School Mathematics* (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 NAEP Achievement Level Descriptions (ALDs; see Appendix A) 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 NAEP mathematics content and practices. Assessment developers need to create a pool of items that reflects the Framework and the range of achievement levels. Because consideration of 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 NAEP Mathematical Practices and content objectives is 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.

All released NAEP items used as exhibits in the Framework and in these specifications were accessed using the online NAEP Questions Tool (NCES, n.d.). Some examples are from other sources, including example items from the Smarter Balanced Assessment Consortium (SBAC), released items from the Partnership for Assessment of Readiness for College and Careers (PARCC), and adaptations of tasks from policy and curriculum documents. The source for each item is cited in related text description about the item.

NAEP Mathematical Practice 1: Representing

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

Focus for Item Developers

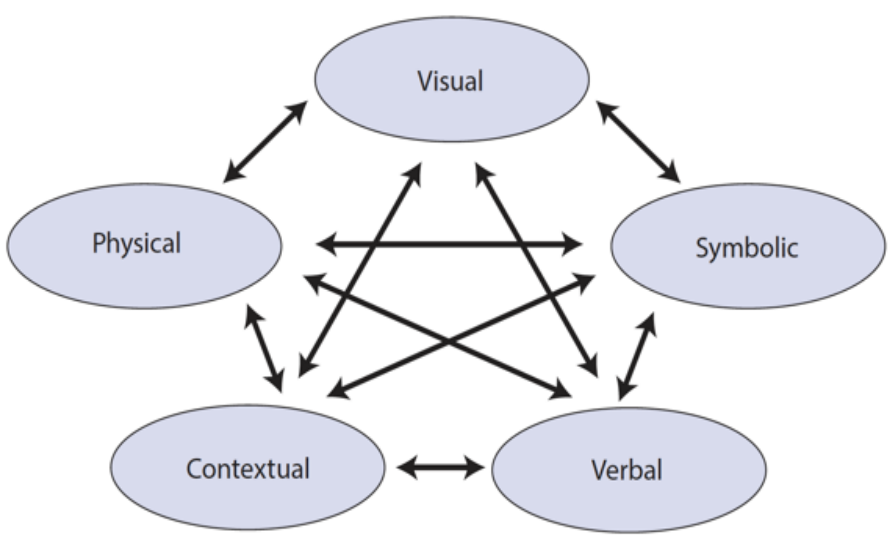
Examples of ways that students can engage in the NAEP Mathematical Practice of Representing include, but are not limited to,

- constructing visual representations of numbers, shapes, and data;
- translating from one mathematical representation to another;
- using representations as tools to solve problems; and
- building on, analyzing, and explaining representations created by others.

Each item associated with this practice should focus mathematical activity on

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

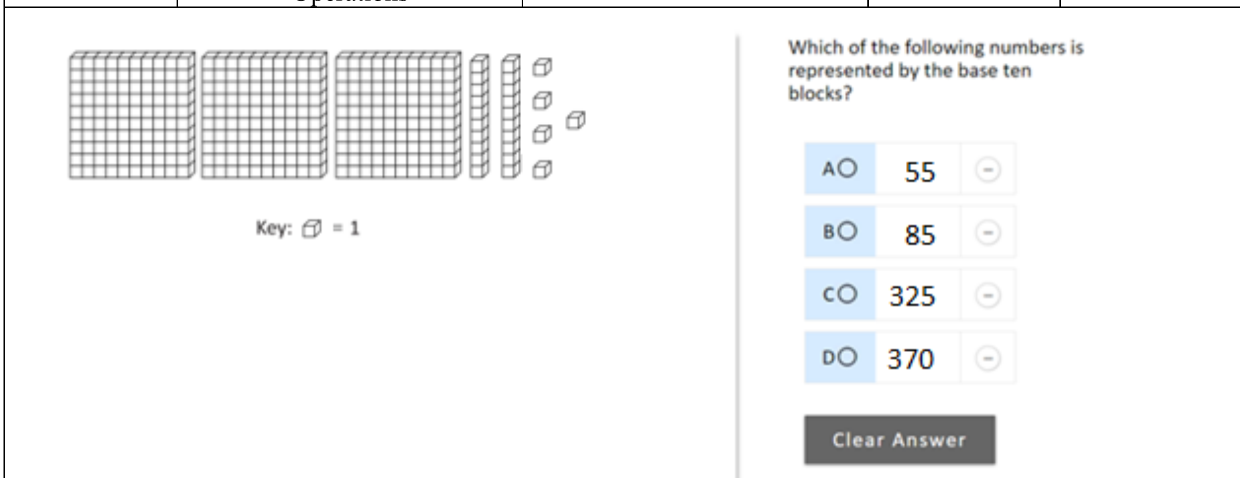


According to the National Research Council (NRC, 2009), students, especially young ones, benefit from using physical objects or acting out processes during problem solving. Base 10 blocks (or blocks/tiles representing other bases), fraction strips/bars, red–black integer tiles, and algebra tiles are 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 grade 4 item in Illustration 3.1 is adapted from Exhibit 3.3 in the Framework. The item provides an image of base 10 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. The item is framed to elicit a basic level response.

Illustration 3.1. Representing Example: Base 10 Blocks
adapted from **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



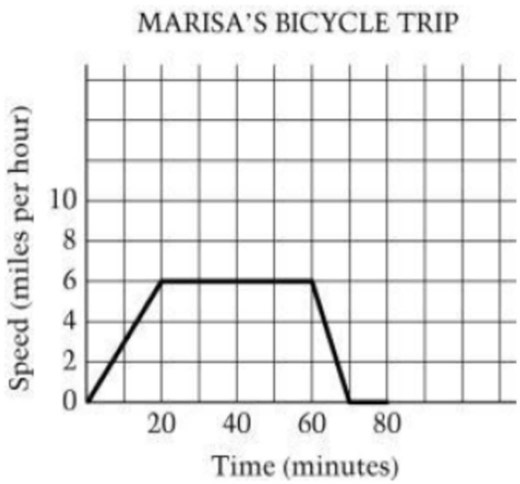
Scoring Information	
Key	C. 325
ALD Notes for Item Developers	
Basic	The item assesses understanding of a visual representation of a familiar set, whole numbers.
Proficient	The item could be revised to present a representation of a whole number using base 10 blocks and ask for a different representation 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.

The item in this illustration is adapted from a NAEP item. The original version of this item appeared in the 2017 grade 4 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 ________

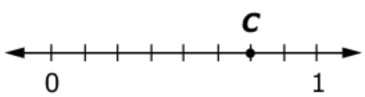
From 60 minutes to 80 minutes Marisa ________

Illustration 3.2 is adapted from Exhibit 3.5 in the Framework. The Smarter Balanced (SBAC) item shown 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
adapted from **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.

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 for the fraction of a mile that Valeria walked, as 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 ask for the selection of a response that describes how the number line could be used to determine the distance Diego walked.

The item in this illustration is adapted from an SBAC item with Item ID 3218, aligned to CCSS-M objective 5.NF.A.2.

Illustration 3.3 is based on Exhibit 3.6 in the Framework. Similar to the previous item, the SBAC item shown 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?

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

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

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

(C) 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?

(D) 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 eights 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.

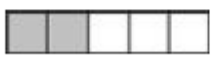
The item in this illustration is based on an SBAC item with Item ID 3274, aligned to CCSS-M objective 5.NF.B.7b.

Translating from one mathematical representation to another is a component of the practice of representing. For instance, the grade 4 item on the left in Illustration 3.4 (example) asks students to write a fraction to describe the shaded part of a figure, a translation from the visual to the numeric. In contrast, the grade 8 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 NAEP Mathematical 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
left: 4	Number Properties and Operations	Representing	Num – 1.e	SCR – FIB
right: 8		Other	Num – 1.i	SR – MC

Example



What fraction of the figure is shaded?

Answer: _______________

Nonexample

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**	2/7, 1/2, 5/9

The grade 4 item in this illustration is based on a NAEP item. The original version of this item appeared in the 2007 NAEP Mathematics Assessment with NAEP Item ID 2007-4M7 #6 M139301.
The grade 8 item in this illustration is based on a NAEP item. The original version of this item appeared in the 2007 NAEP Mathematics Assessment with NAEP Item ID 2007-8M9 #12 M013631.

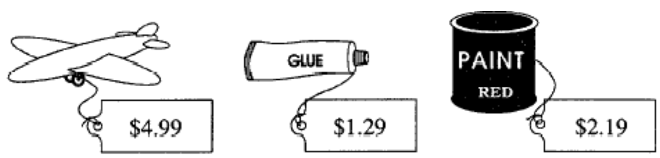
The shaded figure in the example item in Illustration 3.4 is central to the item's assessment of representing. However, the inclusion of an image in an item does not automatically address the NAEP Mathematical Practice of Representing.

Items may include images to convey information that could be provided another way, such as through written text. For a given image to be associated with the practice of representing, the image would need to be a representation with which students engage mathematically and that is critical to the solution process.

Though the item in Illustration 3.5 uses an image to convey information critical to solving the problem (each cost is needed), the image itself is not essential for the mathematical activity required, nor is there a need for students to construct or analyze representations to solve the problem. Therefore, the item in Illustration 3.5 does not assess the NAEP Mathematical Practice of Representing.

Illustration 3.5. Representing Nonexample: Image Does Not Address Representing

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Number Properties and Operations	Other	Num – 3.f	SR – MC



Chen bought one model plane, one tube of glue, and one can of paint. The cost of each item is shown in the figure above. There was no sales tax. How much change should he have gotten back from $10?

A. $1.50

B. $1.53

C. $1.63

D. $1.73

Scoring Information

Key	B. $1.53

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 1990 grade 4 NAEP Mathematics Assessment with NAEP Item ID 1990-4M9 #7 M026131.

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.

Focus for Item Developers

Engaging in the NAEP Mathematical Practice of Abstracting and Generalizing involves one or both of the processes of abstracting or generalizing.

- *Abstracting* is the process of decontextualizing ideas in a given problem or context and expressing, representing, and manipulating them in a manner independent of initial contextual references (Scheiner & Pinto, 2016).
- *Generalizing* is the process or outcome of at least one of the following actions (Ellis, 2011, p. 311):
    - identifying commonality across cases,
    - extending one's reasoning beyond the range in which it originated, or
    - deriving broader results from particular cases.

Abstracting

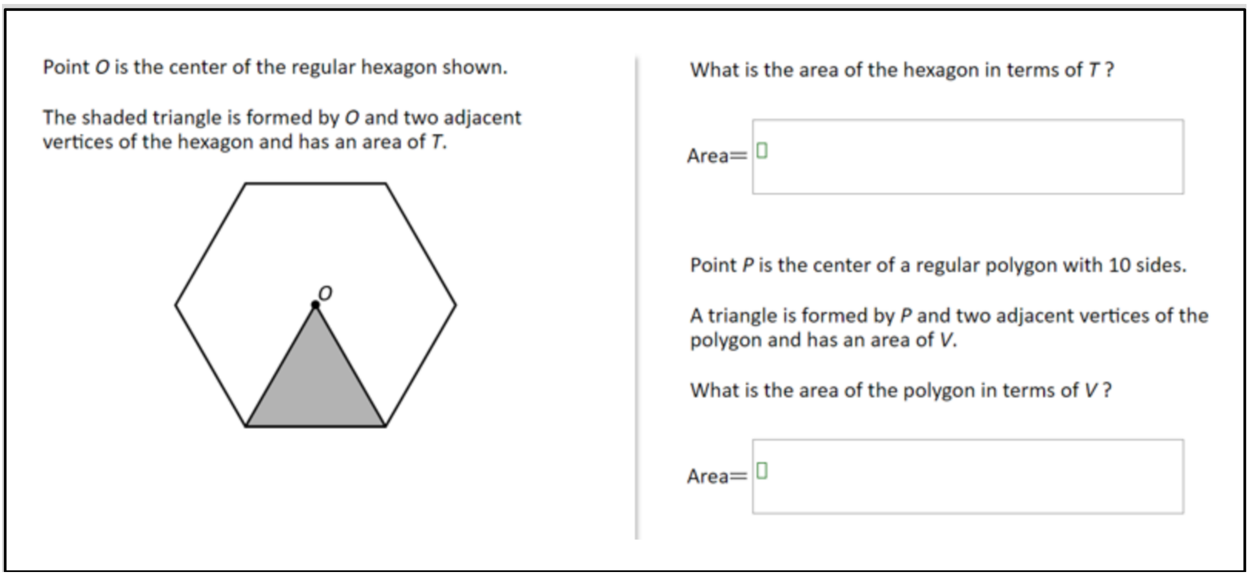
Students learning and doing mathematics also engage in the practice of abstracting and generalizing. An essential element of mathematical learning and problem solving is the ability to reason abstractly and to develop, test, and refine generalizations. In reasoning abstractly, students engage in the 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 grade 8 NAEP 2017 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 previously, 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. When expressing involves both representing and generalizing, more than one NAEP Mathematical Practice could be engaged. In these cases, the practice emphasized to a greater extent in the item should be selected as the primary practice.

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 Exhibit 3.8. The item in Illustration 3.6 is a *NAEP Advanced* variant of the item in Exhibit 3.8. The item shown removes the scaffolding of parts (a) and (b) from the original item. It is worth noting that for items such as Exhibit 3.8, 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 (e.g., as in Illustration 3.6).

Notice that for part c of the grade 12 item in Exhibit 3.8 and the adaptation in Illustration 3.6 students are expected to write a formal algebraic rule for moving from the n^{th} term to the $(n + 1)^{\text{st}}$ term of Sequence I by identifying an explicit rule for the n^{th} term of Sequence II. In other items, students may be tasked with determining a recursive, rather than explicit, rule to find the n^{th} term in a sequence.

Exhibit 3.8. Grade 12 NAEP Number Pattern Item

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 $(n+1)^{st}$ term and the n^{th} term of Sequence I.

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

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

The sequence shown is increasing. Each term in the sequence is greater than the previous term.

3, 5, 9, 17, 33, . . .

The same pattern continues for the terms in the sequence. Write an algebraic expression (rule) to represent the difference between the $(n + 1)^{st}$ term and the n^{th} term of the sequence.

Scoring Information	
Key	2^n
ALD Notes for Item Developers	
Basic	The item could be revised to assess the identification of the pattern of differences (i.e., part [a] in Exhibit 3.8).
Proficient	The item could be revised to assess the extension of the pattern of differences (i.e., parts [a] and [b] in Exhibit 3.8).
Advanced	The item assesses use of structures and patterns to determine a complex rule.

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2005 grade 12 NAEP Mathematics Assessment with NAEP Item ID 2005-12M3 #17 M095401.

Determining a rule for a pattern is a common focus of grade 4 generalization items, such as the adapted 2011 grade 4 TIMSS item shown in Illustration 3.7 (International Association for the Evaluation of Educational Achievement, 2013). 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, 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]

The item in this illustration is adapted from a TIMSS item. The original version of this item appeared in the 2011 TIMSS grade 4 assessment with 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 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, breaking mathematical components of an item apart to identify the building blocks needed to answer a question (Cuoco, Goldenberg, & Mark, 1996; Küchemann & Hoyles, 2009). Thus, a distinction needs to be made between items that ask students to reason structurally and items that prompt students to identify or apply known quantities or properties. Consider the grade 8 item in Illustration 3.8. For each question a set of geometric objects is given and a single example is sufficient 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 and is not an assessment of the NAEP Mathematical Practice of Abstracting and Generalizing (see Chapter 4 for additional information about the grid 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

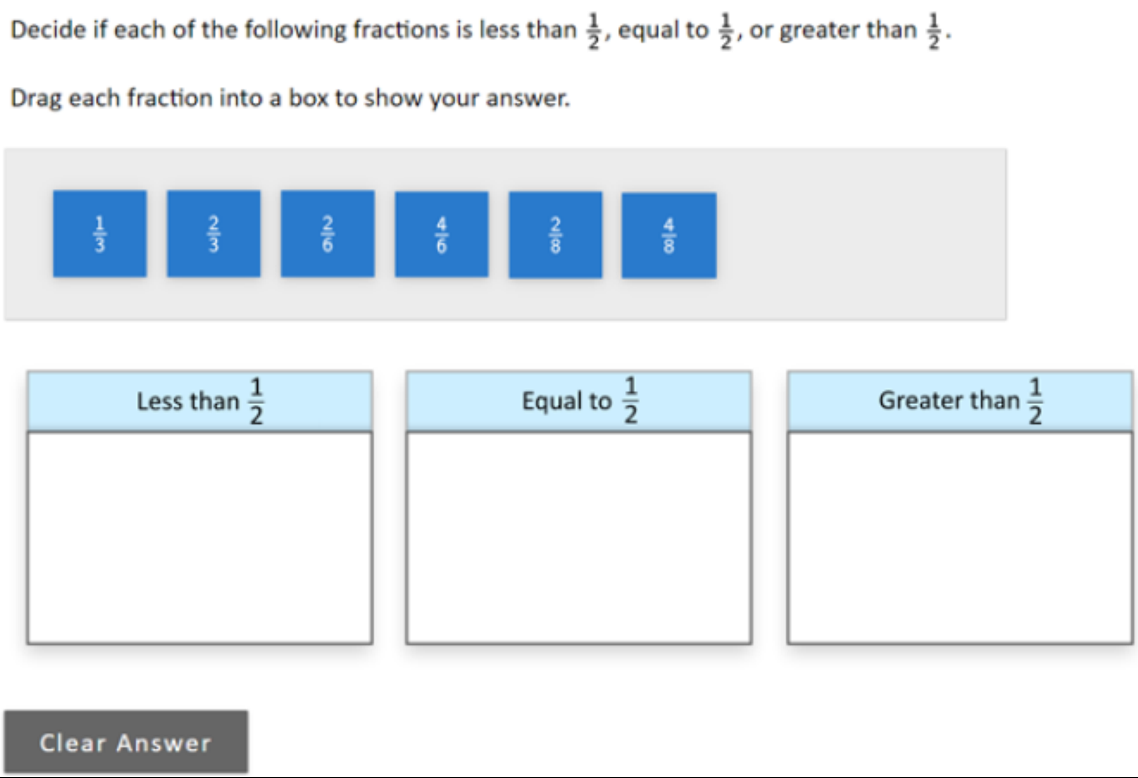
Questions	0	1	2	4	8
At how many points do two different parallel lines meet?	⊙	○	○	○	○
How many right angles are formed by a pair of perpendicular lines?	○	○	○	⊙	○
How many angles that measure less than 180 degrees are formed when two parallel lines are cut by a transversal?	○	○	○	○	⊙

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 8 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #10 M3821MS.

Since the NAEP Mathematical 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. Note that they are not asked to determine the structure of a fraction that is less than, equal to, or greater than 1/2. The thought processes 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. 186, for additional information about the matching item type represented in Illustration 3.9).

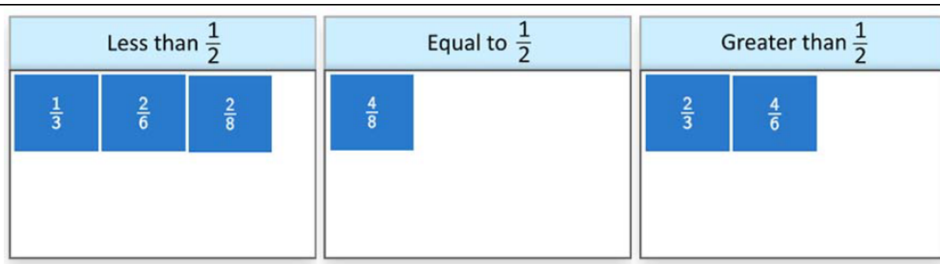
Illustration 3.9. Abstracting and Generalizing Nonexample: Using Benchmarks

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



The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 4 NAEP Mathematics Assessment with NAEP Item ID 2017-4M1 #6 M3714MS.

NAEP Mathematical Practice 3: Justifying and Proving

> *Justifying and Proving: Creating, evaluating, showing, or refuting mathematical claims in developmentally and mathematically appropriate ways.*

Focus for Item Developers

Explaining why something is true or not true is an important aspect of mathematical argumentation. However, not all mathematical arguments involve the NAEP Mathematical Practice of Justifying and Proving.

While the practice includes creating, evaluating, showing, or refuting mathematical claims, distinctions to note for the NAEP Mathematical Practice of Justifying and Proving include:

- *Justifying* involves a deductive argument demonstrating why a statement or claim is true or not true generally (i.e., for all cases).
- *Proving* involves the formal presentation of a valid justification.
- Examples alone do not suffice as a mathematical justification or proof except for proofs by exhaustion or counterexample.
- Explanation of mathematical reasoning aligns with the practice only when the item requires justifying a general statement, not specific instance(s) (e.g., Illustrations 3.12a, 3.14, and 3.15).
- Items to assess generalizing and justifying differ.
    - Given information, a student generalizes (e.g., Illustration 3.6).

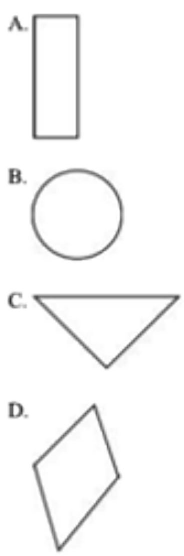
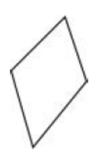
Justifying and proving are essential in all content areas and grade levels. Traditionally, proof was viewed as a form of mathematical argumentation pertaining first to high school geometry and not visited again until pre-calculus courses with proofs of trigonometric identities and proofs by mathematical 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), 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.

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 item in Illustration 3.10, which asks students to choose a counterexample to Alan's claim from a set of response options. The required action reflects emerging development of the practice of justifying and proving.

Illustration 3.10. Justifying and Proving Example: Choose the Counterexample

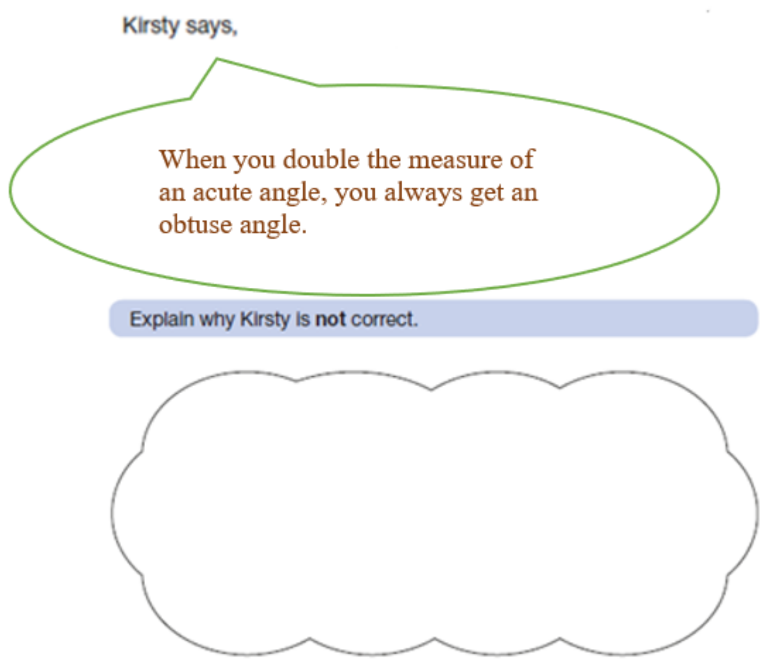
<table>
<tr><th>Grade Level</th><th>Content Area</th><th>Assessed Practice(s)</th><th>Objective ID</th><th>Item Format</th></tr>
<tr><td>4</td><td>Geometry</td><td>Justifying and Proving</td><td>Geom – 1.e</td><td>SR – MC</td></tr>
<tr><td colspan="5">Alan says that if a figure has four sides, it must be a rectangle. Gina does not agree. Which of the following figures shoes that Gina is correct?

</td></tr>
<tr><td colspan="5">Scoring Information</td></tr>
<tr><td>Key</td><td colspan="4">D.
</td></tr>
</table>

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2003 grade 4 NAEP Mathematics Assessment with NAEP Item ID 2003-4M6 #7 M046401.

The adapted *Key Stage* released item in Illustration 3.11 provides a mathematical statement and asks for an explanation of why the statement is incorrect. As with the previous illustration, this item, at a minimum, requires presentation of a counterexample, a single example of when doubling an angle measure results in an angle that is not obtuse. As posed, the item primarily assesses the practice of justifying, even though a student might respond in a way that generalizes the range of angle sizes for which the given statement is false. Since that type of response is not requested, the item does not primarily assess the practice of generalizing.

Illustration 3.11. Justifying and Proving Example/Abstracting and 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%

The item in this illustration is adapted from an England *Key Stage 2* item. The original version of this item appeared as Item 13 in the 2019 administration of *Paper 3: Reasoning*.

Definitions are often used to justify mathematical statements. However, applying a definition is not necessarily engaging in the NAEP practice of Justifying and Proving. The grade 12 item in Illustration 3.12a, adapted from a 2009 NAEP item, provides a mathematical claim, "*y* is a function of *x*," about the particular values in the given table. In generating a correct response to the item, students are likely to use some form of the definition of function to write the requested explanation. However, the item asks about the particular case of the *x*-*y* relation in the table and does not call on students to create an explanation that is generally true (e.g., for a class of cases). As in the sample explanation in the key, students might make a general explanation, but the item does not require it. Therefore, the item does not assess the NAEP Mathematical Practice of Justifying and Proving.

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

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Algebra	Other	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.

The item in this illustration is adapted from a NAEP item. The original version of this item appeared in the 2009 grade 12 NAEP Mathematics Assessment with NAEP Item ID 2009-12M2 #7 M1906E1.

By contrast, consider the variant of Illustration 3.12a shown in Illustration 3.12b. The item in Illustration 3.12b calls for a justification that is generally true for a class of examples, in this case all linear relationships between *x* and *y*. Similarly, Illustrations 3.12c and 3.12d for grade 8 and grade 4, respectively, use items where students are called on to determine and explain the validity of general mathematical relationships. In Illustrations 3.12b, 3.12c, and 3.12d, students might leverage examples to demonstrate why a statement is generally true, yet the items require justifying general statements rather than specific instances.

Illustration 3.12b. Grade 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

The equation $y = cx + d$, where c and d are constants and $c \neq 0$, defines a relationship between the variables x and y. Explain whether or not y is a function of x.

Illustration 3.12b. (continued)

Sample Student Responses

(1) If y is a function of x, then for each x there is exactly one y. So what if it is not a function? Let's pick two values for y and say they correspond to the same x. We'll call these values of y, w and z.
Since w and z correspond to the same x, $w = cx + d$ and $z = cx + d \rightarrow w = z$.
Since w and z are the same number, there are not two values of y that correspond to the same x.
Therefore, y is a function of x.
(2) y is a function of x because for each x, when you multiply by whatever c is you get exactly one number and then you add whatever d is and still get one output. So, there is one output y for one input x. The definition of function is that for every input there is only one output, so y is a function of x.

Illustration 3.12c. Grade 8 Justifying and Proving Example: Supporting with a Definition

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

Ella makes this statement.

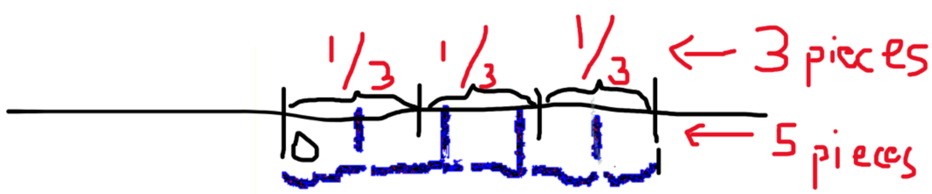
The opposite of the opposite of a number is the number.

Use a number line to explain whether or not Ella's statement is true.

Sample Student Response

(1) [student uses additive inverse]
Ella's statement is true. For example, the opposite of 3 is –3, since the opposite of a number and the number are on the opposite side of 0 on the number line but the same distance from 0. Using the same thinking, the opposite of –3 is 3. Therefore, the opposite of the opposite of 3 is 3.
This reasoning can be applied to any number. Let's say the number is n. The opposite of n is $-n$. The opposite of $-n$ is the same distance from 0 on the number line but on the opposite side from $-n$, so it is n. Therefore, the opposite of the opposite of n is n.

(2) [student uses multiplicative inverse]
If the opposite of a number is 1/number, then flipping it again would give the number back. This is because, on a number line, for example, 1 over 3 means the fraction that goes into one, three times, like in the picture. There are 3 pieces of 1/3. And 1 over 5 means the fraction that goes into one five times, three are 5 parts and each is 1/5, etc., so 1/number would go into 1 that number of times.



The opposite of 1/number would be asking how many times does 1 go into number. Well the answer is always the number – for example, 1 goes into 3, 3 times, 1 goes into 5, five times, and 1 goes into number, number times. So, the opposite of number is 1/number and the opposite of 1/number is number so the opposite of the opposite of a number is number and Ella is saying a true thing.

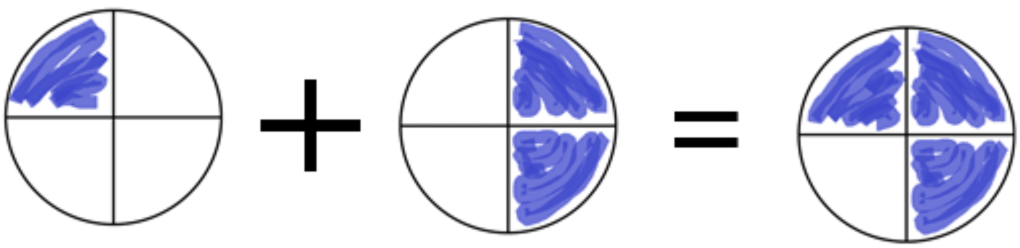
Illustration 3.12d. Grade 4 Justifying and Proving Example: Supporting with a Definition

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

Terry is adding two fractions with the same denominator. In his answer, Terry adds the numerators but keeps the denominator the same. Explain whether or not Terry's answer is correct.

Sample Student Response

Terry's answer is correct because fractions are out of a whole.
Here is an example:



This picture shows that 1/4 + 2/4 = 3/4. This works because the wholes are the same size and the parts are the same size. When you combine the parts, the whole doesn't change. You just add the parts. You could do the same thing, no matter how many parts, as long as each whole was the same number of parts.

A proof can have many different forms, including narrative, pictorial, diagram, two-column, or algebraic forms. The form used to represent a mathematical proof is valid as long as it communicates the proof's essential features, namely, that it contains logically connected mathematical statements that are based on valid definitions and theorems. Consider the grade 4 item in Exhibit 3.10. Some students may use specific examples in their arguments, but 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.

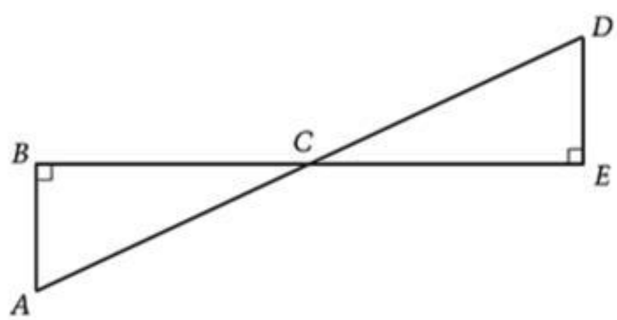
A grade 4 proof for the claim in Exhibit 3.10 could involve demonstrating with either pictures or symbols that the answer can always be separated into two equal parts, because 2 is a factor of 6, or that the answer can always be divided by 2 or cut in half because 2 already divides 6. An argument such 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). 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.

In addition to the various formats one can use to develop or present proofs, there are also other ways of mathematically proving, disproving or justifying a mathematical answer. These include developing deductive arguments, finding counterexamples, 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.

The process of refuting – demonstrating that a statement is false – is a key element of justification because conjecturing can produce both true and false statements. Students must understand that a single counterexample disproves a conjectured generalization.

An example of the value of finding a counterexample 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 $x > –3$ <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> $x < 5$?

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 shown in Illustration 3.13 (based on the item in 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.

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	Tracy is correct. 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 any whole number and 1/2 is less than the whole number, 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 provide a statement about multiplying 6 and any irrational number, which students need to justify or refute.

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 1992 grade 12 NAEP Mathematics Assessment with NAEP Item ID 1992-12M14 #2 M054801.

Understanding that a single counterexample undermines a general claim is an important but difficult aspect of justification. 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.

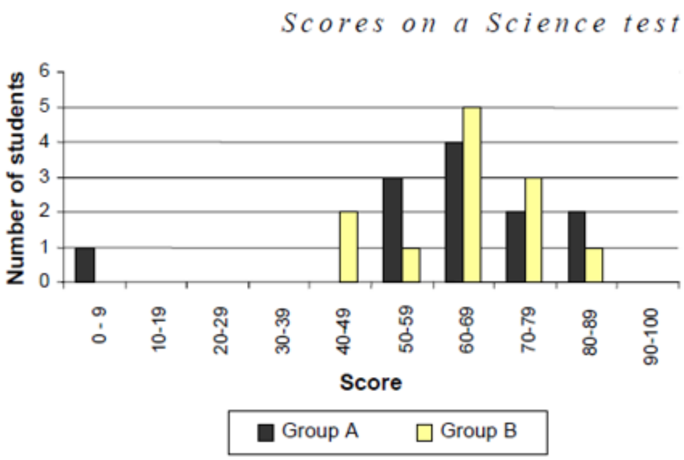
Explaining why something is true or not true is an important aspect of mathematical argumentation. However, not all mathematical arguments involve the NAEP Mathematical Practice of Justifying and Proving. Consider Illustration 3.14, a released 2006 PISA Data Analysis, Statistics, and Probability item (OECD, 2006). The item provides a data set and asks for a mathematical argument to counter a claim by a teacher that Group B did better. The requested mathematical argument is based on the single set of data represented in the graph, which is a specific instance and not a general case. Therefore, though the item in Illustration 3.14 may assess skill with inferential reasoning and identifying evidence for making a mathematical argument, it does not assess the NAEP Mathematical Practice of Justifying and Proving.

Illustration 3.14. Justifying and Proving Nonexample: Reasoning from Data

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
12	Data Analysis, Statistics, and Probability	Other	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 above.

The item in this illustration is based on a 2006 PISA item with Item ID M513Q01 – 0 1 9.

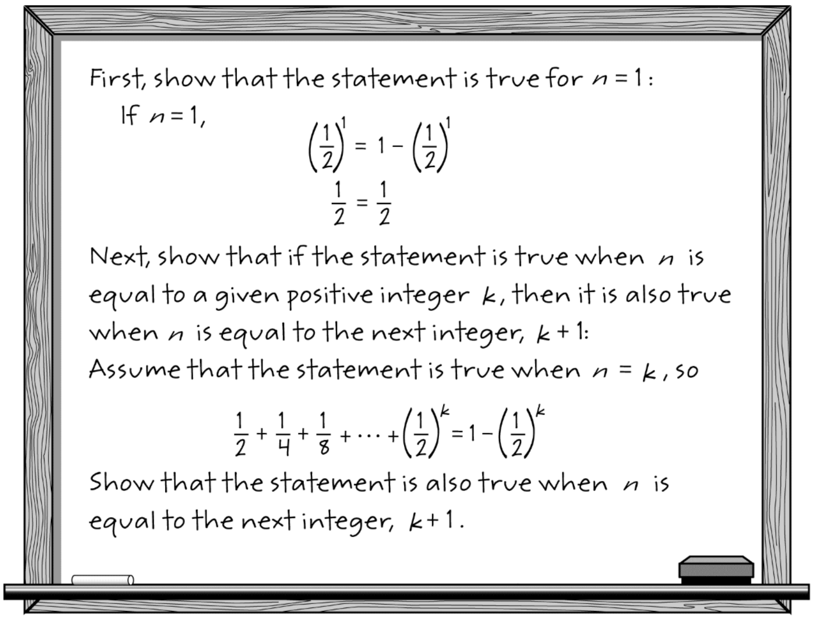
Some NAEP items require a specific mode of proof, such as the grade 12 Number Properties and Operations item in Exhibit 3.14.

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.

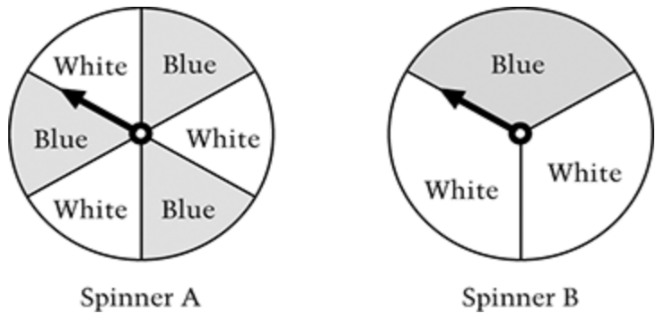
Here, a student must use the tools of mathematical induction to complete the provided argument:

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 deciding 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.

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.

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2011 grade 4 NAEP Mathematics Assessment with NAEP Item ID 2011-4M9 #15 M1609E1.

Asking students to explain *why* the spinner they chose gives Lori the greater probability of landing on blue foregrounds justifying; however, as given, a correct response to the item might be example-based rather than appealing to the general case. 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.

A task in which students evaluated these arguments would assess justifying and proving. It would provide students an opportunity to distinguish between an example-based justification (e.g., Calista's) and those based in what will generally be true of results from each spinner.

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.

Conversely, explaining one's reasoning for how a problem is solved by showing or describing steps in determining an answer is not sufficient for assessing the NAEP Mathematical Practice of Justifying and Proving. Consider the grade 8 item in Illustration 3.15. Responding to the item's prompts for a number of quarts and "how you found your answer" could be satisfied by a student writing down the calculations used to arrive at a conclusion (the number of quarts the student claims Tyler drinks in 7 days) using the given information (indisputable statements such as the information provided about Tyler's milk intake). However, the item does not explicitly call for the reasons why a student might multiply 24 by 7, then divide by 32. Furthermore, the item does not ask students to support why a proposed solution process would work for any problem of this type. For items that ask students to "show" or "explain" to measure justifying, a deductive argument about a general claim (in this case, a mathematical process) must be a part of the item's measurement intent.

Illustration 3.15. Justifying and Proving Nonexample: No General Claim

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Measurement	Other	Meas – 2.b	SCR

Tyler drinks 24 fluid ounces of milk each day for 7 days. How many quarts of milk does he drink in the 7 days? Do not round your answer. (1 quart = 32 fluid ounces)

Answer: ____________________ quarts

Show how you found your answer.

Scoring Information	
Key	5.25 (or equivalent) quarts Possible work shown: $24 \times 7 = 168$ $168 \div 32 = 5.25$

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2013 grade 8 NAEP Mathematics Assessment with NAEP Item ID 2013-8M7 #9 M1687E1.

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, applying the mathematization to reach a solution, and checking the viability of the solution in developmentally and mathematically appropriate ways.*

Focus for Item Developers

The NAEP Mathematical Practice of Mathematical Modeling involves one or more components of the modeling cycle described by Garfunkel and Montgomery (2016, pp. 12–13). Components of the modeling cycle are referenced throughout this section, using descriptions and the letters from this list:

(a) identifying the problem;
(b) making assumptions that may simplify the problem and then identifying variables;
(c) mathematizing the situation;
(d) analyzing and assessing solutions;
(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;
(f) if there seems to be a feasible real-world solution, implementing the model; and
(g) reporting out results.

Each item associated with this practice should address more than one of these

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 bivariate scatterplot, 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 previously described, 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. In 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 focus on multiple steps of the cycle of mathematical modeling driven by the 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.

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, Office of Disability Employment Policy (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 of 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.

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 components (a) and (b) 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 8^{th} 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 sizes of pizzas should we order? What is the cost of each size of pizza?" Here students would need to devise a survey (identify the problem) and 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 is 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). The first three parts of the task are a scaffold to the modeling-heavy work of parts 4 and 5. Parts 3 and 4 engage students in aspects of modeling components (b), (c), and (d) when identifying variables, mathematizing situations, and analyzing and assessing solutions. Part 5 engages students in components (d), (e), and (g) of the modeling cycle.

Exhibit 3.18. Grade 12 Example: Modeling Income Tax Scenario

A state’s tax model is described below.

- Individuals with an income of $10,000 or less per year pay no income tax.
- Individuals with income greater than $10,000 per year pay a 6% tax on all income over $10,000.

(1) What would a resident who made $40,000 pay in tax? What percent of this resident’s total income is paid in tax?
(2) What would a resident who made $50,000 pay in tax? What percent of this resident’s total income is paid in tax?
(3) Determine a method for calculating the percent of any resident’s total income that is paid in tax.
(4) Is there a highest percent of total income that a resident could pay in tax? Defend your position on this percent.
(5) The state is considering the new tax model described below:
  - Individuals with an income of $10,000 or less per year pay no income tax.
  - Individuals with an income greater than $10,000 per year
    - pay 5% on all income over $10,000 up to $50,000, and
    - pay 7% on all income over $50,000.

  Explain whether or not the new tax model benefits individuals in the state who pay income tax. As part of your response, compare the new tax model to the existing tax model.

Access to digital tools, such as equation editors, graphing tools, and spreadsheet tools, would be important in the assessment of students’ modeling practices on tasks like Exhibit 3.18. For example, in parts 3 and 4, the percent income paid in tax can be expressed as the ratio of tax T to income I, or T/I (identify variables). When students compute the tax on income I, with the given 6% rate after the first $10,000 of income, they arrive at $T = 0.06(I – \$10{,}000)$ (mathematize the situation). A symbolic model for the percent income paid in tax could be $T/I = 0.06(I – 10{,}000)/I$. To answer questions about the highest possible tax rate, students could create a graphical model of the percent income paid in tax 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 the percent income paid in tax or even a graph of the percent income paid in tax as a function of income (analyzing and assessing the solution). Modeling carries through to parts 4 and 5 as students compare the new model to the original model. This comparison could be explored through the use of a spreadsheet tool that allows students to rapidly compute the total tax for a given income based on each model. In part 5, the item includes component (g) of the modeling cycle with students reporting out on whether or not the new tax model is recommended.

Modeling processes also often arise in data analysis and statistics. The task in Exhibit 3.19 is an example taken from the online 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 complete modeling cycle from (a) to (g) and closely follows the statistical investigation process as outlined by Bargagliotti and colleagues (2020): identifying a statistical question for investigation, gathering appropriate data, analyzing the data, and 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, Anhalt, Cortez, Turner, & Simic-Muller, 2019). In this task, students are asked to think as a member of a team 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	ECR

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) Write 5 things that the science club members need to consider as they work to determine the number of bottles of water their community would need in a water crisis.

2) The science club members want to learn what they can from the water crisis in the news. 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** questions to which the members need answers so they can determine how many bottles of water the community in the news actually needed.
   - Explain how knowing answers to these two questions will help the science club members in the process of answering their question.

3) There were about 96,000 residents in the community in the news 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. It 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

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2013 grade 8 NAEP Mathematics Assessment with NAEP Item ID 2013-8M3 #2 M1499E1.

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 (NCES, n.d.) of the item shown in Illustration 3.18. In this item, students use a given representation to solve a problem, assessing the NAEP Mathematical Practice of Representing. However, the NAEP Practice of Mathematical Modeling is not assessed. In part, this is 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.d	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

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2013 grade 8 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. Attending to the requirements for representing and modeling will be useful in item development and will allow for distinguishing NAEP Mathematical Practices assessment intent.

Illustration 3.19 shows the first part of a released PARCC item. When considering only this part of the item, the NAEP Mathematical Practice of Mathematical Modeling is not assessed.

The request for a "model" in the item in Illustration 3.19 calls for students to create a symbolic representation of the relationships among costs without requiring substantive engagement in the modeling cycle. Components (a), (b), and (c) of the modeling cycle are provided in the item stem; components (d) and (e) are neither required by nor used as scaffolding within the item; and no opportunity for components (f) or (g) exists in the item as written. Attention to the nuance in the assessment of modeling is included in this document because item writers may work on item development for several different assessments at the same time. This illustration is intended to highlight the distinctions between assessment intent for NAEP Mathematical Modeling items and that for other assessments.

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

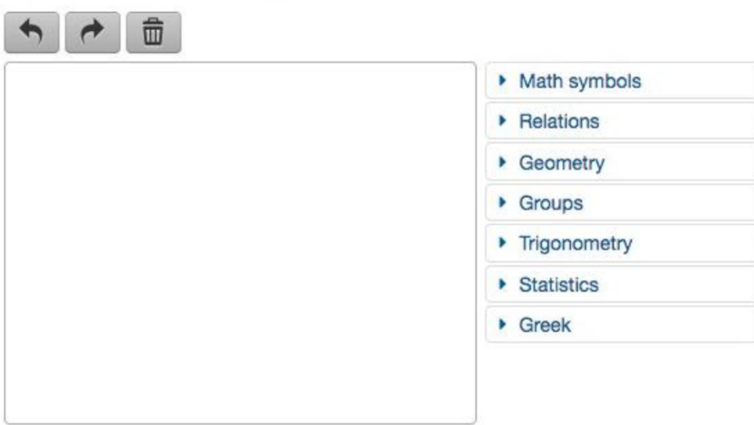
Item	Company X	Company Y
base rental charge	$29.95 per day	$19.95 per day
mileage charge	59 cents per mile	79 cents per mile
drop-off charge	$150	included
insurance	$18 per day	$26 per day

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.



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

The item in this illustration is based on a 2018 PARCC item with Item ID VH145748, aligned to evidence statement HS-D.CCR.

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.

Focus for Item Developers

The NAEP Mathematical Practice of Collaborative Mathematics involves the engagement of mathematical knowledge and skills within a collaborative context. The three measurable skills associated with this practice are:

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

Each item associated with this practice should address one or more of these skills.

Collaborative mathematics in the world of work refers to the talk and actions people engage in with one another as they participate in a necessary collaboration, where the mathematical task is too complex or messy for an individual to meet its demands alone (Fiore et al., 2017). The degree of complexity is different for collaborative mathematics tasks used in schools and assessments. It is true that tasks are designed to require collaboration (e.g., multiple parts, multiple roles, multiple strategies, comparisons of strategies), but it is not necessarily true that the mathematics is too complex or messy for one student; rather, the task may be such that it is designed to require multiple people.

As a practice, collaborative mathematics exists alongside other mathematical practices. That is, as students work together toward 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. 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 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 a highly contextualized activity, tied to cultural ways of working together both in and out of the classroom. As stated in the opening of this 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 Programme for International Student Assessment (PISA), 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 (e.g., "Brad") and the assessed student ("you") through item response choices.

Exhibit 3.20. Example PISA Collaborative Problem-Solving Item

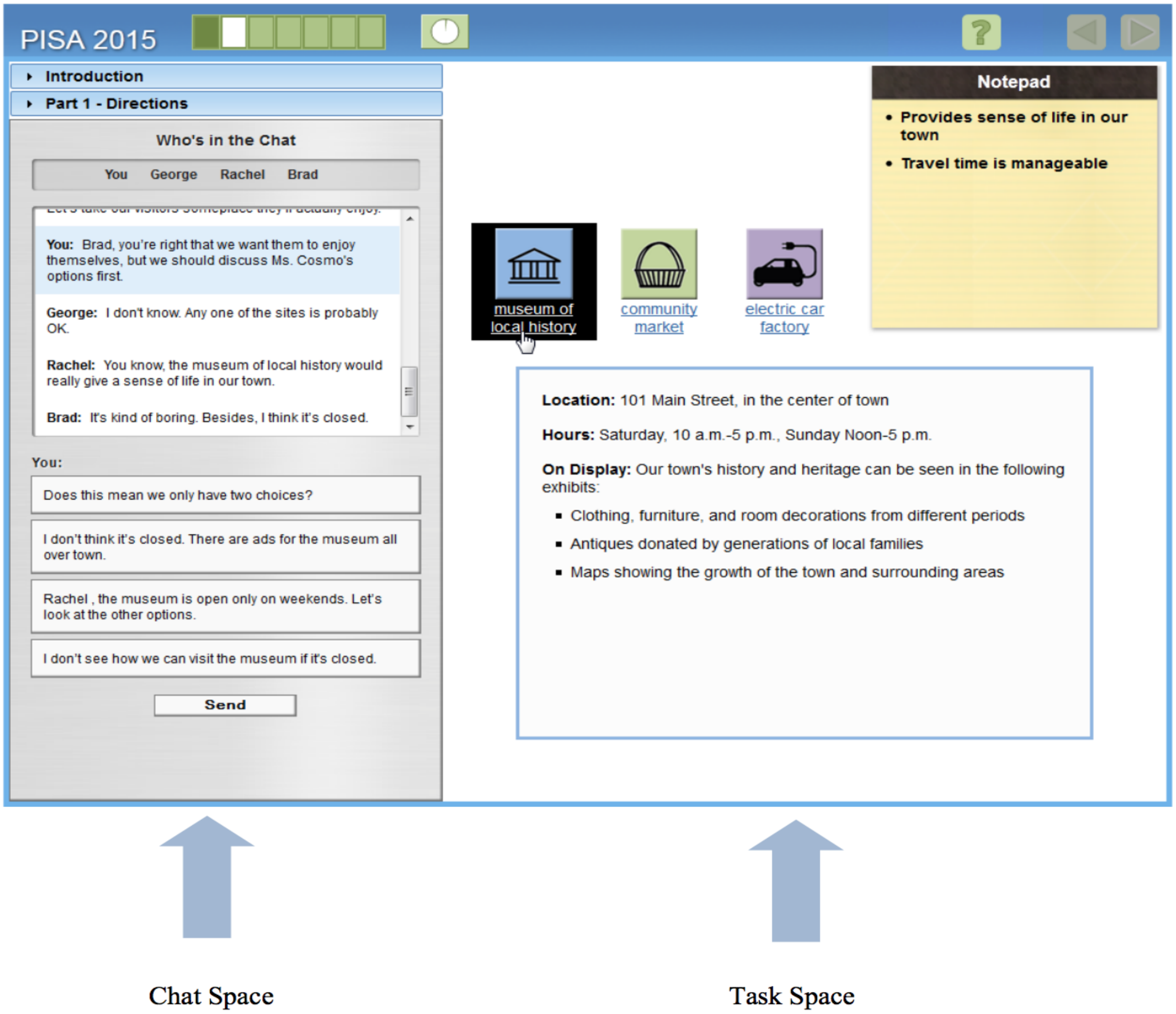
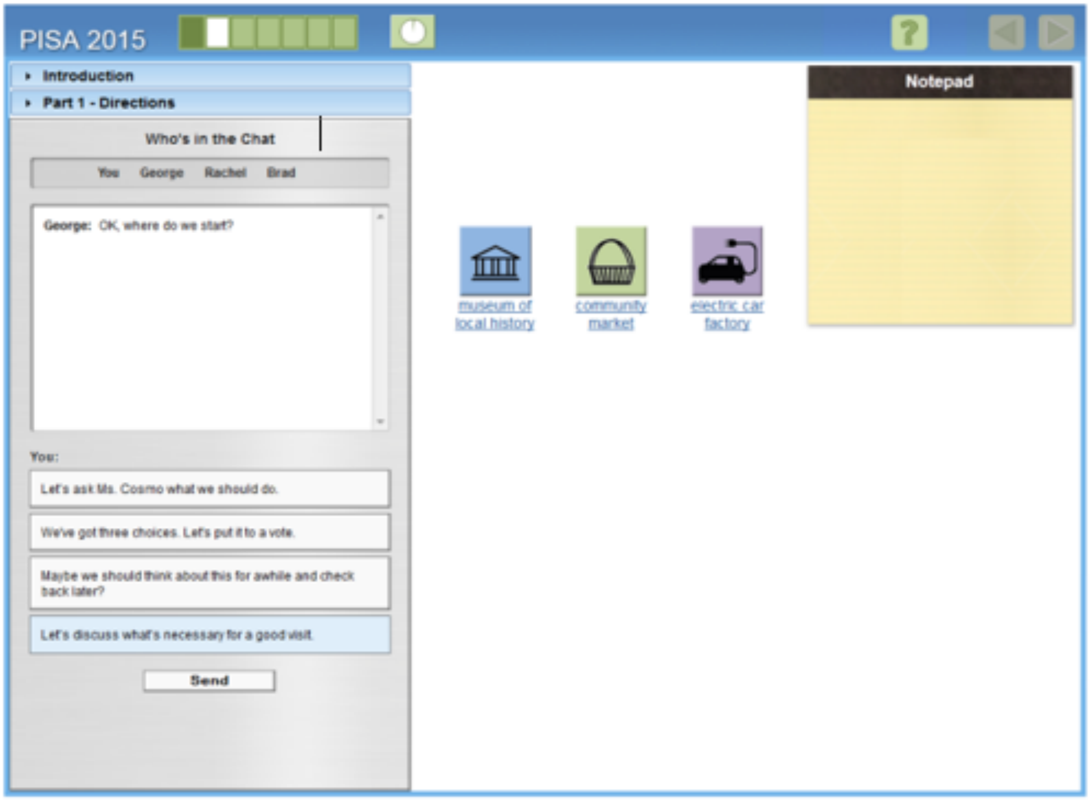
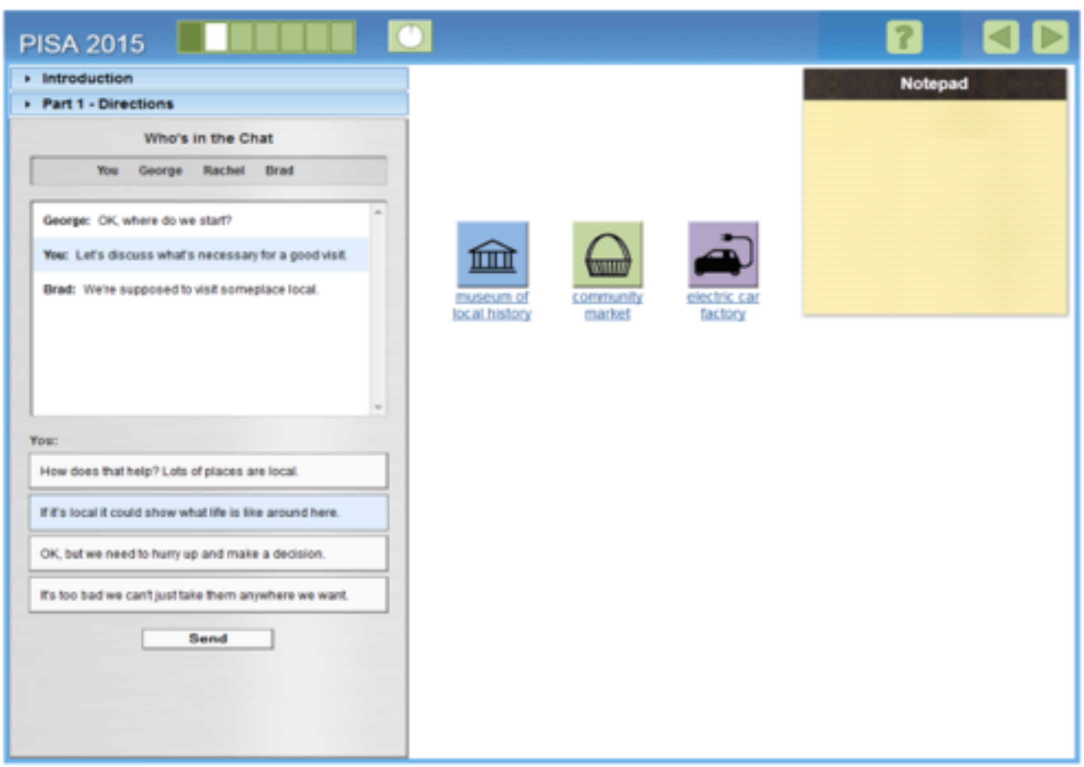


Exhibit 3.21. Example PISA Collaborative Problem-Solving Interaction



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. These competencies are assessed at varying levels of the collaborative skill (OECD, 2017).

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, and
- responding productively to others' mathematical ideas.

Collaborative mathematics items may assess one or more of these aspects as a student engages with others during the assessment (e.g., a human, a computer-based avatar, or named characters introduced in the item stem). Measurement targets are at the intersection of the assessed student's cognitive and social processes within a collaborative context. For the NAEP Mathematics Assessment, the collaborative process will most often begin for the assessed student after an initial presentation of mathematical context and content.

Features of items can include negotiating mathematical ideas through such activities as:

- expressing agreement, disagreement, or uncertainty;
- requesting clarification;
- elaborating on or revoicing others' ideas;
- identifying conflicts or gaps in mathematical thinking; and
- revising one's own thinking.

Negotiation may or may not entail conflict, but it does entail the processes through which team members accommodate and resolve differences on the way to coming to agreement (Dillenbourg & Baker, 1996; Fiore et al., 2017; Hesse, Care, Buder, Sassenberg, & Griffin, 2015). The agents involved in the collaboration coordinate their item- or task-relevant interactions, developing shared understandings, and constructing solutions.

Items can also involve establishing and maintaining team discourse relevant to the item or task at hand by (Flor, Yoon, Hao, Liu, & von Davier, 2016; Hao, Liu, von Davier, Kyllonen, & Kitchen, 2016):

- identifying goals;
- communicating next steps;
- evaluating teamwork; and
- checking understanding.

The structure of collaborative items or tasks allows for interaction between team members in a way that informs their thinking in item- or task-relevant ways (Dillenbourg, 1999). Individual thinking "can be inferred from the actions performed by the individual, communications made to others, intermediate and final products of the problem-solving tasks, and open-ended reflections on problem-solving representations and activities" (OECD, 2017, p. 135).

The assessed student and agents with whom the student engages can take on different roles in the collaborative process. A student may take on multiple roles within one interchange, depending on how an action is structured (Chiu, 2000):

- *Facilitator*: guides the group, helping to maintain focus and productivity
- *Proposer*: communicates claims
- *Supporter*: communicates agreement
- *Critic:* communicates disagreement
- *Recorder*: synthesizes group communications

For each of the three measurable collaborative skills described in this section, the table in Illustration 3.20 lists potential student actions and associated student roles.

Illustration 3.20. Potential Student Actions and Roles Associated with the Three Measurable Collaborative Mathematics Skills

Collaborative Mathematics Skill	Potential Student Actions	Potential Student Roles
Attending to and making sense of the mathematical contributions of others	Student asks the teammate to repeat a statement.	Facilitator
	Student asks the teammate to clarify a statement.	Facilitator
	Student rephrases/completes the teammate's statement.	Recorder
	Student identifies the goal of the conversation.	Facilitator/Recorder
	Student expresses confusion/frustration or lack of understanding.	Facilitator/Critic
	Student expresses progress in understanding.	Facilitator/Supporter
	Student checks on understanding.	Facilitator
Evaluating the mathematical merit of the contributions of others	Student expresses agreement with teammates.	Supporter
	Student expresses disagreement with teammates.	Critic
	Student expresses uncertainty of agreement or disagreement.	Supporter/Critic
	Student identifies a conflict in their own idea and the teammate's idea.	Critic
	Student uses relevant evidence to point out some gap in the teammate's statement.	Critic
	Student expresses what is missing in the teamwork to solve the problem.	Critic
	Student evaluates whether certain group contribution is useful or not for the problem solving.	Supporter/Critic
	Student points out some gap in a group decision.	Facilitator/Critic
	Student identifies a problem in problem solving.	Facilitator/Critic
Responding productively to others' mathematical ideas	Student elaborates on their own statement.	Supporter
	Student changes their own idea after listening to the teammate's reasoning.	Supporter
	Student suggests the next step for the group to take.	Proposer
	Student reflects on what the group did.	Recorder

Other aspects of the assessment of collaborative mathematics should be informed by results from the special studies described in Appendix E. These aspects include, but are not limited to, whether the agents with whom the assessed student engages are human or computer-based, process data collected, and methods of data collection.

Attending to and Making Sense of the Mathematical Contributions of Others

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.

Negotiation skills such as requesting clarification and revoicing others' ideas have been shown to be a sign of effective collaboration (Hao et al., 2016). 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/an avatar, or to justify its mathematical appropriateness.

The item in Illustration 3.21 is adapted from Exhibit 3.12 in the Framework (discussed previously in the justifying and proving section of this chapter). In both the original item and the adapted item, students are asked to make sense of the mathematical contributions of others as they evaluate the correctness of given statements. The adapted item includes an opportunity for students to consider original and revoiced statements. Also note that the names in the adapted item are different from 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 3.21. Collaborative Mathematics Example: **Revoicing**
adapted from **Exhibit 3.12**

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

Kala and Samir discussed the two questions shown.

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

Kala stated that the answers to the two questions are different.

Samir agreed with Kala, stating that the compound inequality $-3 < x < 5$ and the compound inequality $x > -3$ or $x < 5$ have different solution sets.

Explain how Kala's statement and Samir's statement are equivalent. You may use words, symbols, or graphs in your explanation.

The item in this illustration is adapted from a NAEP item. The original version of this item appeared in the 2013 grade 12 NAEP Mathematics Assessment with NAEP Item ID 2013-12M99 #1 M1934E1.

Evaluating the Mathematical Merit of the Contributions of Others

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).

Expressing agreement or disagreement is a negotiating skill associated with collaboration (Hao et al., 2016). 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. When answering this item correctly, the assessed student takes on the role of supporter, as the correct response indicates agreement with Connor's statement along with an explanation for the agreement (Chiu, 2000).

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 begins with a collaborative situation and 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 claim.

Exhibit 3.22. Adapted 2018 Grade 4 SBAC Number Properties Collaborative Mathematics Item

> Together, you and Connor are finding 8 × 16.
> Connor says, "We can find the product if we multiply 8 and 15 and then add 8."
> Which sentence could you say to Connor to **best** explain that his statement is correct or incorrect?
>
> A. I think you are incorrect, because we should add 16 instead of 8.
> B. I think you are correct, because 15 is an easier number to multiply by than 16.
> C. I think you are correct, because 8 × 16 is the same as 15 groups of 8, plus 1 group of 8.
> D. I think you are incorrect, because 8 × 16 is the same as 4 groups of 8, plus 4 groups of 8.

Another negotiating skill associated with collaboration is the use of relevant evidence to point out a gap in a teammate's statement (Hao et al., 2016). Illustration 3.22, based on Exhibit 3.23, shows another grade 4 item from the SBAC collection. Like the previous item, the item begins with a collaborative situation within which 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. In this item, the assessed student takes on the role of critic, as the item stem indicates disagreement with Jose's statement and the completion components of the item support the disagreement (Chiu, 2000).

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 complete a sentence that shows why Jose's statement is false. Although the item tells the student that Jose's statement is incorrect, the assessed student needs to understand Jose's statement before responding. The item also addresses the practice of justifying and proving, through the required completion of a counterexample to refute Jose's statement.

Illustration 3.22. 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	Justifying and Proving Collaborative Mathematics	Num – 5.b	SR – matching

You and Jose talk about the number of factors all whole numbers have.

Jose says that all whole numbers except 1 have an even number of factors because factors always come in pairs.

Jose's statement is incorrect.

Complete the sentences to help Jose see that his statement is not always correct. Drag numbers into the empty boxes to complete the sentences.

1 2 3 4 5 6 7 8 9 10	What about the number ☐? It has ☐ factors.

Scoring Information

Key	4; 3 OR 9; 3

ALD Notes for Item Developers
Notes for Basic and Advanced address both Collaborative Mathematics and Justifying and Proving.

Basic	The item could be revised to decrease collaborative demands. For example, the revision could provide a true statement about factors and ask for a revoiced version of the statement. The item could be revised to decrease justifying and proving demands. For example, the revision could 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 increase collaborative demands. For example, the revision could provide multiple statements from additional peers about factors across which making connections is required. The item could be revised to increase justifying and proving demands. For example, the revision could require the determination of the validity of Jose's statement along with completion of sentences that use the definition of factor to justify or refute Jose's statement.

The item in this illustration is based on an SBAC item with Item ID 3322, aligned to CCSS-M objective 4.OA.B.4.

The item in Illustration 3.22 involves the NAEP Mathematical Practices of Collaborative Mathematics and Justifying and Proving. As stated previously, the collaborative practice is often intertwined with other mathematical practices in the development of an item, but it should be possible to identify a primary practice focus. It is at the developer's discretion to determine which practice should be indicated as the primary practice. For this item, the practice focus is justifying and proving because the counterexample is fundamental to the completion of the item. The situating of the item in a collaborative context is not a requirement for arriving at the counterexample.

Consider, again, Illustration 3.13, 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. That is, the assessment happens in the context of examining the justifying activity of Pat. 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. Again, the assessed student has the opportunity to read or hear (through voiceover) Tracy and Pat's own utterances. This conversational format is preferable to items that might offer paraphrased positions that the assessed student is tasked to evaluate.

Responding Productively to Others' Mathematical Ideas

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 3.23a has potential, but, as written, does not assess the practice of collaborative mathematics. The established context is not inherently collaborative, except for the fact that students are asked to make sense of Mark's nonstandard first step and provide guidance for the next step.

While the item in Illustration 3.23a does not assess collaborative mathematics, it does assess the NAEP Mathematical Practice of Representing. In the item, students are presented with a symbolic representation of subtraction with which they engage 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 interpret presented representations.

Illustration 3.23a. 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

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 4 NAEP Mathematics Assessment with NAEP Item ID 2017-4M1 #8 M3744E0.

The item in Illustration 3.23b provides a revision that places the mathematical action in an inherently collaborative context. In this item, the assessed student takes on the role of proposer, offering a next step in the subtraction process (Chiu, 2000). The collaborative demands of the revision could be increased by programming a computer-based agent to offer steps in the subtraction process until the team of Mark and the assessed student arrives at the difference.

Illustration 3.23b. Collaborative Mathematics Example: Determining a Difference Collaboratively

You and Mark take turns suggesting steps to find the difference shown.

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

Mark’s first step is to subtract 43 from 143.

Based on Mark’s first step, what next step do you suggest? Explain why you suggest this step.

Scoring Information	
Key	Student suggests a valid next step with valid support. Examples of valid steps and possible supports could be: • Subtracting 43 from 48 so that the combination of Mark's step and this step makes an expression equivalent to the original expression. • 143 – 48 is the same as 143 take away 43 and then take away 5 more. Mark found the first part is 100, so to finish you have to take away 5 more, to get 95.

The practice of collaborative mathematics can also be assessed through a scenario-based task. Illustration 3.24 presents the outline of a collaborative task, built from a classroom-based situation that could also involve mathematical modeling. Following the suggested student actions for each item in the task, a component of collaborative problem solving is given (see Illustration 3.20 for student performance associated with the collaborative skills of negotiating ideas and regulating problem solving listed in Hao et al. [2016]). 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 3.24. Collaborative Mathematics and Mathematical Modeling Example: Outline of a Scenario-Based Task Situated in a Classroom Setting

> Start with a video clip or an avatar that sets up the task. Teacher shows students a transparent bin containing same-sized cubes. The bin cannot be opened by the students, but they can see that 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 left over? Students begin to converge around questions involving the number of cubes in the bin, and about ways of determining the number of cubes in the bin. As part of the collaborative process, the assessed student is included as one of the students in the classroom.
>
> Ideas for item content, below, reflect components of collaborative problem solving (Hao et al., 2016). Additional video or avatar communications should be provided between items.
>
> 1. The assessed student is asked to generate statements that define the goal of the task. (Student identifies the goal of the conversation.)
> 2. The assessed student is asked to provide a constructive response to a question about distributing cubes to groups. (Student suggests the next step for the group to take.)
> 3. After a process for distributing cubes to groups is determined, the assessed student is asked to describe what else the team needs to do. (Student expresses what is missing in the teamwork to solve the problem.)
> 4. When a classmate expresses frustration with the next step in the process of determining the number of cubes, the assessed student is asked to revoice the process to assist the classmate in understanding the process. (Student rephrases/completes the teammate's statement.)
> 5. The assessed student is asked to provide a representation for the number of cubes in the bin. When a classmate asks the student how that representation was determined, the assessed student provides support for the response by summarizing decision-making throughout the task. (Student reflects on what the group did.)

Developing items and tasks that involve the NAEP Mathematical Practice of Collaborative Mathematics 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 3.24 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 students' knowledge while they engage 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 the item being designed 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 E. 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 3.22* 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.	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, or orientation between figures 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. 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. First, 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 & Hennessy, 2008). Other potential threats to assessment validity are 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. Those chapters, combined with the guidelines in this chapter, focus on realizing the intent of the Framework in developing items used on the assessment.

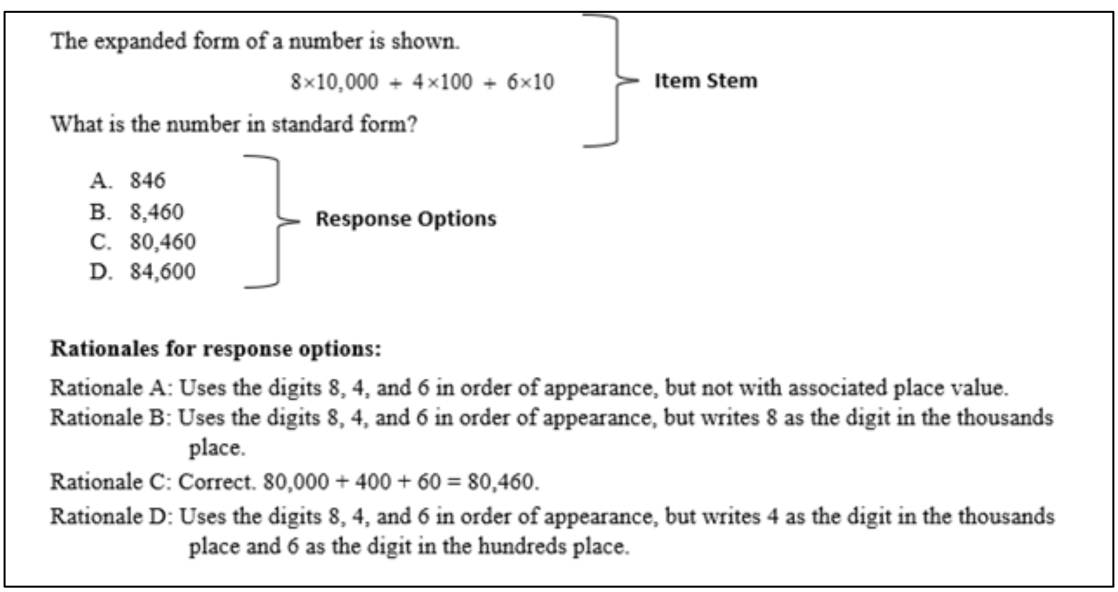
The guidelines offered here highlight only some of the critical considerations in item development, concentrating on topics specific to the NAEP Mathematics Assessment. Item writers should refer to directions for developing items provided by the Governing Board and its designees in addition to the information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

Item Characteristics

The specific components of an item are determined by the item format. 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, students. The item stem should provide all of the necessary information for students to respond, clearly laying out for the students what is being asked and the expected response method. The response method is determined by the item format.

Illustration 4.1 is 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 used 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?

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

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). The Board's policy includes 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 (Governing Board, 2014b, 2014c). Examples of scenario-based TEL tasks can be found at https://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 (more on these in the Item Types section, later in this chapter). 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 Exhibit 3.25 at the end of 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 mathematical rigor is maintained. 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 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 (Fawcett, Risko, & Kingstone, 2015). When multimedia content is included in a scenario-based task, developers need to ensure that the multimedia content 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), 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).

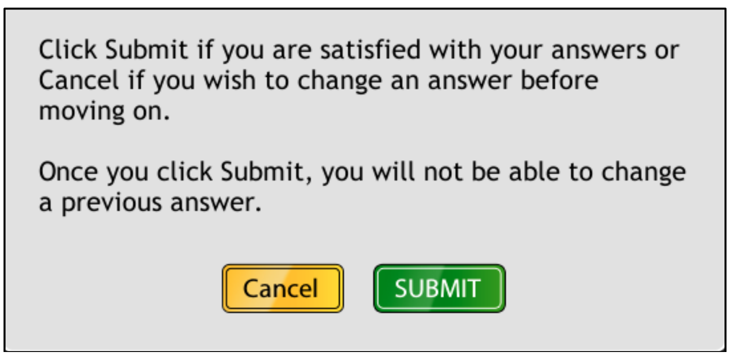
5. How many chairs can be in a stack and still fit in the storage room? ______ chairs
   Show how you figured it out.

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. When dependencies such as this occur, scoring needs to account for a correct answer to item 4 based on an incorrect answer to item 3. On a digitally based assessment, the task can 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. The revised digital version of item 3 from Exhibit 4.1 could read:

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.

Upon completion of the item, the student is notified that once the answer is submitted, it cannot be changed. The image below shows text displayed to students during administration of TEL tasks, which 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. For example, item 4 could be revised to read:

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.

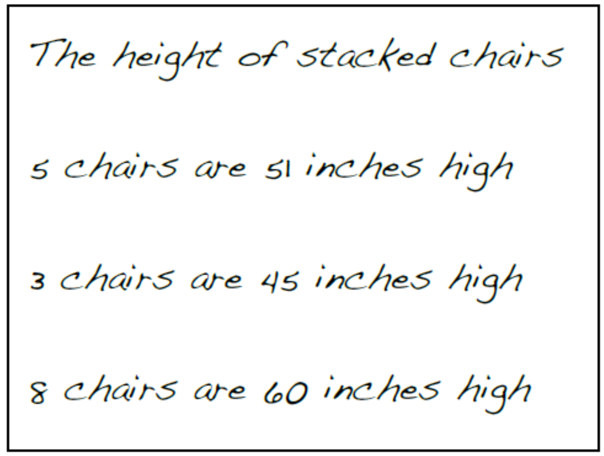
For additional examples that avoid dependencies between related item parts, see the TEL Andromeda Task: https://www.nationsreportcard.gov/tel/tasks/andromeda/.

A richer version of the *Stacking Chairs* task, as a scenario-based task, is provided in Illustration 4.2. The context puts the students in the task as part of a team determining whether chairs can be stacked in a storage room. Item text in Parts B, C, and D presents 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. Alternative *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.



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

Part B (Grade 8).
The team wants a way to determine the height of a stack of chairs when the number of chairs in the stack is known. 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 $h = 3n + 36$.*]*

Part B (Grade 12).
The team wants a way to determine the height of a stack of chairs when the number of chairs in the stack is known. 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.

[Scoring Information: 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 one 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* $h = 3n + 54$.*]*

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.]*

Part C (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, when each cart has the same number of chairs stacked on it, each cart will 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.]*

Part D (Grade 8).

The team will put 200 stacking chairs into 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

Part A (Grades 8 and 12).
Objective Alignment: Algebra, 1.a
NAEP Mathematical 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.

Part B (Grades 8 and 12).
Objective Alignment: Algebra, 3.b
NAEP Mathematical Practice Alignment: Abstracting and 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.

Part C (Grade 8).
Objective Alignment: Algebra, 4.a
NAEP Mathematical 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 is intended as a scaffold to the open-ended item in Part D. In Part C, students need to provide an explanation for how the equation could be used to determine the height of a stack of chairs. The height of one or more stacks of chairs is a component in the process of determining whether or not all of the chairs will fit in the storage room, which is the focus of Part D.

Part C (Grade 12).
Objective Alignment: Algebra, 4.c
NAEP Mathematical Practice Alignment: Justifying and 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.

Part D (Grade 8)
Objective Alignment: Algebra, 4.c
NAEP Mathematical Practice Alignment: Mathematical Modeling

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 request for constraints and limitations apply in general to the locating of chairs in the room and associates this item to component (b) of the mathematical modeling cycle (p. 126) as the responses serve to identify additional information needed to complete the task.

One of the affordances of scenario-based tasks is in the ability to leverage digital tools to make the content, and thereby the evidence produced, more accessible and authentic. For this particular task, the inclusion of a virtual measuring tape could allow students to measure heights of stacks of varying numbers of chairs. The measuring could be within a scale drawing context, or could allow for realistic measurements within a virtual environment. In either case, the student would need to line the measuring tool up properly to measure the height. To assist with this, the digital environment could have the measuring tape click in place (visually and audibly) so that a more accurate measurement could be made.

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 noted that TIMSS problem solving and inquiry tasks (PSIs), which have characteristics similar to NAEP scenario-based tasks, were challenging and time consuming to build (Mullis & Martin, 2016). Therefore, to aid in task development for NAEP mathematics, some illustrated suggestions are offered, building on existing NAEP TEL specifications and from existing NAEP items as starting points.

The 2014 TEL specifications (Governing Board, 2014c) suggested 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. An example from the *TEL Assessment and Item Specifications* is shown in Illustration 4.3a. An adaptation for NAEP mathematics item development 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?*

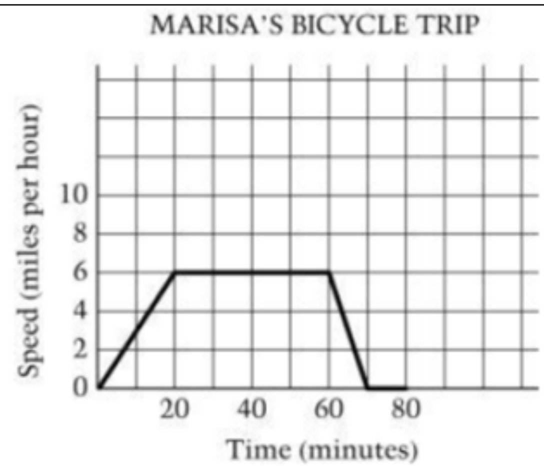
Illustration 4.3b. NAEP Mathematics Sample Scenario Shell

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 Practice(s)	*Which of the NAEP Mathematical Practices will be measured?*
NAEP Mathematics Objective(s)	*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 (Exhibit 3.4) and included again in Illustration 4.4 for reference.

Illustration 4.4. 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 *[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	interactive item component
NAEP Mathematical Practice(s)	Representing
NAEP Mathematics Objective(s)	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. Since 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 could continue to be refined (e.g., using a satellite version of a map where the student can visually see the topography and then draw the route Marisa took on the map, based on the graph; as the student draws the route, there could be a timer or clock on the side that adjusts as the route is drawn; changing the original problem from time to distance, then including an odometer on the side).

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, based on 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 in Illustration 4.6 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 make this task scenario-based.

In the draft item in Illustration 4.6, the focus on Algebra as a content area and the focus on 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* task 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. Content that is *not* appropriate to assess at grade 4 is as important to consider as content that is appropriate. 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	interactive item component
NAEP Mathematical Practice(s)	Representing
NAEP Mathematics Objective(s)	Number Properties and Operations 1.i – Order or compare whole numbers, decimals, or fractions using common denominators or benchmarks.

Note that as the scenario shell developed, the objective was changed from Algebra 2.a to Number Properties and Operations 1.i. This change stemmed from a desire to focus on a provided representation, instead of on translation between representations. To adapt for grade 4, a diagram might be presented showing 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 pair of locations, 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.

Alternatively, if preserving the alignment to Algebra 2.a is essential, the item could require students to match clips with actions, such as "Marisa is speeding up" or "Marisa has stopped." This revision would focus on translating between representational forms.

Bicycle Trip Example: Grade 12 Scenario-Based Task

To reimagine the task for grade 12, content that is not addressed at grade 8 but is 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 that 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	interactive item component
NAEP Mathematical Practice(s)	Representing
NAEP Mathematics Objective(s)	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.

Alternatively, students could view the clips and be asked to make their own graph of Marisa's trip, showing speed versus time. This revision would also focus on translating between representational forms, and, therefore, would also align to Algebra objective 2.a.

Identification and revision of a story concept foundation for a scenario-based task is likely to happen in parallel with the selection of target mathematics objective(s). 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). 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 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. 182).

What a Scenario-Based Task Is Not

The inclusion of multiple parts is not sufficient to make an item set 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 two parts. A correct response to each part requires use of the table 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 (also, there are no multimedia aspects and no tools enabled to solve the problem).

Illustration 4.9. Nonexample of a Scenario-Based Task

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

VISITORS TO THE MUSEUM AND PARK LAST WEEK

Day	Number of Visitors to the Museum	Number of Visitors to the Park
Monday	63	107
Tuesday	72	93
Wednesday	58	60
Thursday	81	88
Friday	95	64
Saturday	451	208
Sunday	407	225

The table shows the number of visitors who went to the museum and the number of visitors who went to the park on each day last week.

On Tuesday, how many more visitors went to the park than went to the museum?

visitors

On Saturday and Sunday combined, how many more visitors went to the museum than went to the park?

visitors

Scoring Information

Key	(a) 21 (b) 425

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 4 NAEP Mathematics Assessment with NAEP Item ID 2017-4M9 #3 M3461E1.

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 that 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 designing a content focus given the constraints of a framework, 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 work in evidence-centered design for assessment, design patterns leverage commonalities 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 the development of scenario-based tasks for the 2025 NAEP Mathematics Assessment.

Item Types

Since 1992, the NAEP Mathematics Assessment has used two types of items: 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 include those that require students to provide a text-based or numerical response. Both selected response and constructed response items may contain interactive item components (IICs). IICs may be embedded in an item (e.g., virtual ruler) or in the response field (e.g., number line).

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, which are frequently viewed as artificial but have high reliability, and traditional constructed response items, which 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 items and 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 mean 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

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.

- 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 stem of the item.
- 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).
- Zone: 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 task.

A new selected response item type included for the 2025 NAEP Mathematics Assessment involves the use of discourse and collaboration responses. Items of this type 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 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 responses 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(s) 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 3.22 Illustration 4.12
zone	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	Illustration 4.15

		menus that appear in various sections of an item.	

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 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 (e.g., when single-selection items on the same assessment 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 the multiple-selection format is more challenging than single-selection 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 identify and 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

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 4 NAEP Mathematics Assessment with NAEP Item ID 2017-4M1 #7 M3706MS.

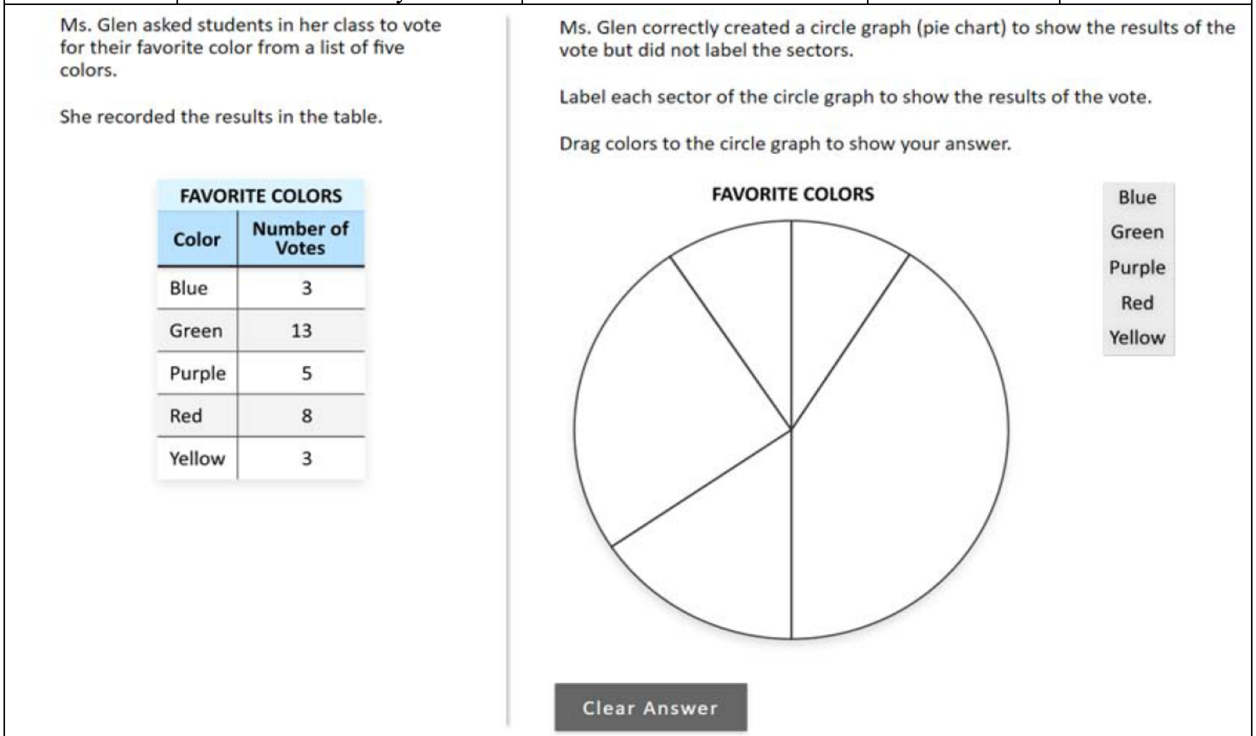
Matching. Matching items take many forms, but each involves the dragging and dropping of one or more objects. For example, a matching item may require the dragging of text, numbers, or figures into indicated spaces; the ordering of presented text, numbers, or figures; or the matching of a subset of objects 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. In addition to accessibility concerns (see p. 208 for more on accessibility), 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.

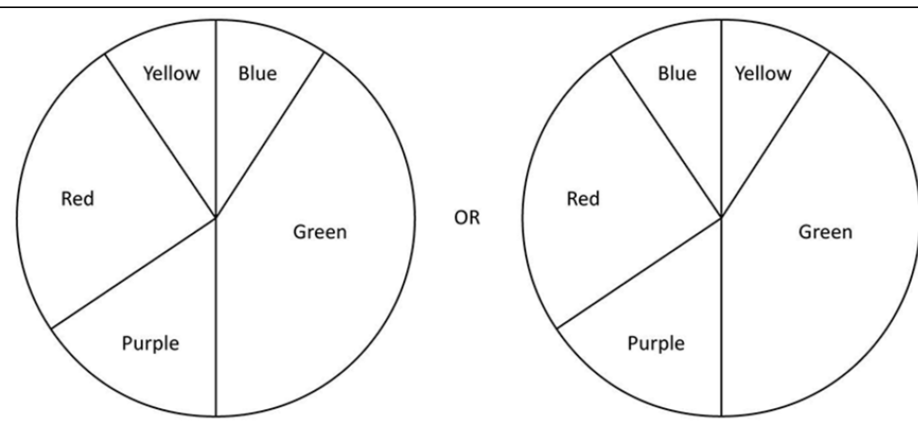
Illustration 4.12. **Selected Response Example:** **Matching Item**

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Representing	Data – 1.b	SR – matching



Scoring Information

Key



Correct	Five colors correctly placed
Incorrect	Incorrect response

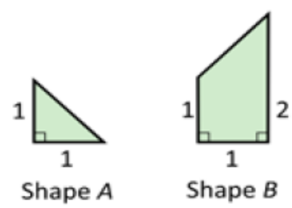
The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 8 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #2 M3806MS.

Zone. 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 encountered on some paper-and-pencil assessments, such as plotting a point on a number line. As with matching items, writers should consider the number and type of student actions required in light of accessibility and 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 select. 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 are used to clearly convey the size and shape of each figure.

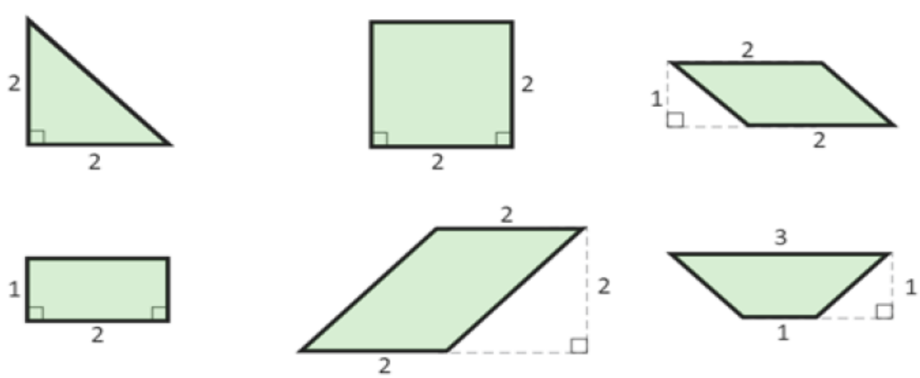
Illustration 4.13a. Selected Response Example: Grade 8 Zone Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Geometry	Other	Geom – 2.d	SR – zone



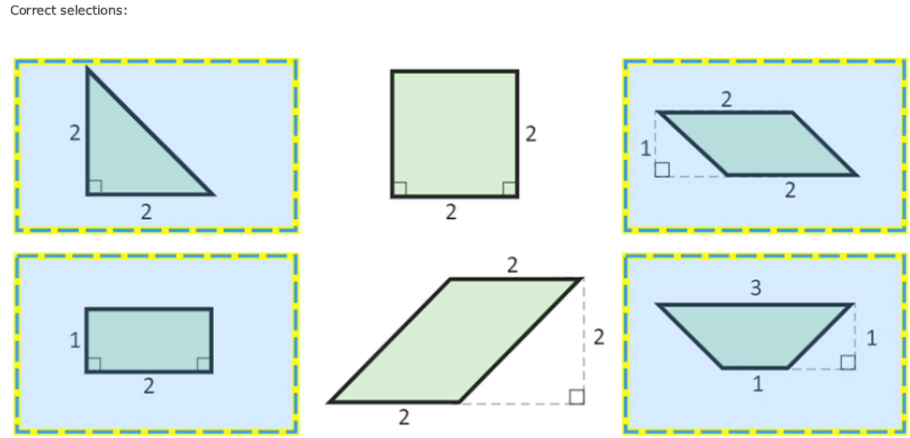
Which of the following figures can be formed by placing shape *A* adjacent to shape *B* so that they share a side?

Select all the correct answers.



Clear Answer

Illustration 4.13a. (continued)

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

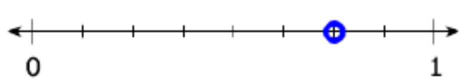
The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 8 NAEP Mathematics Assessment with NAEP Item ID 2017-8M3 #12 M3814EM.

The item in Illustration 4.13b presents a number line on which students can select a point. Although information regarding the numbers and locations 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: Grade 4 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	

The item in this illustration is based on a 2017 PARCC item with Item ID VF889661, aligned to evidence statement 3.NF.3a-2.

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 and type of student actions required in light of accessibility and 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

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 8 NAEP Mathematics Assessment with 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, in terms of accessibility, 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 the student must hear and manage.

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 option that completes the other statement.

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... ▾] the slope of Function B.

- greater than
- less than
- equal to

The y-intercept of Function A is [Choose... ▾] the y-intercept of Function B.

- greater than
- less than
- equal to

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.

The item in this illustration is based on a 2019 PARCC item with Item ID VH139356, aligned to evidence statement 8.F.2.

Constructed Response

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.

- 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, an analysis of a graph or table of values or an algebraic expression, 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 (new): Students respond by manipulating or using a physical object. The state of the object upon item completion is the response (see page 203 for additional details).

The table in Illustration 4.16 describes constructed response item formats, indicates other names by which an item format might be known, and provides the locations of exhibits and illustrations within this document 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 claim.</td></tr>
<tr><td>NAEP Item Formats</td><td>Abbreviations</td><td>Description</td><td>Location(s) of Example Item(s)</td></tr>
<tr><td>short constructed response</td><td>SCR</td><td>Ask 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.</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>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.</td><td>Illustration 4.18a
Illustration 4.18b</td></tr>
<tr><td>object-based response</td><td></td><td>Ask students to manipulate or use a physical object to provide a response.</td><td>Illustration 4.19</td></tr>
</table>

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 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, should depend on the mathematical construct 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 a scoring rubric as they are developing the item, so that both the item and the 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 (e.g., as outlined in the ALDs). Writers should sketch condensed sample responses for each score category, even before pilot use. Similarly, a mathematical justification or explanation for each category in a rubric 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 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 referred to as short constructed response (SCR) items. These are short-answer items that require students to give 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. SCR items may be scored as correct, incorrect, or partially correct, depending on the nature of the problem and the information gained from students' responses.

Most fill-in-the-blank (FIB) items with one response box are SCR items. FIB items 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). FIB items 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 for 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.

$228 \div 4 = \square$

Scoring Information	
Key	57
Correct	Answer of 57
Incorrect	Incorrect response

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 4 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 an FIB 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 if students had 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	

Illustration 4.17b. (continued)

<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>

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 8 NAEP Mathematics Assessment with NAEP Item ID 2017-8M9 #7 M3553E1.

Some 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). This item is presented again here with scoring information. Note that, like 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

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2009 grade 4 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, 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 that they take the student to respond and the time and effort that scoring the response takes. Extended constructed response items typically have five scoring categories: Extended, Satisfactory, Partial, Minimal, and Incorrect. In some cases, it may be appropriate to have four scoring categories for an extended constructed response item, depending upon the construct assessed and the nature of expected student responses to the item.

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 an FIB 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

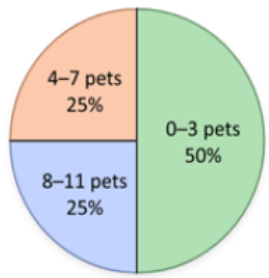
Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
8	Data Analysis, Statistics, and Probability	Representing	Data – 1.c	ECR – composite

A teacher surveyed 8 students in a class to find out how many pets they each have.

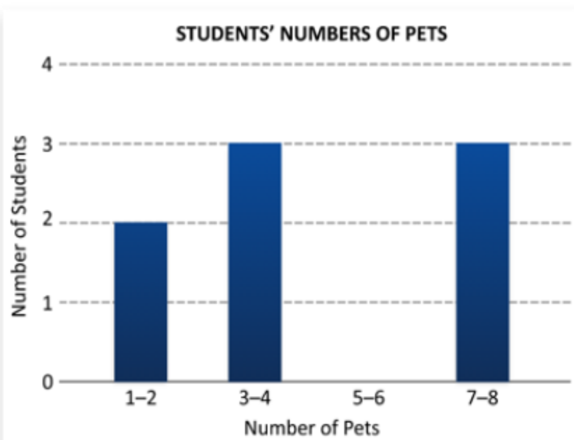
Boyd and Jenny represented the results of the survey in two different ways.

Boyd made a circle graph (pie chart) to show the results.

STUDENTS' NUMBERS OF PETS



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.

Illustration 4.18a. (continued)

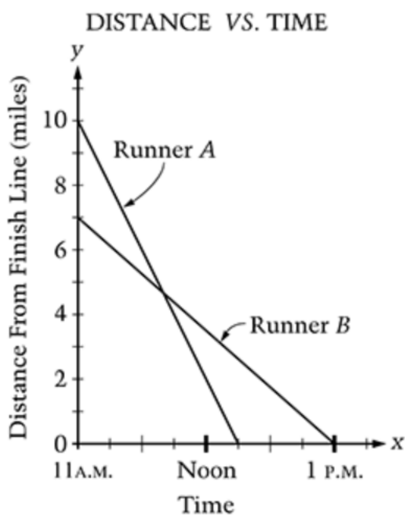
Scoring Information	
Key	(a) Correct selections: B. 0, 1, 2, 3, 5, 6, 9, 11 D. 1, 1, 1, 1, 4, 4, 8, 8 (b) Answer: 1, 1, 3, 3, 4, 7, 8, 8 OR 1, 2, 3, 3, 4, 7, 8, 8 OR 2, 2, 3, 3, 4, 7, 8, 8
Extended	Two correct selections and no incorrect selections for part (a) with a correct data set for part (b)
Satisfactory	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 Two correct selections and more than one incorrect selection for part (a) with a correct data set for part (b) OR One correct selection and one or more incorrect selections for part (a) with a correct data set for part (b) OR No correct selections 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 correct selection and no incorrect selections for part (a) with an incorrect data set for part (b)
Incorrect	Two correct selections and more than one incorrect selection for part (a) with an incorrect data set for part (b) OR One correct selection and one or more incorrect selections for part (a) with an incorrect data set for part (b) OR No correct selections for part (a) with an incorrect data set for part (b)

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 2017 grade 8 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 item in Illustration 4.18a, 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 – Composite



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 shown 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	

Illustration 4.18b. (continued)

<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 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 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>
</table>

The item in this illustration is based on a NAEP item. The original versions of this item appeared in the 2009 grade 12 NAEP Mathematics Assessment with NAEP Item IDs 2009-12M2 #9 M1809CL, M180901, M180902, and M180903.

Object-Based Response. 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 2025 and beyond 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 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, 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 aims to assess angle measurement, for which students have 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 NAEP Mathematics Assessment). In contrast, an item that asks students to represent the number 126 with base 10 blocks, where students manipulate physical "smart" base 10 blocks, would collect evidence that the student can represent a number in base 10. The submitted state of the base 10 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.

The item in Illustration 4.19 is based on a NAEP item that was accompanied by physical shapes which students manipulated to create their response. To respond to the item, students were asked to trace the shape that resulted from physically manipulating the provided pieces. The provided shapes in an object-based response item could be "smart" physical objects the students would manipulate and whose final state would be captured as the response to the item.

Illustration 4.19. Foundation for an Object-Based Response Item

Grade Level	Content Area	Assessed Practice(s)	Objective ID	Item Format
4	Geometry	Representing	Geom – 3.b	Object-Based Response
[Students are provided a set of physical shapes. Among them are congruent right triangles, each labeled *Q*.] You will need two shapes labeled *Q*. Please find those two shapes now. Use the two shapes labeled *Q* to make a square. [The functionality that allows the shape the student constructs to be captured will determine the completion of the item.]				

The item in this illustration is based on a NAEP item. The original version of this item appeared in the 1996 grade 4 NAEP Mathematics Assessment with NAEP Item ID 1996-4M10 #3 M061903.

Potential Scoring Advances

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 are 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). The Board's policy includes 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 can be used in a composite item. Some examples of composite items from this chapter are located in Illustration 4.9 which utilizes the FIB item format in each of the two parts, and Illustration 4.18a, which utilizes multiple-selection multiple choice and FIB 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* reflect 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 lengths 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 mathematics tools they will use during the assessment. The 2019 tutorials for each grade level can be found on the Internet at the following links:

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 multistate-consortium 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" also refers to the use of a digital emulator for calculation and 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) 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). In contrast, allowing for the use of a calculator when solving a multistep 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).

On-Screen Mathematics Keyboard

The item in Illustration 4.20 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 mathematics 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 mathematics keyboard so that the determined answer can be entered without indicating the number type for the correct response. The on-screen mathematics 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 mathematics keyboard at all three grade levels.

Illustration 4.20. Short Constructed Response Example: On-Screen Mathematics 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			

The item in this illustration is based on a 2016 PARCC item with Item ID M20834, aligned to evidence statement 7.SP.8a.

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 to 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.

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 students) and some English language learners (ELL students), 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 detailed information about accommodations, see the Governing Board's *NAEP Testing and Reporting of Students with Disabilities and English Language Learners Policy Statement* (2014a) 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 reproduces the distribution of items by grade and content area. See Chapter 2 for further details.

Exhibit 4.2. Percentage Distribution of Items by Grade and Content Area

Content Area	Grade 4	Grade 8	Grade 12
Number Properties and Operations	45*	20	10
Measurement	20	10	30
Geometry	15	20	
Data Analysis, Statistics, and Probability	5	20	25
Algebra	15	30	35

* 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 reproduced 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. Percentage Distribution of Items by NAEP Mathematical Practice

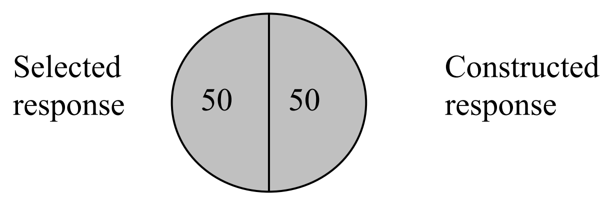
NAEP Mathematical Practice Area	Percentage of Items
Representing	10–15
Abstracting and Generalizing	10–15
Justifying and Proving	15–25
Mathematical Modeling	10–15
Collaborative Mathematics	10–15
Other	15–45

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 response types such as matching, zone, in-line choice, grid, and limited option responses. These items are machine scored. Constructed response includes short and extended constructed response. 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



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* (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 E.

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.

NAEP 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 A. 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 *Developing Student Achievement Levels for the National Assessment of Educational Progress Policy Statement* (2018a). The Governing 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 (Governing Board, 2017b) requires reporting according to various student populations (see section 303[b][2][G]), including:

a. Gender,
b. Race/ethnicity,
c. Eligibility for free/reduced-price lunch,
d. Students with disabilities, and
e. 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* (AERA, APA, & NCME, 2014) recommend 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. (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 in the following section.

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 to receive the greatest emphasis in the 2025 NAEP Mathematics Assessment's contextual questionnaires (in order of priority).

- *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 NAEP 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.
- *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.
- *Student mathematical identity*. Research demonstrates that students' perceptions of their mathematical identity directly relates to their mathematics learning. The 2025 NAEP Mathematics 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.
- *Instructional resources*. A range of resources influences 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.

- *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 on 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 on which 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.

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.

Achievement Level Descriptions (ALDs): 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 on the NAEP assessment.

CCSS-M: Common Core State Standards: Mathematics.

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.

Collaborative Mathematics: A NAEP Mathematical Practice that involves 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.

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.

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.

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.

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.

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* (Garfunkel & Montgomery, 2016). 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, or refuting mathematical claims in developmentally and mathematically appropriate ways.

Mathematical argumentation: The action or process of reasoning systematically in support of an idea, action, or theory.

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 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, Thames, & Phelps, 2008).

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, 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 Programme 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 NAEP 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 LEVEL DESCRIPTIONS

The NAEP 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 previous frameworks. The differences reflect not only changes to the mathematics knowledge, skills, and abilities assessed (mathematics content areas and mathematical 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 in each grade-level ALD.
- 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 previous 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 activities such as enacting knowledge of mathematical facts, using procedural fluency, and engaging in mathematical practices that are not included in the five identified for the NAEP Mathematics Assessment.
- Although Chapter 4 of the Framework provides examples of digital tools (e.g., graphing tools) that may be common in 2025 and beyond in schools, 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 have commonalities, and the content of each achievement level can be described generally. Descriptions at each achievement level, for all grade levels, are as follows:

- Descriptions at the *NAEP Basic* level focus on emerging understanding of grade-appropriate concepts and introductory engagement with mathematical practices.
- Descriptions at the *NAEP Proficient* level focus on application of grade-appropriate concepts and skillful engagement with mathematical practices.
- Descriptions at the *NAEP Advanced* level focus on extension of grade-appropriate 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 B, 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 strategies or solutions (e.g., justify solutions to word problems through numeric representations and operations); make mathematical sense of a problem scenario; select and use visual, physical, or symbolic representations, as needed, to lead to solutions; 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 strategies or solutions (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 solutions; 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 strategies or solutions (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 solutions; 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

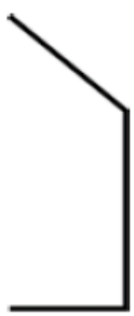
NAEP Basic, Grade 4

In this item, students are given an incomplete representation of a shape and asked to identify an associated complete shape, 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 multistep computation and interpretation within the context of the situation, addressing *NAEP Proficient* level language "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 each of items 1 through 4, refer to the following three figures.

Figure 1

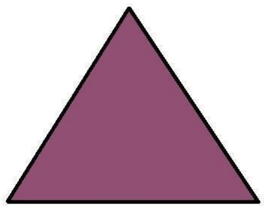


Figure 2

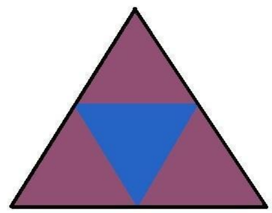
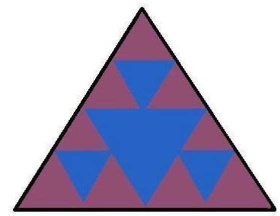


Figure 3



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-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-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.

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 multistep 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).

To assist item writers when coordinating decisions about item content and NAEP Mathematical Practices, the tables in this appendix include the objectives from Chapter 2 and, in an additional column, examples of the NAEP Mathematical Practice(s) likely to be assessed by each objective. The assignment of NAEP Mathematical Practices to objectives is based on the verb usage in the objectives and their alignment to the description of a particular practice or practices.

The listed "Inherent Practice(s)" are not definitive, nor are they intended as a boundary on how an objective and a practice might be assessed together. See Chapter 3 for detailed discussion and illustration of the NAEP Mathematical Practices.

Some cells in the "Inherent Practice(s) column" do not name a NAEP Mathematical Practice, and instead contain "Other" or "Variable."

- Other: The content of this objective lends itself to recall, procedural fluency, or mathematical practice(s) other than NAEP Mathematical Practices.
- Variable: More than one practice is likely to work well with the objective. The direction taken with the development of context, activity, prompts, or questions in an item aligned to the listed content objective would determine which, if any, of the NAEP Mathematical Practices is also assessed.

As in Chapter 2, the symbol * is used to identify 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) and # is used to indicate opportunities for a mathematical literacy focus.

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

Num – 1. Number sense (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
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
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

<table>
<tr><th colspan="4">Num – 3. Number operations</th></tr>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th><th>Inherent Practice(s)</th></tr>
<tr><td>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.</td><td>a) Perform computations with rational numbers.</td><td>a) Find integer or simple fractional powers of real numbers.</td><td>Other</td></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.</td><td></td><td>b) Perform arithmetic operations with real numbers, including common irrational numbers.</td><td>Other</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.</td><td></td><td>c) Perform arithmetic operations with expressions involving absolute value.</td><td>Other</td></tr>
<tr><td></td><td>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.</td><td>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.</td><td>Abstracting and Generalizing</td></tr>
</table>

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 rationals 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 instruments 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) Choose appropriate units for geometric measurements (length, area, perimeter, volume) and apply units in expressions, equations, and problem solutions.	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 or between measurement systems, given a relationship between the units.	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) Understand that numerical values associated with measurements of physical quantities are approximate, subject to variation, and must be assigned units of measurement.	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.	# f) Construct or solve problems 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 $\sin^2\theta + \cos^2\theta = 1$ for angles θ between 0° and 90°; recognize this identity as a special representation of the Pythagorean theorem.	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
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 problems.	# c) Represent problem situations with geometric figures to solve 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, quadrilaterals, other polygons; circles; 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

Geom – 3. Relationships between geometric figures (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
		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 a 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, 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 and Generalizing

Data Analysis, Statistics, and Probability

Data – 1. Data representation			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
Representations of data are indicated for each grade level in the next row. For some objectives only a subset of the representations is applicable, indicated by a parenthetical list at the end of 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.	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 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).	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/bimodal).	Representing
b) Given a distribution of whole number data in a context, identify and explain the meaning of the greatest value, of 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) 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

Data – 2. Characteristics of data sets (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
		# 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
		# 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 a 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) 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
	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

Data – 4. Probability (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
	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

Alg – 2. Algebraic representations (continued)			
Grade 4	**Grade 8**	**Grade 12**	**Inherent Practice(s)**
		e) Make inferences or predictions using an algebraic model of a situation.	Abstracting and Generalizing; Justifying and Proving; Mathematical Modeling
	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

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 that this vision is executed fully, NAEP should advance several strands of research and development, which correspond with the studies described in this appendix. 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. 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. 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 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.

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 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 scenario-based tasks 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 as well as through 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 scenario-based task on a NAEP Science or 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 scenario-based task 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.

Scenario-based tasks offer a promising avenue for measuring the NAEP Mathematics Practices, so the first phase of Study 1 will involve examining the measurement information that scenario-based tasks 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 mathematical practices through *response* data (i.e., students' scored responses to a discrete item or to the standardized group of items administered through an scenario-based task). If students' navigations through and interactions with these items' contexts are summarized only through response data, available measurement information may be under-utilized. 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 that 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, such as responding to a multiple choice item, or for relatively unstructured tasks, such as interacting with a digital environment). In fact, process data may increase the measurement information that can be gleaned from discrete items and scenario-based tasks, 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 (e.g., Andrews et al., 2017), 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 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*

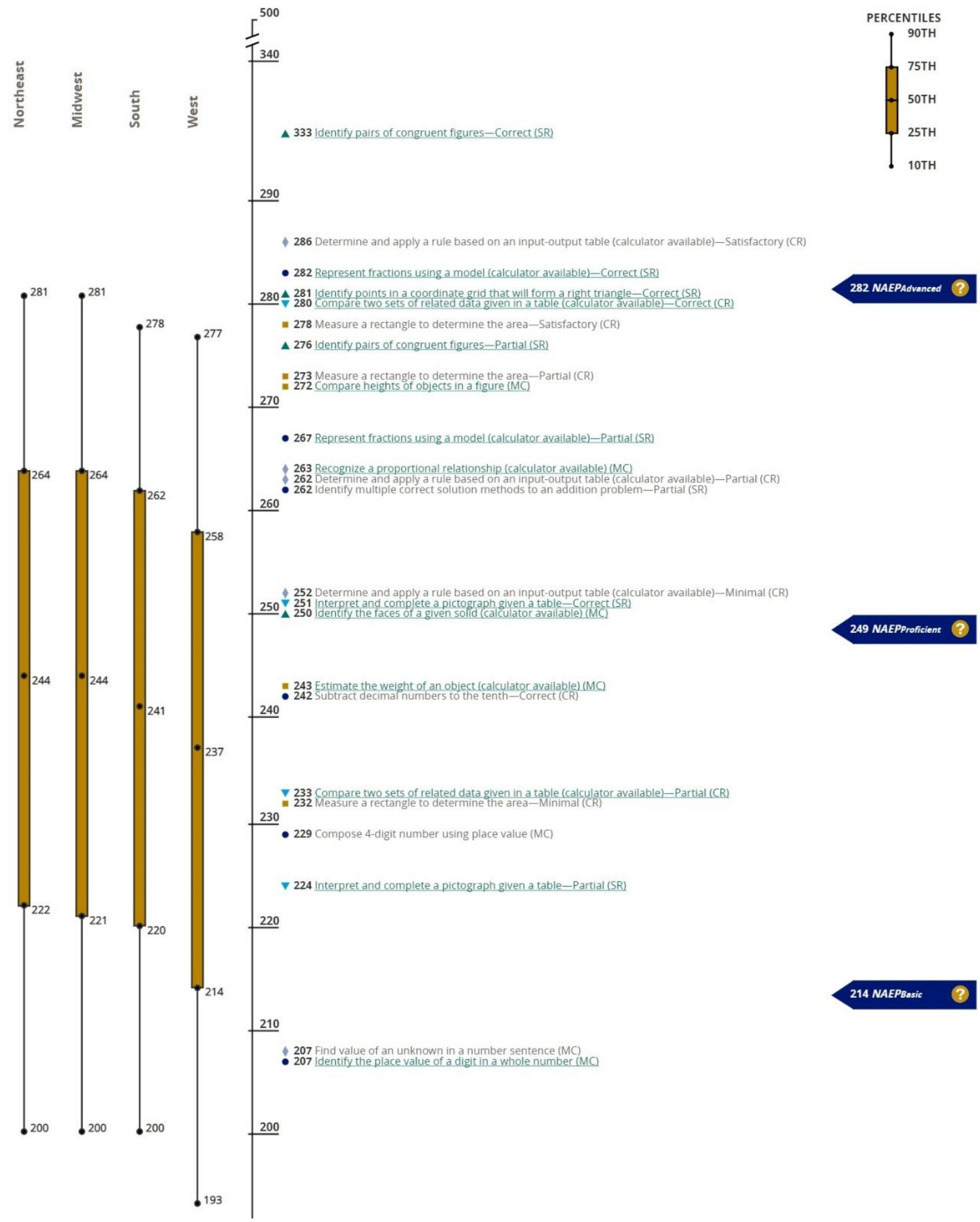
▲ **281** Identify points in a coordinate grid that will form a right triangle—Correct (SR)
▼ **280** Compare two sets of related data given in a table (calculator available)—Correct (CR)
■ **278** Measure a rectangle to determine the area—Satisfactory (CR)
▲ **276** Identify pairs of congruent figures—Partial (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)

Source: https://www.nationsreportcard.gov/itemmaps/?subj=MAT&grade=4&year=2017

Item maps can be augmented to summarize student performance or to 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. 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 E.1 and E.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



Source: https://www.nationsreportcard.gov/itemmaps/?subj=MAT&grade=4&year=2017&jurisdiction=NT&variable=CENSREG).

After a limited set of sample reports is created to represent 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 and simplicity, and on any areas that raise the risk of confusion or unintended interpretation. 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 that NAEP can use when considering the adoption of other large-scale programs' reporting methods. In addition, this study will produce one or more candidate reporting methods for consideration by other NAEP programs seeking to report on domains without scale scores.
2. NAEP will determine 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 that 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 of 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, particularly in grade 12. First, the Framework provides a definition of mathematical literacy:

> 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 number/hashtag sign (#), if there are everyday applications of the objective to situations in a person's life as a community member, citizen, worker, or consumer. These identifications 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 align 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 that 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 mathematics content area and each mathematical practice.
2. To signal the importance of mathematical literacy, NAEP will develop considerations for valid and straightforward reporting of mathematical literacy results.

APPENDIX F: PANEL PROCESS FOR NAEP CONTEXTUAL VARIABLES ANALYSIS

The 2025 NAEP Mathematics Framework Development Panel was charged with reviewing and making recommendations for changes to student, teacher, and school administrator surveys designed to measure the NAEP mathematics-specific contextual variables. Those recommendations appear in Chapter 5 of the 2025 NAEP Mathematics Framework. The goal of the review and recommendations was increasing the usefulness of these data for the interpretation of NAEP Mathematics results in 2025 and beyond. In making these recommendations, the Development Panel adhered to guidance from the 2025 NAEP Mathematics Framework Visioning Panel:

> A major goal of NAEP [reporting] should involve understanding NAEP results through analyses of contextual variables, including opportunities to learn . . . to emphasize the fact that what students know and can do is profoundly shaped by the social, human, and material resources available for their learning, the contexts in which they live and learn, as well as teacher and student orientations, motivations, and beliefs. (p. 7)

Overview

Item development for the 2017 NAEP Mathematics Survey Questionnaires addressed two areas of context: opportunity to learn and non-cognitive student factors. The broad categories of contextual factors within these areas were arranged around four issues:

- *Teacher preparation*, including content knowledge and subject-specific training, professional development, and non-cognitive teacher factors;
- *Student factors*, including mathematics activities outside of school, self-related beliefs, interest and motivation, grit for mathematics, and desire for learning for mathematics;
- *Resources for learning and instruction*, including people resources, product resources, and time resources; and
- *Organization of instruction*, including curriculum content, instructional strategies, use of technology in instruction, and use of formative assessment.

Importantly, attention to each of these issues and associated sub-categories was justified on the basis of its impact on student performance *specifically* in mathematics, its connection to possible interventions in and outside the classroom, and the extent to which information about it could appropriately be captured through survey questionnaires. For the 2025 NAEP Mathematics Framework update, the Development Panel was guided by the Visioning Panel's recommendation to use recent mathematics education literature on opportunity to learn and to explore the contextual variables associated with differing levels of performance among and within subgroups.

The Development Panel's subsequent review process entailed examining specific questions in the current student, teacher, and administrator surveys in each category and sub-category; analyzing these questions with respect to the extent to which they reflected recommendations from the Visioning Panel and implications of the current research on opportunities to learn; and then articulating recommendations for revising the focus of the NAEP mathematics-specifics surveys in order to strengthen what could be learned from the contextual variables data would support effective reporting of NAEP results.

Rationales for New Survey Emphases

In reviewing the current surveys, the Development Panel addressed not only the mathematics content students are taught, but also the opportunities students have to engage with that content in meaningful ways (NCTM, 2014). Do students engage in authentic mathematics activity focused on making sense of rich non-routine tasks? Do they have access to the kinds of tools that support the exploration of mathematical representations and how these connect to important mathematical ideas? Do they engage in meaningful discourse about their mathematical thinking and the mathematical thinking of others? In other words, to what extent do students have opportunities to engage with important *mathematics content and practices*, particularly the mathematical practices that are now recommended to be part of the NAEP framework? All of this colors how authentically students engage with opportunities to learn (e.g., Sword, Matsuura, Cuoco, Kang, & Gates, 2018). For these reasons, the Development Panel suggests addressing the extent to which students have opportunities to engage in authentic mathematical activities that provide opportunities to think and reason like mathematicians, both during mathematics instruction and also with regard to the nature of assigned homework (e.g., Rosario et al., 2015).

The Development Panel's review was informed by the importance of determining students' relationship to schooling and mathematics (e.g., Boaler, 2010; Strutchens, 2000). How do students think about the strengths they bring to their mathematical work, including the extent to which they have developed a strong mathematical identity and sense of agency? Do all students feel welcome in their mathematics classroom? Do their ideas matter to the teacher and their classmates? Are they seen as capable mathematical thinkers who have contributions to make during group work and whole class discussions? The answers to these questions shape how individual students see themselves as mathematical thinkers, inform the kind of mathematical identities individual students develop, and impact how deeply each student engages in the opportunities to learn that arise during classroom instruction (Steele, Spencer, & Aronson, 2002). For these reasons, the Development Panel suggests addressing *student mathematical identity* through questions addressing student participation in activities (such as discussion of mathematical ideas or evaluation of how a mathematics problem is framed), how students see themselves as mathematics learners, what they think it means to do mathematics, and what they think it takes to be a successful mathematics student.

In reviewing the literature and existing contextual issues, the Development Panel explored how to capture the extent to which students are given opportunities to draw on knowledge and skills acquired through mathematical experiences *outside* of the mathematics classroom as they engage with and make sense of activities *inside* the mathematics classroom (Aguirre, Mayfield-Ingram, & Martin, 2013; Boaler, 2002; Civil, 2007; Langer-Osuna, Moschkovich, Noren, Powell, & Vazquez, 2016; Lewis, 2014; Martin, 2006; Tomaz & David, 2015). Such funds of knowledge, acquired outside and used inside the classroom, have been defined as the skills and knowledge that have been historically and culturally developed to enable an individual or household to function within a given culture. This includes a wide range of ethnic and language communities, and importantly, such communities as street vendors, artists, video-gamers, and even the homeless who have established ways of interacting mathematically with their environments.

Recently, it also has come to include individuals with dis/abilities who have created their own ways of successfully mathematizing their environments (Tan & Krastberg, 2017). To what extent are these wide-ranging experiences leveraged to strengthen student opportunity to learn in the mathematics classroom? The panel suggests addressing questions about mathematical activity outside of the mathematics classroom, connections between what students are learning during mathematics instruction and students' experiential worlds outside of school, and *instructional organization and strategies* related to these (Crespo, Celedon-Pattichis, & Civil, 2017; Fernandez, Crespo, & Civil, 2017; White, Fernandez, & Civil, 2017).

The Visioning Panel recommended that teacher questions on the contextual surveys parallel many of the questions on the student surveys in order to be able to explore consistencies or inconsistencies across student and teacher perspectives. While acknowledging the limitations of the contextual surveys, the Development Panel also notes the importance of seeking ways to understand the answers to such comparative questions to situate reported results. What do students and teachers think it means to do mathematics, learn mathematics, and teach mathematics? How do teachers think about the extent to which there are opportunities to engage in authentic mathematical activity through meaningful contexts that draw out and build on students' funds of knowledge? How are teachers thinking about the mathematical strengths and mathematical identities students bring to learning? Similarly, it may be valuable for administrator questions on the school contextual variables survey to parallel questions on the teacher survey, given the roles of administrators as instructional leaders (Cobb et al., 2018; NCTM, 2014). To what extent do administrators and teachers share views on what counts as mathematical activity, including what it means to do mathematics, what it means to teach and learn mathematics, as well as questions addressing the development of mathematical identity and agency? What are the implications of teachers and administrators having different perspectives on these questions for NAEP score reporting?

One subcategory of resources for learning and instruction not specifically addressed in the NAEP Visioning Panel Guidelines document, but already present in NAEP surveys, is family perspectives on mathematics teaching and learning. Given the important influence of families on mathematics identity and agency, including the beliefs families hold about what it means to do and learn mathematics and who has the capacity to succeed in mathematics, the panel suggests exploring how students characterize how they and their families interact around what it means to do and learn mathematics (e.g, Aguirre et. al, 2013; Civil, 2007; Civil & Bernier, 2006; Martin, 2006).

Reflecting New Survey Emphases in New Categories

To first build an understanding of the context surveys, the Development Panel looked at the existing questions across grade levels and respondent groups (student, teacher, administrator) to determine where student, teacher, and school questions were the same and where there were differences. The next step was to cluster questions in ways that allowed an examination of the extent to which the collection of questions addressed issues of importance to the question of opportunity to learn. Finally, the third step was to look carefully across the student, teacher, and school questions to review consistency and appropriateness in how questions attended to opportunity to learn from student, teacher, and school perspectives.

As a result of this initial exploration of the mathematics-specific student surveys across grades 4, 8, and 12 the Development Panel organized existing questions into seven groups:

- Mathematics Content
- Student Engagement and Identity
- Views of Mathematics Teaching and Learning
- Features of Classroom Instruction
- Use of Technology
- Engagement in Mathematics Outside of School
- Perspectives on Family Beliefs About Mathematics Teaching and Learning

For the mathematics-specific teacher surveys for grades 4 and 8, the Development Panel identified six categories of items:

- Mathematics Content
- Views of Mathematics Teaching and Learning
- Features of Classroom Instruction Including Mathematics Teacher Learning and Support
- Use of Technology
- Student Engagement in Mathematics Outside of School
- Perspectives on Family Beliefs About Mathematics Teaching and Learning

For the mathematics-specific sections of the school surveys for grades 4, 8, and 12 completed by a school administrator, the Development Panel organized questions according to the following categories:

- School Mathematics Program
- Views of Mathematics Teaching and Learning
- Features of Classroom Instruction including Mathematics Teacher Learning and Support
- Use of Technology
- Student Engagement in Mathematics Outside of School
- Perspectives on Family Beliefs About Mathematics Teaching and Learning

As illustrated in Table 1 (next page), all of these categories align with the Development Panel's focus on opportunities to learn and are consistent with the four categories of issues identified in the 2017 NAEP Mathematics Survey Questionnaires. Also, as discussed in Chapter 5 of the framework, it is important to note that the panel recommends that decisions about contextual variables address categories in the following priority ordering (from highest priority to lowest priority):

- *Teacher preparation*, including content knowledge and subject-specific training, professional development, and non-cognitive teacher factors;
- *Student factors*, including mathematics activities outside of school, self-related beliefs, interest and motivation, grit for mathematics, and desire for learning for mathematics;
- *Resources for learning and instruction*, including people resources, product resources, and time resources; and
- *Organization of instruction*, including curriculum content, instructional strategies, use of technology in instruction, and use of formative assessment.

Exhibit F.1. Issues Addressed by Surveys and What They Include

	Resources for Learning and Instruction	Organization of Instruction	Teacher Preparation	Student Factors
Mathematics curriculum content	X	X	X	
Views of mathematics teaching and learning	X	X	X	X
Features of classroom instruction	X	X	X	X
Use of technology	X	X	X	
Student engagement and identity	X	X	X	X
Student engagement in mathematics outside of school				X
Family beliefs about mathematics teaching and learning				X

Attention to opportunity to learn requires foregrounding constructs such as student engagement and identity in the particular context of a mathematics classroom. In classrooms where there is limited student engagement in mathematical practices, including the NAEP Mathematical Practices, students have few opportunities to share and discuss their mathematical thinking, student perspectives are not likely to be valued and leveraged as important mathematical contributions, and the result is students being less engaged and less able to realize their full capacity to do mathematics. It is also the case that students can shift in their levels of engagement and sense of mathematical identity as they move from one classroom to another, depending on the norms that are in place that shape what it means to teach and learn mathematics, whose voices are heard, and whose mathematical thinking matters.

Asking similar survey questions across the issues is important, as many aspects of the issues identified cut across categories and sub-categories. For instance, to what extent do schools have the kinds of resources that allow teachers to engage students in mathematical activity that creates the opportunity for meaningful discussion and debate? To what extent is instruction organized in ways that allow students to meaningfully interact with classmates in small groups as they pursue solutions to tasks in ways that make sense to them? To what extent are teachers well prepared to plan and facilitate these kinds of instructional approaches, especially since these approaches often require a deeper understanding of mathematics and how students are likely to think about that mathematics? And to what extent do school administrators value and support this kind of mathematics instruction? Asking similar questions across the categories of issues may help uncover pieces that are in place and those that are not in ways relevant to reporting of NAEP scoring information.

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

Suzanne M. Wilson, Panel Chair*
Professor and Head, Department of Curriculum and Instruction, Education
University of Connecticut
Storrs, CT

Nadja Young
Senior Manager, Federal Government Civilian Projects
SAS Institute
Cary, NC

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
Assessment Specialist
Senior Math Content Specialist
WestEd
Columbia, MO

Elizabeth Dyer
Mathematics Content Associate
Research Associate
WestEd
Redwood City, CA

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

Fen Chou
Program Director, Standards, Assessment, and Accountability

Scott Norton
Deputy Executive Director, Programs

National Assessment Governing Board Staff

Michelle Blair
Project Officer
Assistant Director for Assessment Development

Sharyn Rosenberg
Assistant Director for Psychometrics

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., 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.

Aguirre, J., Mayfield-Ingram, K., & Martin, D. (2013). *The impact of identity in K-8 mathematics: Rethinking equity-based practices.* Reston, VA: National Council of Teachers of Mathematics.

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.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing.* Washington, DC: Authors.

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.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Barbosa, J. C. (2006). Mathematical modeling in classroom: A socio-critical and discursive perspective, *ZDM*, 293–301.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education (GAISE) report II*. Alexandria, VA, & Reston, VA: American Statistical Association &: National Council of Teachers of Mathematics.

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.

Boaler, J. (2002). Learning from teaching: Exploring the relationship between reform curriculum and equity. *Journal for Research in Mathematics Education, 33*(4), 239–258.

Boaler, J. (2010). *The elephant in the classroom: Helping children learn and love maths.* London: Souvenir Press.

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, NY: 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.

Chiu, M. M. (2000). Group problem-solving processes: Social interactions and individual actions. *Journal for the Theory of Social Behavior*, *30*(1), 27–49.

Civil, M. (2007). Building on community knowledge: An avenue to equity in mathematics education. In N. S. Nasir and P. Cobb (Eds.), Improving access to mathematics: Diversity and equity in the classroom (pp. 105-117), New York: Teachers College Press.

Civil, M. (2016). STEM learning research through a funds of knowledge lens. *Cultural Studies of Science Education*, *11*(1), 41–59. Retrieved from https://doi.org/10.1007/s11422-014-9648-2

Civil, M. & Bernier, E. (2006). Exploring images of parental participation in mathematics education: Challenges and possibilities. *Mathematical Thinking and Learning*, *8*(3), 309-330.

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.

Cobb, P., Jackson, K., Henrick, E., Smith, T.M., and the MIST Team. (2018). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Cambridge, MA: Harvard Educational Press.

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, NY: 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, A. R. (2016). *Psychometric properties of technology-enhanced item formats: An evaluation of construct validity and technical characteristics* (Doctoral dissertation). Retrieved from https://doi.org/10.17077/etd.922fbj4d

Crespo, S., Celedon-Pattichis, S., & Civil, M. (2017). *Access and equity: Promoting high-quality mathematics in grades 3-5.* Reston, VA: National Council of Teachers of Mathematics.

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.

Dillenbourg, P. (1999). What do you mean by 'collaborative learning'? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Oxford, England: Elsevier.

Dillenbourg, P., & Baker, M. (1996). Negotiation spaces in human-computer collaborative learning. In COOP Group (Ed.), *Proceedings of the International Conference on Cooperative Systems* (pp. 187–206). Juan-les-Prince, France: Author.

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, 24*(1), 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, NY: 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. B. (2011). Generalizing-promoting actions: How classroom collaborations can support students' mathematical generalizations. *Journal for Research in Mathematics Education*, *42*, 308–345.

Ellis, A., Bieda, K., & Knuth, E. (2012). *Developing 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, NY: State University of New York Press.

Fawcett, J. M., Risko, E. F., & Kingstone, A. (Eds.) (2015). *The handbook of attention*. Cambridge, MA: MIT Press.

Fernandez, A., Crespo, S, & Civil, M. (2017). *Access and equity: Promoting high-quality mathematics in grades 6-8.* Reston, VA: National Council of Teachers of Mathematics.

Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., Kyllonen, P., Massey, C., O'Neil, H., Pellegrino, J., Rothman, R., Soulé, H., & von Davier, A. A. (2017). *Collaborative problem solving: Considerations for the National Assessment of Educational Progress*. Washington, DC: National Center for Education Statistics.

Flor, M., Yoon, S.-Y., Hao, J., Liu, L., & von Davier, A. (2016). Automated classification of collaborative problem solving interactions in simulated science tasks. In *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 31–41). San Diego, CA: Association for Computational Linguistics.

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *The High School Journal, 91*(1), 29–42.

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.

Garfunkel, S., & Montgomery, M. (Eds.). (2016). *GAIMME: Guidelines for assessment and instruction in mathematical modeling education*. Philadelphia, PA: Society for Industrial and Applied Mathematics & Consortium for Mathematics and Its Applications

(U.S.), National Council of Teachers of Mathematics. 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, NY: Routledge.

González, N., Moll, L. C., & Amanti, C. (Eds.). (2005). *Funds of knowledge: Theorizing practices in households, communities, and classrooms.* New York, NY: 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.

Hao, J., Liu, L., von Davier, A., Kyllonen, P., & Kitchen, C. (2016). *Collaborative problem-solving skills versus collaboration outcomes: Findings from statistical analysis and data mining*. Paper presented at the 9th International Conference on Educational Data Mining, Raleigh, NC.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput, & 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.

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and* *teaching of 21st century skills: Methods and approach* (Vol. 2, pp. 37–56). Dordrecht, Netherlands: Springer.

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, NY: 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*. Chestnut Hill, MA, & Amsterdam, Netherlands: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, & International Association for the Evaluation of Educational Achievement (IEA), IEA Secretariat.

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. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 153–170). New York, NY: Routledge.

Küchemann, D., & Hoyles, C. (2009). From empirical to structural reasoning: Tracking changes over time. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 171–191). New York, NY: Routledge.

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.

Langer-Osuna, J. M., Moschkovich, J., Noren, E., Powell, A. B., & Vazquez, S. (2016). Student agency and counter-narratives in diverse multilingual mathematics classrooms: Challenging deficit perspectives. In R. Barwell et al. (Eds), *Mathematics education and language diversity*. Switzerland: Springer International.

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*. New York, NY: 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., & 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.

Lewis, K. E. (2014). Difference not deficit: Reconceptualizing mathematics learning disabilities. *Journal for Research in Mathematics Education, 45*(3), 351-396.

Levels of Conceptual Understanding in Statistics (LOCUS). (2019). *Question browser*. Retrieved from https://locus.statisticseducation.org/professional-development/questions/by-grade/grade-6?page=&type=prodev_const_response_question

Martin, D. B. (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.

Martin, D. B. (2006). Mathematics learning and participation as racialized forms of experience: African American parents speak on the struggle for mathematics literacy. *Mathematical Thinking and Learning, 8*(3), 197-229.

Martin, D. B. (2009). *Mathematics teaching, learning, and liberation in the lives of black children*. New York, NY: 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, & Amsterdam, Netherlands: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, and International Association for the Evaluation of Educational Achievement.

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 (Governing Board). (n.d.). *The nation's report card.* Retrieved from https://www.nationsreportcard.gov

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

Governing Board. (2007). *Assessment and item specifications for the NAEP 2009 Mathematics Assessment*. Washington, DC: Author.

Governing Board. (2014a). *NAEP testing and reporting on students with disabilities and English language learners policy statement*. Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/policies/naep_testandreport_studentswithdisabilities.pdf

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

Governing Board. (2014c). *Technology and engineering literacy assessment and item specifications for the 2014 National Assessment of Educational Progress.* Washington, DC: Author. Retrieved from https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/technology/2014-technology-specification-prepublication.pdf

Governing Board. (2017a). *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

Governing Board. (2017b). *The NAEP law.* Retrieved from https://www.nagb.gov/about-naep/the-naep-law.html

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

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*. Retrieved from https://nces.ed.gov/NationsReportCard/nqt

NCES. (2019). *Trends in International Mathematics and Science Study*. Retrieved from https://nces.ed.gov/timss/

National Council 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). *Principles 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. (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). (2003). Mathematical literacy. In *PISA 2003 assessment framework*. Retrieved from http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33707192.pdf

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, France: OECD Publishing, https://doi.org/10.1787/9789264305274-5-en.

OECD. (2019). *Programme for International Student Assessment*. Retrieved from 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.

Rosario, P., Nunez, J. C., Vallejo, G., Cunha, J., Nunez, T., Mourao, R., & Pinto, R. (2015). Does homework design matter? The role of homework's purpose in student mathematics achievement. *Contemporary Educational Psychology, 43*,10-24.

Saxe, G. B. (1988). Candy selling and math learning. *Educational Researcher, 17*(6), 14–21.

Scheiner, T., & Pinto, M. M. F. (2016). Images of abstraction in mathematics education: Contradictions, controversies, and convergences. In C. Csíkos, A. Rausch, & J. Szitányi (Eds.). *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 155–162). Szeged, Hungary: PME.

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.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York, NY: Academic Press.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & 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.

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, & G. Solano-Flores (Eds.), *Cultural validity in assessment* (pp. 19–37). New York, NY: 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.

Strutchens, M. E. (2000). Confronting beliefs and stereotypes that impede the mathematical empowerment of African American students. In M. E. Strutchens, M. L. Johnson, and W. F. Tate (Eds.), *Changing the faces of mathematics: Perspectives on African Americans* (pp. 7-14). Reston, VA: National Council of Teachers of Mathematics.

Sword, S., Matsuura, R., Cuoco, A., Kang, J., & Gates, M. (2018). Leaning on mathematical habits of mind. *Mathematics Teacher, 3*(4), 256-263.

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., Reckase, M., & Rowley, G. (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, Netherlands: International Association for the Evaluation of Student Achievement.

Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie & S. J. Derry (Eds.), *The computer as a cognitive tool* (pp. 229–258). Hillsdale, NJ: Lawrence Erlbaum Associates.

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 & National Council of Supervisors of Mathematics (NCSM). (2016). *Mathematics education through the lens of social justice: Acknowledgment, actions, and accountability*. Retrieved from https://www.todos-math.org/socialjustice

Tomaz, V. S., & David, M. M. (2015). How students' everyday situations modify classroom mathematical activity: The case of water consumption. *Journal for Research in Mathematics Education*, *46*(4), 455–496.

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, NY: 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, & 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.

White, D. Y., Fernandez, A., & Civil, M. (2017). *Access and equity: Promoting high-quality mathematics in grades 9-12.* Reston, VA: National Council of Teachers of Mathematics.

Wickstrom, M. H., & Aytes, T. (2018). Elementary modeling: Connecting counting with sharing. *Teaching Children Mathematics, 24*(5), 300–307.