TO: National Assessment Governing Board Members

FROM: Governing Board Staff

DATE: May 3, 2021

RE: Status Report and Next Steps for NAEP Reading Framework update

This memo repeats updates sent to Board members on April 23, 2021.

Background

The Governing Board has been undertaking updates to the NAEP Reading Framework for nearly two years. This process started with a unanimous March 2019 Board action to initiate an update to the framework, which provided specific guidance to a panel of subject-matter experts and NAEP stakeholders. An initial draft of the proposed framework update was released for public comment in June 2020. Using extensive public comment, engagement with stakeholders, and several Board policy discussions on the framework, the draft was revised and submitted for discussion at the March 2021 Board meeting.

Current Status

Several points have emerged in the Board's continued discussions on the Reading Framework update:

1) Board members are all committed to maintaining the level of quality and rigor that has long characterized the NAEP assessments.

2) While the initial charge to the panel may not have explicitly asked them to prioritize the stable reporting of trend, the Board is now prioritizing trend.

3) Consensus is emerging among Board members on several aspects of the proposed framework update (see Table 1 at the end of this memo). This includes proposed updates in the following areas:
   a. Comprehension Targets
   b. Disciplinary Contexts
   c. Purposes
   d. Text Types
   e. Text Source
   f. Text Format
   g. Text Complexity
   h. Language Structures and Vocabulary
   i. Reporting

4) There are two important areas where the Board has not reached consensus, which relate to the role of background knowledge (specifically, topical familiarity) in reading comprehension. These include:
   a. **Definition of reading:** There has been some concern that the framework updates change the construct that is being measured. Further Board discussion will attempt to clarify the construct in the existing framework (2004) and the construct as defined in the updated framework.

   b. **Universal Design Elements (UDEs):** Essentially, the question is about the degree of change in the framework update regarding UDEs and how they will be used in the

assessment. The Assessment Development Committee (ADC) is proposing to remove new knowledge-based UDEs that have to do with video, audio, and photo introductions to passages, i.e., not adding new knowledge-based elements, but retaining the ones already on the assessment such as written introductions. (See attached the latest revision of the framework update reflecting this removal.[1])

An optional Board webinar was held on April 30, 2021, to address Board member questions about how the current NAEP Reading Assessment relates to the latest draft of the framework. Presenters included NCES and representatives of the framework's Technical Advisory Committee, and more than half of the webinar allowed for Board members to discuss and ask additional questions.

The April 30 webinar was recorded for members who could not attend, and **the Executive Committee requests that all Board members either attend live or view the recording before the May 13 Reading Framework plenary session.**

Documentation in Response to Questions

Given some of the questions that have been raised so far, staff are providing the attached materials:

1) *Revised Draft NAEP Reading Framework Update* (updated 4/21/21), which appears in track changes. It reflects the removal of multimedia knowledge-based Universal Design Elements and clarifying edits relative to the draft reviewed at the March Board meeting, e.g., reducing academic language, removing redundancies, and improving graphics. (Please contact Michelle Blair to receive a copy that shows in track changes *all edits* to the framework since the public comment version was released last June.)

2) *NCES overview of current NAEP Reading Assessment* (updated since March 2021 meeting), which now includes additional information about the NAEP Reading Assessment over time (from paper to digital) and revised assessment development plans to further bolster the likelihood of maintaining trend.

3) *Redacted NCES report on evaluation of universal design elements in Scenario-Based Tasks (SBTs)* (new resource; underline>embargoed); this research has been referred to in several Board discussions, and key findings are summarized in the *NCES overview of the current NAEP Reading Assessment* (listed above). As secure material, ***this report can only be discussed in closed session***. It was provided under separate cover to Board members.

4) *NCES Questions and Answers on the NAEP Reading Framework update* (new resource), which responds to questions asked by several Board members.

5) *Guidance from Framework Technical Advisory Committee* (new resource), which was provided to the Framework Development Panel after the March 2021 Board meeting on key questions regarding topic knowledge in reading comprehension assessment constructs and Universal Design Elements.

---

[1] Multi-media introductions are removed from the framework update and will be proposed as an area for special study. Special studies are listed as part of the Assessment and Item Specifications, a separate document that is reviewed after a framework is adopted and specifically written for NCES implementation purposes.

6) *Results from a survey conducted by CCSSO* (new resource), which gathered more detailed information on states' approaches to topic knowledge.

Next Steps

The Executive Committee would like to enable full and productive deliberations on the Reading Framework at the May 2021 Quarterly Board Meeting. So, it is critical that all Board members review the above materials and raise any questions with ADC or staff so that all issues can be addressed to the extent possible.

TABLE 1: Similarities and Differences Between the 2009–2019 and 2026 NAEP Reading Frameworks

==Green== rows are areas of agreement, i.e., no remaining concerns have been raised by Board members.

==Yellow== rows are areas under discussion.

	Current Framework and Assessment	2026 Framework Update
Definition	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 making meaning with text, a complex cognitive process shaped by students' social and cultural influences. To comprehend, readers: ● Engage with text in print and multimodal forms; ● Employ personal resources that include foundational reading skills, language, knowledge, and motivations; ● Extract, construct, integrate, critique, and apply meaning in activities across a range of contexts.
Comprehension Targets	Locate and Recall Integrate and Interpret Critique and Evaluate	Locate and Recall Integrate and Interpret Analyze and Evaluate Use and Apply
Disciplinary Contexts	Literary Text Informational Text	Literature Contexts Social Studies Contexts Science Contexts
Purposes	Specific purposes communicated to students for scenario-based tasks in digitally based assessment as of 2017	Broad Purposes ● Reading to Develop Understanding ● Reading to Solve Problems Specific purposes for all assessment tasks are communicated to students
Text Types	Literary Texts Informational Texts	Literature Texts Social Studies Texts Science Texts

	Current Framework and Assessment	2026 Framework Update
Text Source	Authentic	Authentic except in rare instances
Text Format	Digital texts as of 2017 ● Static – non-moving print, graphics, or images on screen ● Dynamic – navigation across modes (print, video, other) or nonlinear locations (hypertext link)	Digital texts ● Static – non-moving print, graphics, or images on screen ● Expanded use of dynamic formats – navigation across modes (print, video, other) or nonlinear locations (hypertext link)
Text Complexity	Determined by: ● Expert judgment ● Passage length ● Two or more research-based readability measures	Determined by: ● Expert judgment ● Passage length ● Quantitative and qualitative research-based complexity measures
Language Structures and Vocabulary	Vocabulary assessed Potential for subscore	Language structures and vocabulary assessed No subscore
Universal Design Elements (UDE)	Digitally based assessment as of 2017 includes tools and support features: ● Highlighting and notetaking ● Text-to-speech on Directions and Help screens ● Zoom-in and selection of color schemes ● Sequential directions and transitions ● Look-back buttons to return to relevant section of text ● Graphic organizers ● Item foreshadowing ● Multi-part response frames ● Purpose statements ● Task characters (avatars that act as partners in simulated settings) ● Pop-up notes for definitions of vocabulary ● Resetting by providing correct response to answered questions ● Topic or passage introductions	Types of UDEs and possible examples: ● Task-based UDEs – Highlighting and notetaking – Text-to-speech on Directions and Help Screens – Zoom-in and selection of color schemes – Sequential directions and transitions for reading collection of texts – Look-back buttons to return to relevant section of text – Graphic organizers – Item foreshadowing – Multi-part response frames – **Samples of student writing as examples** ● Motivational UDEs – Explicit connections between broad and specific purposes – Task characters that provide oral or written directions, act as peers

	Current Framework and Assessment	2026 Framework Update

		or experts, or serve as an audience
		● Knowledge-based UDEs
		   – Text providing brief topic previews
		   – Pop-up notes for definitions of words or phrases
		   – Resetting by providing correct response to answered questions
		
		[Differences compared with current framework/assessment are listed in **bold** above; all others are already part of the assessment]

Reporting	Overall scale score and achievement levels (NAEP Basic, NAEP Proficient, NAEP Advanced) Disaggregation by gender, race/ethnicity, socioeconomic status, English learner status, state, region, type of community, public or nonpublic school, and literary and informational texts Data collected from student, teacher, and administrator questionnaires on contextual variables of interest Some data collected from students' test taking behaviors (process data) in digital administrations	Overall scale score and achievement levels (NAEP Basic, NAEP Proficient, NAEP Advanced) Disaggregation by all existing categories, adding ● Disciplinary contexts ● Socioeconomic status within race/ethnicity ● Former English learners (ELs) as well as current ELs and non-ELs Data collected from student, teacher, and administrator questionnaires on expanded set of contextual variables Data collected from students' test taking behaviors (process data) on expanded set of contextual variables

Reading Framework for the 2026 National Assessment of Educational Progress

*** *April 21, 2021 Draft* ***

National Assessment Governing Board
U.S. Department of Education

Visioning Panel
[* indicates the subgroup who drafted this framework as part of the Development Panel]

Peter Afflerbach*
Professor, Reading
University of Maryland
Silver Spring, MD

Carolyn Aguirre
Middle School Teacher / Department Head
New Haven Unified School District
San Leandro, CA

Sarah Aguirre*
Field Education Specialist
University of Texas, San Antonio
San Antonio, TX

Minerva Anaya St John
President
A-SJ Properties, Inc.
McAllen, TX

Nancy Brynelson*
Co-Director
California State University Chancellor's
Office. Center for the Advancement of
Reading and Writing
Gold River, CA

Jinghong Cai
Senior Research Analyst
National School Boards Association (NSBA)
Center for Public Education
Arlington, VA

Gina Cervetti*
Associate Professor, Education
University of Michigan
Ann Arbor, MI

Byeong-Young Cho*
Associate Professor, Korean Language
Education
Hanyang University
Seoul, Republic of Korea

Julie Coiro*
Associate Professor, Education
University of Rhode Island
Quaker Hill, CT

Carol Connor*
President, Society for the Scientific Study of
Reading /
Chancellor's Professor, University of
California, Irvine
Irvine, CA

Elena Forzani*
Assistant Professor, Literacy Education
Boston University
Boston, MA

Josephine Franklin
Associate Director
National Association of Secondary School
Principals (NASSP)
Reston, VA

John Guthrie*
Jean Mullan Professor Emeritus, Human
Development and Quantitative Methodology
University of Maryland, College Park
Chestertown, MD

Bonnie Hain*
Senior Director, Learning and Professional
Services
American College Testing (ACT)
Woodstock, MD

Robin Hall
Director, Language Arts and Literacy
Council of the Great City Schools (CGCS)
Fairburn, GA

Kathleen Hinchman*
Professor, Childhood and Adolescent Literacy
Syracuse University
Syracuse, NY

Christy Howard
Associate Professor, Content Area Literacy
East Carolina University
Raleigh, NC

Panayiota Kendeou
Guy Bond Chair in Reading / Professor
University of Minnesota
Minneapolis, MN

Emily Kirkpatrick
Executive Director
National Council of Teachers of English
(NCTE)
Louisville, KY

Carol Lee*
Edwina S. Tarry Professor, Education and
Social Policy
Northwestern University
Country Club Hills, IL

Karen Malone
Curriculum, Instruction and Assessment
Education Specialist
Bureau of Indian Education, Navajo District
Gallup, NM

Marina Pacheco*
Associate Professor, Curriculum and
Instruction
University of Wisconsin, Madison
Madison, WI

Cindy Parker
English Language Arts Collaborative Advisor
Council of Chief State School Officers
(CCSSO)
Danville, KY

Jim Patterson
Executive Director
The College Board
Coralville, IA

P. David Pearson, Panel Chair*
Professor Emeritus, Education
University of California, Berkeley
Berkeley, CA

Sue Pimentel
Founding Partner
Student Achievement Partners
Fort Myers, FL

Alicia Ross*
Teacher / Educational Consultant
Blue Ridge High School
Throop, PA

Robert Rothman*
Senior Editor
National Center on Education and the
Economy (NCEE)
Washington, DC

Allison Skerrett*
Professor, Curriculum and Instruction
University of Texas, Austin
Austin, TX

Eric Turman
Principal
Reading High School
Reading, PA

Paola Uccelli*
Professor, Education
Harvard University
Belmont, MA

Paul Wenger
President-Elect
National Association of Elementary School
Principals (NAESP)
West Des Moines, IA

Victoria Young
Director, Reading, Writing and Social
Studies Assessments
Texas Education Agency
Austin, TX

Technical Advisory Committee

Derek C. Briggs
Professor, Research and Evaluation
Methodology
University of Colorado, Boulder
Boulder, CO

Michael Kolen
Professor Emeritus, Educational
Measurement
University of Iowa
Estes Park, CO

Howard Everson
Senior Principal Research Scientist
SRI International
New York, NY

Scott Marion
Executive Director
The National Center for the Improvement of
Educational Assessment (NCIEA)
Dover, NH

Joan Herman
Senior Research Scientist, University of
California, Los Angeles /
Co-Director Emeritus, National Center for
Research on Evaluation, Standards, and Student
Testing (CRESST)
Los Angeles, CA

Jennifer Randall
Associate Professor and Director of
Evaluation for the Center for Educational
Assessment, Education
University of Massachusetts, Amherst
Amherst, MA

Kristen L. Huff
Vice President
Curriculum Associates
North Billerica, MA

Guillermo Solano-Flores
Professor, Education
Stanford University
Stanford, CA

WestEd Staff

Matthew Gaertner
Measurement Specialist
Director of Research, Assessment Research and
Innovation
WestEd
Austin, TX

Georgia Earnest García
Reading Content Specialist
Professor Emerita
University of Illinois, Urbana-Champaign
Napa, CA / Champaign, IL

Cynthia Greenleaf
Reading Content Specialist
Senior Research Scientist
WestEd
Albany, CA

Mira-Lisa Katz
Reading Content Specialist
Associate Director in Learning and Technology
WestEd
San Francisco, CA

Mark Loveland
Deputy Project Director
Senior Research Associate
WestEd
Redwood City, CA

Matthew Rudoff
Assessment Specialist
Manager, English Language Arts Assessment
WestEd
San Francisco, CA

Megan Schneider
Content Team Coordinator
Program Associate
WestEd
Redwood City, CA

Steven Schneider
Project Director
Senior Program Director, Science,
Technology, Engineering, and Mathematics
Research and Entrepreneurship
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

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 other subjects. The Nation's Report Card provides 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. By law and by design, NAEP does not produce results for individual students or schools. The National Assessment Governing Board (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 2026 NAEP Reading Framework describes the content and design of the 2026 NAEP Reading Assessment; it is intended for a general audience. A second document, the *Assessment and Item Specifications for the 2026 NAEP Reading Framework*, serves as the "test blueprint" with information about passage selection, item development and other aspects of test development; it is intended for a more technical audience, including NCES and the contractors that will develop the NAEP Reading Assessment. In accordance with Governing Board policy, the 2026 NAEP Reading Framework focuses on "important, measurable indicators of student achievement to inform the nation about what students know and are able to do without endorsing or advocating a particular instructional approach."

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 blueprints for the content and design of the assessments, and setting achievement levels. The NAEP Reading Assessment is given in English every two years to students in grades 4 and 8, and every four years to students in grade 12. The assessment measures reading comprehension by asking students to read grade-appropriate materials and answer questions based on what they have read.

Current NAEP Reading Assessment in a Digital Environment

The NAEP Reading Assessment has been administered on a digital platform since 2017. NAEP's move to dynamic and innovative technologies provides an opportunity for an engaging assessment experience for students and more meaningful data about students' skills and knowledge for educators. With digitally based assessments, students are asked to receive, gather, and report information just as they do in many aspects of their everyday lives. These assessments also are constructed to reflect the principles of Universal Design of Assessments (UDA) (National Center on Educational Outcomes, 2016). The principles of UDA are intended to increase assessment validity and accessibility and to provide a more accurate understanding of what students know and can do (Thompson, Johnstone, & Thurlow, 2002; Thompson, Thurlow, & Malouf, 2004). Examples of three of the seven UDA principles include precisely defined constructs, accessible, non-biased items, and maximum readability and comprehensibility.

The current NAEP Reading Assessment is organized according to assessment blocks. These feature either discrete items (stand-alone text passages and related questions) or scenario-based tasks (simulated settings in which students read passages while following various steps to accomplish a particular purpose or solve a problem). Scenario-based tasks (SBTs) can include many innovative features, such as:

- Task characters (avatars acting as simulated task partners)

- Increased guidance enabling students to navigate more complex items

- Item resetting in which students, after locking in answers, receive information about the correct response, so they can avoid carrying misconceptions into the next portion of the task

Schools and students participating in NAEP assessments are supported in various ways so they can successfully engage with the digitally based assessment. The digital platform provides students with support features that are intended to replicate the types of support provided during reading instruction and practice in school and at home or the workplace. For both discrete and SBT assessment blocks, tools available to all students include annotation via an on-screen pencil or highlighter, selection of color themes, and zoom-in. In addition, a text-to-speech capability is available on the Directions and Help screens (but not available for the reading passages or questions). Texts or questions may include hyperlinks, such as pop-up notes to click for more information (typically a definition of a selected word), a look-back button that takes students back to the relevant sentence or location in the text, multi-part response frames, and more. Not all support features are available in every block, but all blocks include some support features.

At the beginning of the assessment session, students interact with a tutorial that presents all the information needed to take the assessment on the digital platform; the tutorial explains how to progress through the reading passage and how to indicate or provide answers to questions, as well as how to use the tools. Students try out the tools and then enter and edit responses in a brief practice session. After the tutorial, students engage with two assessment blocks, each including one or more texts and approximately 10 questions. Texts may include images, graphics, or even a short video, and assessment items include both selected response and constructed response formats. The digital platform allows for a greater variety of formats, including selecting key words or sentences in a passage, dragging and dropping responses to complete a sequence or chart, completing a matrix or grid, and selecting more than one correct response. Hybrid items combine selected and constructed responses.

When students finish answering assessment questions, they participate in a digital survey, answering both general and reading-related questions. Student surveys collect demographic data and students' perceptions about access to technology and their reading habits and experiences in school, home, and the community. Together, the assessment blocks and survey take roughly 90 minutes. Teachers and administrators also complete surveys. Data collected as students navigate the digital assessment can provide valuable information about how students process texts and information during the assessment. For example, process data can reveal the time students take to read texts and respond to questions, how often they return to the text as they answer questions, and their use of optional digital tools.

While maintaining the essential structure and purpose of previous paper-and-pencil assessments, the development and implementation of digitally based assessments is key in maintaining NAEP's position as a leader in large-scale assessment.

Development of the 2026 NAEP Reading Framework

In 2018, the Governing Board conducted a review of the current NAEP Reading Framework. In accordance with the Board policy, the review included commissioned papers and discussions with an array of reading educators and experts. Based on the review, at its March 2019 meeting, the Governing Board determined that the Reading Framework needed updating to address advances in research in reading. The process of updating the 2026 NAEP Reading Framework was guided by Governing Board policies that specify that the work be undertaken by a Visioning Panel of educators; experts in reading, learning and development, and assessment; 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. In 2019, the Board charged the Visioning and Development Panels with developing recommendations for updating the framework 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.

To undertake this charge the Visioning Panel reviewed the considerable developments in reading research, literacy standards, and assessment that have taken place since the Board adopted the 2009–2019 NAEP Reading Framework in 2004. The Visioning Panel also considered input from a special panel of state literacy leaders as well as a paper, commissioned by NCES and authored by the NAEP Validity Studies (NVS) Panel, that examined the degree to which NAEP's assessments in mathematics, reading, and writing reflected both the content standards and the assessments implemented by ~~the~~ states. In this report, the NVS Panel recommended that NAEP "should continue to develop and implement reading blocks that use new formats similar to scenario-based tasks or other alternatives that prioritize purpose-driven, performance-oriented, multisource tasks" (Valencia, Wixson, Kitmitto & Blankenship, 2019). Accordingly, the Visioning Panel set guidelines for drafting an updated NAEP Reading Framework that would:

- Expand the construct of reading;
- Expand the definition of text;
- Extend the range of comprehension tasks that require knowledge application;
- Augment and expand the cognitive targets and the approaches 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.

At the heart of the Visioning Panel's guidelines was a commitment to equity, guided by two priorities in accordance with the most recent standards of fairness and equity in large-scale assessment to accomplish the following:

(1) Measure disparities in students' reading achievement in a way that minimizes test bias to the maximum extent (American Educational Research Association, American Psychological Association, and National Council of Measurement in Education, 2014; International Testing Commission, 2019; Task Force on Assessment of the International Reading Association, 2010); and

(2) Describe disparities in "access to resources and opportunities, including the structural aspects of school systems that may impact opportunity and exacerbate existing disparities in family and community contexts and contribute to unequal outcomes" in reading (the National Academies of Sciences, Engineering, and Medicine, 2019, p. 3).

The Visioning Panel thus wanted to ensure that updates to the 2009–2019 framework would enable students to draw on their accumulated knowledge and experiences to complete assessment tasks. To that end, the Visioning Panel asked the Development Panel to update the framework in a manner that would enhance the assessment's validity and fairness while minimizing bias. The Panel also called for assessment texts and tasks to be broadly representative of the knowledge and experiences of the nation's students and the many ways in which they engage with reading in today's world.

To address the Visioning Panel guidelines, the Development Panel considered frameworks for other large-scale literacy assessments, such as the Programme for International Student Assessment (PISA) and the Progress in International Reading Literacy Study (PIRLS). The Development Panel attended to educational and societal developments, including advances in technology and new types of texts (digital and multimodal), and they incorporated findings from new research in three areas: disciplinary literacy; the role of affect, motivation, and agency in shaping readers' performance; and the role of social and cultural experiences in human development and learning, particularly in reading comprehension. The Panel augmented its attention to principles of Universal Design of Assessments to address the experiences of the nation's increasingly diverse students in more inclusive ways, many states' recent adoption of new standards and assessments, and innovations in digitally based assessments. These broad developments in research, policy, and practice guided the drafting of this framework update for the 2026 administration of the NAEP Reading Assessment.

The Updated NAEP Reading Framework

This updated framework for the 2026 NAEP Reading Assessment addresses 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. The 2002 report of the ~~Rand~~RAND Reading Study Group identified three key components of reading comprehension—reader, text, and activity—and situated them in sociocultural contexts. The term sociocultural refers to the social and cultural features and practices of contexts, such as schools, homes, and communities, where students learn to read and engage in reading (Lee, 2020; Pacheco, 2015, 2018; Skerrett, 2020). This sociocultural perspective is important to reading comprehension assessment because it acknowledges that these practices influence how readers approach, engage with, and make meaning from texts (Mislevy, 2016; 2019).

Since the ~~watershed Rand~~RAND report, an even broader 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 lifelong learning (National Academy of Sciences, 2018). In this emerging consensus, learning—and reading—are still, at their cores, cognitive processes. However, cognitive acts, including reading, are influenced by the particular contexts in which texts are written and in which reading takes place.

The understanding of reading comprehension informing the 2026 NAEP Reading Framework is an outgrowth of earlier and current cognitively oriented work in reading comprehension (Anderson & Pearson, 1984; Kintsch, 1998; RAND Reading Study Group, 2002; Pearson, et al., 2020). Descriptions of the cognitive activities involved in constructing meaning have increasingly implicated social and cultural dimensions over time, dimensions that were also foreshadowed in NAEP reading frameworks adopted by the Governing Board in 1992 and 2004. Research evidence has highlighted that, like all human learning, reading comprehension is a meaning-making activity ~~imbued with~~that involves socially and culturally specific characteristics and practices~~.~~ (Bronfenbrenner & Morris, 2006; Lee, 2016b, 2020; National Academy of Sciences, 2018; Zelazo, 2013).

Drawing from previous frameworks and these research understandings, this updated NAEP Reading Framework attends to four key features of reading comprehension—contexts, readers, texts, and activities. The cognitive processes involved in reading are shaped by social interaction and mediated by many aspects of cultural practice, including the traditions and modes of speaking, that are part of students' daily lives (Nasir & Hand, 2006). At the heart of the 2026 NAEP Reading Framework is the definition of reading comprehension:

Reading comprehension is making meaning with text, a complex cognitive process shaped by students' social and cultural influences. To comprehend, readers:

- Engage with text in print and multimodal forms;
- Employ personal resources that include foundational reading skills, language, knowledge, and motivations; and
- Extract, construct, integrate, critique, and apply meaning in activities across a range of contexts.

Readers draw on a range of resources to make sense from text:

- What readers know about a topic;
- What readers know about texts and how they work;
- Internal processes, or foundational skills, needed to render text sensible, including phonemic awareness, letter-sound knowledge, and word- and sentence-reading skills;
- Higher order cognitive processes, such as attention, working memory, language comprehension, inferential reasoning, and comprehension monitoring; and
- Socially and culturally situated knowledge and practices from home, community, and school contexts.

The definition of reading comprehension included in the 2026 NAEP Reading Framework acknowledges and incorporates the cognitive roots of previous reading frameworks.

Also, the definition 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. It embraces the understanding that social and cultural practices also influence texts, including who reads and writes them and under what circumstances, how they are generated, how they appear, and how they are used. And finally, the definition emphasizes the integration of reading with other communication practices and the application of reading to tasks that address wide-ranging purposes and contexts.

Advances in measurement and in digitally administered assessment of reading comprehension, already initiated by NAEP in 2017, allow for a large-scale assessment that is more accessible to a greater number of individuals (National Center on Educational Outcomes, 2016). These advances have also allowed the assessment design to address the sociocultural aspects of the cognitive processes known as reading comprehension. Enacting the definition of reading comprehension in the 2026 NAEP Reading Assessment—described in this and subsequent chapters of the updated Framework—will enable NAEP to:

- Develop assessments with greater ecological validity (e.g., reading with purpose, applying what one learns from reading to a new task, benefiting from the presence of Universal Design elements that are typically available when reading outside of an assessment context);

- Draw on a greater range of texts and tasks representative of students' diverse experiences;

- Report on a broader array of the resources that students bring to bear in the act of reading (knowledge, language, motivations, prior experiences, agency, opportunities to learn); and

- Increase the precision of inferences about student reading achievement in the U.S.

Overview of the Updated NAEP Reading Framework's Key Components

The new framework maintains many aspects of the 2009–2019 NAEP Reading Framework. It also introduces some changes in the assessment design that are based on current scientific research in human development and learning, including reading comprehension. A continuing commitment to equity, non-biased and valid assessments, and the principles of Universal Design of Assessments were central to the updates in the 2026 NAEP Reading Framework. The advent of digitally based assessments in 2017 has allowed NAEP to provide an engaging assessment experience for students and explore new testing methods and question types. Framework updates also reflect trends in international reading comprehension assessments, such as the Programme for International Student Assessment (PISA) and the Progress in International Reading Literacy Study (PIRLS).

Comprehension Targets

Like its predecessors, the 2026 NAEP Reading Assessment engages students in reading texts and responding to questions that assess their comprehension of these texts. ~~The 2026 NAEP Reading Assessment invites students to read texts and respond to questions that assess their comprehension of these texts.~~ Comprehension Targets are used to generate test items that assess four important dimensions of reading comprehension. Three of these—Locate and Recall, Integrate and Interpret, and Analyze and Evaluate—are similar to the cognitive targets used in

the 2009–2019 Framework. One new target—Use and Apply—reflects a frequent and authentic purpose in disciplinary and workplace reading. Assessment of students' comprehension of vocabulary and language structures is systematically woven throughout the comprehension items.

Other Key Components

Disciplinary contexts for reading have taken on an expanded role in the 2026 NAEP Reading Framework to mirror the increased focus in schools on reading comprehension within disciplines, as well as in state standards and large-scale reading comprehension assessments. Two broad purposes for reading comprehension—reading to develop understanding and reading to solve a problem—will be delineated to systematically sample students' reading performance in literature, science, and social studies contexts. Texts, too, are sampled to address purposes within disciplines, affordances offered by digital and multimodal formats, and text complexity criteria for each tested grade. Finally, task-based, motivational, and knowledge-based Universal Design Elements are included as appropriate to support precise measurement of students' reading comprehension in ecologically valid ways.

Reporting 2026 NAEP Reading Assessment Results

Results of the NAEP Reading Assessment 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*). They are reported in the aggregate for the nation, states, and select large urban districts participating in the NAEP Trial Urban District Assessment; they are not reported for individual students, classrooms, or schools.

The 2026 NAEP Reading Framework updates the reporting system to emphasize equity, rigor, precision, and validity. The aim is to provide more nuanced reporting and useful data to key stakeholders across the nation. Currently, results of the NAEP Reading Assessment are disaggregated by gender, race/ethnicity, socioeconomic status, English learner status, state, region, type of community, public and nonpublic school, and literary and informational texts. Building on this system, the 2026 Framework proposes to disaggregate results by disciplinary contexts—literature, social studies, and science—rather than literature and informational texts. In addition, reporting categories are expanded to include (1) socioeconomic status within race/ethnicity and (2) former English (ELs) learners, in addition to current ELs and non-ELs, in order to describe student performance in more precise and detailed ways.

The framework also proposes to measure an expanded set of contextual variables, as is current practice, via student, teacher, and administrator questionnaires and by expanded the increased use of digital process data to provide further precision and explanation ofmore information on student performance. The contextual variables are clustered by two sets of reader characteristics: (1) cognition and metacognition and (2) engagement and motivation; and by two sets of environmental characteristics: (1) perceptions of school and community resources and (2) perceptions of teacher, instructional, and classroom supports. Ultimately, the framework envisions a reporting system that has enhanced explanatory capacity to assist educators in accessing, interpreting, and acting on the valuable information provided in NAEP reports and databases.

Comparison of the 2009–2019 NAEP Reading Framework and the 2026 NAEP Reading Framework

The framework for the 2026 NAEP Reading Assessment updates the framework developed and used for the 2009–2019 assessments. Building from this previous framework and on digital innovations, updates include:

- Expansion of the definition of reading comprehension to explicitly acknowledge the sociocognitive processes of reading. Reading comprehension is defined as making meaning with text and four key features are highlighted—contexts, readers, texts, and activities.

- Emphasis on three additional, research-based concepts: (1) how social and cultural experiences shape learning and development; (2) how reading varies across disciplines; and (3) the increasing use of digital and multimodal texts.

Key similarities and differences between the two frameworks are presented in exhibit 1.1. While updated, the continuity between the current framework and assessment and the 2026 NAEP Reading Framework is substantial.

Exhibit 1.1. Similarities and Differences Between the 2009–2019 and 2026 NAEP Reading Frameworks

	Current Framework and Assessment	**2026 Framework Update**
Comprehension Targets	Locate and Recall Integrate and Interpret Critique and Evaluate	Locate and Recall Integrate and Interpret Analyze and Evaluate Use and Apply
Disciplinary Contexts	Literary Text Informational Text	Literature Contexts Social Studies Contexts Science Contexts
Purposes	Specific purposes communicated to students for scenario-based tasks in digitally based assessment as of 2017	Broad Purposes ● Reading to Develop Understanding ● Reading to Solve Problems Specific purposes for all assessment tasks are communicated to students
Text Types	Literary Texts Informational Texts	Literature Texts Social Studies Texts Science Texts
Text Source	Authentic	Authentic except in rare instances
Text Format	Digital texts as of 2017 ● Static – non-moving print, graphics, or images on screen	Digital texts ● Static – non-moving print, graphics, or images on screen

	Current Framework and Assessment	**2026 Framework Update**
	● Dynamic – navigation across modes (print, video, other) or nonlinear locations (hypertext link)	● Expanded use of dynamic formats – navigation across modes (print, video, other) or nonlinear locations (hypertext link)
Text Complexity	Determined by: ● Expert judgment ● Passage length ● Two or more research-based readability measures	Determined by: ● Expert judgment ● Passage length ● Quantitative and qualitative research-based complexity measures
Language Structures and Vocabulary	Vocabulary assessed Potential for subscore	Language structures and vocabulary assessed No subscore

Universal Design Elements (UDE)	Digitally based assessment as of 2017 includes tools and support features:
	● Highlighting and notetaking
	● Text-to-speech on Directions and Help screens
	● Zoom-in and selection of color schemes
	● Sequential directions and transitions
	● Look-back buttons to return to relevant section of text
	● Graphic organizers
	● Item foreshadowing
	● Multi-part response frames
	● Purpose statements
	● Task characters (avatars that act as partners in simulated settings)
	● Pop-up notes for definitions of vocabulary
	● Resetting by providing correct response to answered questions
	● Topic or passage introductions
		Types of UDEs and possible examples:
		● Task-based UDEs
		  – Highlighting and notetaking
		  – Text-to-speech on Directions and Help Screens
		  – Zoom-in and selection of color schemes
		  – Sequential directions and transitions for reading collection of texts
		  – Look-back buttons to return to relevant section of text
		  – Graphic organizers
		  – Item foreshadowing
		  – Multi-part response frames
		  – ~~Student exemplars as mentor texts~~
		  – Samples of student writing as examples
		● Motivational UDEs
		  – Explicit connections between broad and specific purposes
		  – Task characters that provide oral or written directions, act as peers or experts, or serve as an audience
		● Knowledge-based UDEs
		  – Text~~, videos, or photographs~~ providing brief topic previews
		  – Pop-up notes for definitions of words or phrases
		  – Resetting by providing correct response to answered questions

	Current Framework and Assessment	**2026 Framework Update**
Reporting	Overall scale score and achievement levels (NAEP Basic, NAEP Proficient, NAEP Advanced) Disaggregation by gender, race/ethnicity, socioeconomic status, English learner status, state, region, type of community, public or nonpublic school, and literary and informational texts Data collected from student, teacher, and administrator questionnaires on contextual variables of interest Some data collected from students' test taking behaviors (process data) in digital administrations	Overall scale score and achievement levels (NAEP Basic, NAEP Proficient, NAEP Advanced) Disaggregation by all existing categories, adding ● Disciplinary contexts ● Socioeconomic status within race/ethnicity ● Former English learners (ELs) as well as current ELs and non-ELs Data collected from student, teacher, and administrator questionnaires on expanded set of contextual variables Data collected from students' test taking behaviors (process data) on expanded set of contextual variables

The remainder of the framework is organized to provide greater detail about the proposed content and design of the assessment and the reporting of results:

- Chapter 2 presents the **2026 NAEP Reading Assessment**, including the definition of reading comprehension and major assessment components.

- Chapter 3 describes the **Development of the 2026 NAEP Reading Assessment**, including specific design elements.

- Chapter 4 explains the **Reporting of NAEP 2026 Results**, including the expansion of reporting categories, contextual variables, and explanatory reporting capacity.

The 2026 NAEP Reading Framework recommends updates necessary to deliver assessments that are relevant, fair, and valid measures of student achievement in the U.S. The 2026 Framework builds on the current NAEP framework and operational assessment, especially the advances made possible by digitally-based assessment, by drawing on current understandings of reading comprehension and assessment. Chapter 2 provides a detailed description of the components that will be included in NAEP Reading assessments that students will take beginning in 2026. The chapter begins with the 2026 NAEP Definition of Reading Comprehension, ~~traces~~presents the definition's origins in policy and scholarship on reading comprehension, and ~~culminates in~~concludes with a description of the components of the assessment.

The NAEP Definition of Reading Comprehension

The 2026 NAEP Reading Framework attends to four key features involved in reading comprehension—contexts, readers, texts, and activities. The cognitive processes involved in reading are shaped by social interaction and mediated by many aspects of cultural practice, including the traditions and modes of speaking, that are part of students' daily lives (Nasir & Hand, 2006). At the core of the 2026 NAEP Reading Framework is the definition of reading comprehension:

> **Reading comprehension is making meaning with text,** a complex cognitive process shaped by students' social and cultural experiences. To comprehend, readers:
>
> - Engage with texts in print and multimodal forms;
> - Employ personal resources that include foundational reading skills, language, knowledge, and motivation; and
> - Extract, construct, integrate, critique, and apply meaning in activities across a range of contexts.

Key Terminology in the Definition

Each feature of the definition (contexts, readers, texts, activities) is important to understand how readers make meaning in the presence of texts.

Contexts. A central principle of the 2026 NAEP Definition of Reading Comprehension is that, as a human meaning-making activity, reading comprehension is situated within, and shaped by, social and cultural contexts. Social contexts, the settings within which individuals interact with one another, are governed by particular norms and expectations for the roles that different participants take up (e.g., student and teacher; youngest and eldest sibling). Social contexts are also inherently cultural. Cultural socialization occurs in classrooms, families, communities, and many other social contexts. With repeated ways of acting, interacting, knowing, believing, and valuing being passed down across generations all social groups develop cultures (Nasir & Hand, 2006).

Experiences students have in these contexts shape every aspect of reading comprehension: understanding of what to do, how to engage with text, and how to respond to and learn from reading. Contexts influence everything that readers bring to reading—including

the language, knowledge, motivations, and cognition that are acquired and refined in home, community, and school settings. Contexts shape the texts readers read. Although there is a common thread to the cognition involved in reading across contexts, much of the process of comprehension is influenced by context ~~and situated within particular settings and practices~~ (Scribner & Cole, 1981; Skerrett, 2020).

Readers. Each reader is a distinctive human being who brings a unique and diverse repertoire of cultural, cognitive (including metacognitive), motivational, and linguistic resources to every encounter with text. These resources are developed through experiences in multiple settings and communities and applied as readers make sense of text. For instance, first graders will use their knowledge of the stories they have listened to at home and in daycare settings to understand the stories they now have to read on their own. Adolescents in the U.S. would face a challenge when reading an unfamiliar text about the game of cricket in India, using their knowledge of other sports to make sense of the text. Bilingual readers often use what they know about reading in one language to read in another language (August & Shanahan, 2006; García & Godina, 2017). Readers' motivations and purposes are also impacted by their previous experiences and by the particular contexts in which the reading is being performed. They read to enjoy and be carried away by stories, to appreciate an author's use of language, to learn about themselves and the natural and social worlds in which they live, or to gather information and insight to act on the world. They read by themselves and with others; silently or orally; and lightly for a general impression or closely to prepare for a debate.

The Specialized Role of Readers' Knowledge. Many different kinds of knowledge play important roles in reading comprehension (Willingham, 2006). The categories of knowledge include world knowledge, knowledge of the topics of texts readers encounter, knowledge of text genres and structures, and linguistic knowledge, including vocabulary and syntax. In the process of extracting meaning, readers use this knowledge to clarify potential sources of ambiguities, including use of pronouns, words with multiple meanings, and ambiguous syntax. These forms of knowledge enable readers to make connections between adjacent ideas in texts even when authors do not make these connections explicitly. In more transparently construction-oriented processes, readers use knowledge to fill in gaps left by the author. Readers also use ~~frameworks of~~ knowledge ~~(e.g., a birthday party)~~ related to key ideas or themes in the text to construct mental models of meaning.

Of all of the types of knowledge involved in reading comprehension, the role of topic knowledge is probably the best understood. Contemporary cognitive models of reading describe the essential role of topic knowledge in text comprehension (Graesser, Singer, & Trabasso, 1994; Kintsch, 1998; McCarthy & McNamara, 2021; van den Broek, Risden, Fletcher, & Thurlow, 1996). These models represent the relationship between knowledge and comprehension as one in which existing knowledge is continually activated and integrated with textual information as readers develop a propositional understanding and, ultimately, a coherent mental representation of the text. Moreover, a large body of research has documented the impact of readers' topic knowledge and domain knowledge on reading comprehension across grade levels and text genres (e.g., Pearson, Hansen, & Gordon, 1979; Taft & Leslie, 1985; Alexander, Kulikowich, & Schulze, 1994). These studies also explain that while topic knowledge often influences readers' ability to recall information from text and to answer text explicit comprehension questions, the most consistent impact of topic knowledge is on readers' abilities to respond to questions that require bridging inferences (connecting information within texts) and more global inferences

(such as understanding concepts or themes). Readers may be generally skilled at such mental operations but not able to do so when texts focus on unfamiliar topics.

Texts. Texts are artifacts generated by authors to communicate their ideas. Texts take many forms, drawing on multiple genres and combinations of genres. They relay vastly different content to address many kinds of purposes. They draw on a wide array of modalities (e.g., static print, nonlinear hypertext, images, videos), sometimes combining modalities into multimodal forms (e.g., print with images or links to videos). They may be printed on paper or published in digital forms. They also differ in complexity, a term that usually refers to the density and nuance of texts' ideas and language structures.

Texts are composed according to conventions tied to cultural traditions and social practices. These traditions and practices are developed within and across such disciplines as literature, science, or history. Such conventions include genre traditions ~~of~~ favored by disciplines and modalities that are selected because of the ways they communicate certain kinds of ideas. Texts also vary in terms of the people, points of view, and experiences that are or are not represented. This means that texts may be readily understood by readers who find the ideas familiar or compelling but more challenging to others.

Activities. Activities include all the things readers do as they comprehend text and communicate and apply their understanding after reading. For example, readers *read the lines*, making sense of individual propositions in a text; they *read between the lines*, drawing inferences that connect ideas in one part of the text with ideas in another; and they *read beyond the lines*, using what they know to fill in gaps and draw more global meanings, such as themes and concepts. Evidence of comprehension-related activity comes from the things readers do to communicate and apply their understanding. For example, readers discuss their understanding of text and engage in activities in which they apply their understanding, such as preparing for a debate. They offer evaluations of texts, and they apply what they learn from their reading to solve problems and act in the world. They also use foundational skills, such as decoding, word recognition, and fluency (Vorstius, Radach, Mayer, & Lonigan, 2013). While these activities enable comprehension, they do not provide direct evidence of comprehension; thus, they are not directly assessed in the NAEP Reading Assessment.

Reading comprehension depends on who is doing the reading, what they are reading, why and where they are reading, how they have been prepared for the reading, with whom they are reading, and what schools and society will take as evidence of successful comprehension. Because all of these factors influence a complex process like reading comprehension, assessments must be sufficiently complex in their design and implementation (Mislevy, 2016).

Roots of the Definition

The NAEP Definition of Reading Comprehension and the resulting assessment are grounded in important developments in reading comprehension theory, research, practice, and policy over the three decades since the first NAEP Reading Framework was published in 1992. This definition draws on robust features from earlier NAEP reading frameworks and research describing cognitive processes involved in reading comprehension. It also attends to recent sociocultural understandings of learning and development, to disciplinary reading, and to an expanding conceptualization of what counts as text in today's society.

NAEP's definitions of reading comprehension in both the 1992-2007 Reading Framework and the 2009-2019 Reading Framework reflected dominant cognitive models of their times. The ~~Construction-Integration~~construction integration (C-I) ~~Models~~models proposed by theorists such as Kintsch (1998), Perfetti (1999), and van den Broek (van den Broek, Risden, Fletcher, Thurlow, Britton, & Graesser, 1996~~),~~) are still regarded as the most valid and useful cognitive accounts of reading comprehension. These models emphasize the multiple levels of meaning readers create, including a representation of the surface form that reflects accurate decoding; a text-base that includes all of the key ideas in the text plus the text-based inferences that link ideas within texts; and a situation model that represents the integrative links readers make between ideas expressed in the text and the knowledge they bring to reading.

Although earlier NAEP Reading frameworks were grounded in cognitive models of comprehension, they also acknowledged the importance of readers' purposes and the contexts in which they read and learned to read. In the first Reading Framework published in 1992, reading comprehension was defined as "… a complex process that involves an interaction among the reader, the text, and the context in which something is read" (p. 6). Purpose was mentioned when describing characteristics of good readers, who "can read a variety of texts for different purposes" (p. 9). The 2002 RAND Model of Reading Comprehension, which was heavily influenced by C-I models, was explicitly cited in the 2009-2019 Framework. Related to the features in the 2026 Definition of Reading Comprehension, the RAND model posited that reader, text, and activity reside in a sociocultural context, describing how "the identities and capacities of readers, the texts that are available and valued, and the activities in which readers are engaged with those texts are all influenced by, and in some cases determined by, the sociocultural context" (pp. 11-12). The 2009-2019 Framework also introduced the centrality of "using meaning as appropriate to type of text, purpose, and situation" (p. 3). The 2026 NAEP Reading Assessment will continue NAEP's longstanding focus on reading comprehension, rather than foundational skills or writing.

Updating the NAEP Reading Framework

The 2026 NAEP Reading Framework is updated to reflect three research-based developments that help to ensure that the NAEP Reading Assessment is a precise, fair, and accurate measure of reading comprehension. The first is how sociocultural experiences shape learning and development, including the learning and development of reading comprehension and, consequently, its assessment. The second is how reading varies across disciplines. The third ~~development is~~regards the ~~increasing~~ use of digital and multimodal texts.

Literacy scholarship has documented that cognitive actions associated with reading comprehension reflect the language and literacy practices (broadly, any activities through which students make and communicate meaning) of schools and communities (Frankel, Becker, Rowe, & Pearson, 2016; Heath, 1982; Lee, 2017; Scribner & Cole, 1981; Smagorinsky, 2001; Street, 1984), including disciplinary communities (Goldman, et al, 2016; Moje, 2007). This insight mirrors the broad consensus that has emerged across the learning sciences that learning is sociocultural in nature (Brown, Collins, & Duguid, 1989; Nasir & Hand, 2006). This finding is reflected in a 2018 report of the National Academies of Sciences, Engineering, and Medicine [NASEM]. The report explains that "each learner develops a unique array of knowledge and cognitive resources in the course of life that are molded by the interplay of that ~~learners'~~learner's cultural, social, cognitive, and biological contexts" (NASEM, p. 33).

This NASEM finding is also reflected in other large-scale assessments. PIRLS, the international assessment of reading for fourth grade students, notes that "social interactions about reading in one or more communities of readers can be instrumental in helping young students gain an understanding and appreciation of texts and other sources of information" (Mullis & Marten, 2021, p. 7). PISA, an international assessment for many subjects for 15-year-olds, similarly states that reading "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).

Scholars who study assessment closely (Greeno, 1998; Mislevy, 2016, 2019; Pellegrino, 2013) also note the importance of attending to contextual factors that shape student performance in any domain of expertise or learning. Measurement scholar Mislevy's (2019) summary of the implications of recognizing these factors for educational assessment is far-reaching:

> Situative, sociocognitive (SC) psychology is forcing a reconception of educational assessment. The SC perspective emphasizes the interplay between across-person linguistic, cultural, and substantive patterns that human activity is organized around and within-person cognitive resources that individuals develop to participate in activities. Rather than seeing assessment primarily as measurement, we are increasingly seeing it as an evidentiary argument, situated in social contexts, shaped by purposes, and centered on students' developing capabilities for valued activities... Implications follow for current challenges such as assessing higher order skills, performance in digital environments, and diverse student populations. (p. 164)

This perspective builds on longstanding understandings from scholarship in psychology and education. Over 30 years ago, Cronbach (1990) predicted that the psychology of individuals would have to take into account the highly contextualized framing of learning implied by Bronfrenbrenner's (1979) ecological approach. He noted that to fully understand individual development, psychologists and educators would have to engage in systematic analysis of the interactions among the attributes of students and the characteristics of the settings in which their learning is fostered and assessed. For many engaged in the study of assessment, a perspective that accounts for contextual facets of the assessment space is needed to assess more complex constructs. One of these complex constructs is reading comprehension, which can be assessed with greater relevance, precision, fairness, and validity by better reflecting contemporary understandings about the nature of the process.

A second update in the 2026 NAEP Reading Framework is the recognition of recent research demonstrating that reading and texts are shaped by disciplinary contexts. While a core set of academic literacy skills and strategies can be applied across areas of study, there are important differences in disciplinary reading practices. These include differences in the genres and discourse conventions and structures of texts, what counts as explanation, argument, and evidence, and the kinds of reasoning needed to formulate new understandings (Goldman, et al., 2016; Moje, 2007; Shanahan & Shanahan, 2008; Snow, 2010). These differences, which are related to the core activities in each discipline, require readers to employ different resources as they read and respond to text.

Also newly explicit in the 2026 Framework is recognition of the multimodal nature of texts used across all aspects of society. The widespread presence and rapid evolution of computers, smart devices, and software platforms have changed society's ideas about what

counts as text and its uses. Students read digital/multimodal texts in and out of school. Even though there is a common thread to reading in print and multimodal texts, there are also substantial differences, particularly around navigation (Coiro, 2020; Hartman, Morsink, & Zheng, 2010; Serafini & Gee, 2017). The implication is that the NAEP Reading Assessment must sample multiple modes of text.

These updates allow the 2026 NAEP Reading Framework to account more precisely for how well U.S. students comprehend what they read in texts and situations that more closely approximate reading practices in today's society. By building on past frameworks and research traditions while embracing more recent developments in assessment, NAEP ~~honors its mission of~~will continue to both ~~leading~~lead and ~~reflecting~~reflect reading assessment in the nation.

The NAEP 2026 Reading Assessment and the Definition of Reading Comprehension

The NAEP Definition of Reading Comprehension provides the foundation for how NAEP will assess reading comprehension. Each of the four aspects of the NAEP Definition of Reading Comprehension—contexts, readers, texts, and activities—is reflected throughout the 2026 NAEP Reading Assessment. The remainder of this chapter describes and explains key components of the NAEP Reading Assessment as well as their relationship to the definition. (See Exhibit 2.1.)

Components. The section begins with the core component of the assessment, the reading comprehension assessment items. After describing the items, the chapter takes on the challenge posed by Cronbach (1990) and Mislevy (2019), which is to address the variability inherent in complex domains of learning, including reading comprehension. ~~Five~~To that end, five additional ~~sets of new or~~ updated ~~assessment~~ components are ~~introduced~~also presented: disciplinary contexts, purposes, texts, universal design elements, and contextual variables. Taken together, these components ensure that NAEP will assess students' reading comprehension in ways that reflect the NAEP Definition of Reading Comprehension. ~~It also allows~~ and the natural variation that readers encounter in reading in home, school, community, and workplace settings. In this way, NAEP ~~assessment to account~~accounts for a wide range of factors that influence reading comprehension~~,~~ ~~mitigating~~ and mitigates potential bias that might result from a narrower operationalization of reading comprehension. ~~That is, building planned variation into every facet of the assessment provides opportunities for readers with varied backgrounds to find connections to their knowledge and experiences. Although it continues to be the case that students read the same texts and complete the same tasks and that their responses are evaluated in the same way, these assessment components help to create a more equitable standardized assessment.~~

Comprehension Items: The Role of Comprehension Targets

As in previous NAEP assessments, the 2026 NAEP Reading Assessment will engage students in reading sets of texts and responding to questions that assess their comprehension of these texts. Comprehension Targets are used in NAEP to generate the questions, ~~or test~~i.e., the assessment items, that students respond to as they take the test. Students' answers to these questions provide the observable data that NAEP uses to represent how effectively students engage in important comprehension processes, such as recalling texts and forming connections

among ideas within and across texts, when reading various kinds of texts. Three of the four targets~~,~~ *Locate and Recall*, *Integrate and Interpret*, *Analyze and Evaluate*~~,~~ are closely aligned with those in the 2009-2019 NAEP Reading Framework. ~~One~~An additional target, *Use and Apply,* ~~is an update that reflects~~has been added to reflect the importance of applying comprehension to new situations.

~~Although different, the~~Each comprehension ~~targets involve~~target involves inferences that readers tend to find more or less challenging in general~~, items~~. Items based on each target will range in difficulty, depending on the particulars of the questions in relation to the texts they are designed to probe. Building on the attention to vocabulary in the 2009-2019 Framework, the 2026 assessment also attends to structures of language within ~~the~~each comprehension ~~targets~~target.

Locate and Recall. The first ~~set of~~ Comprehension ~~Targets~~Target is Locate and Recall. In order to comprehend, readers need to identify important information and form connections among ideas in the text as they move through it. 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".

Items assessing the Locate and Recall ~~targets~~target typically focus on information stated directly in a single location in a text, such as a sentence, a paragraph, adjacent paragraphs, or a single graphic. However, in some cases, readers may need to navigate across different pages or documents, including hyperlinked and multimodal texts, to find additional information that is relevant to the test item. Test items might ask readers to recall or locate specific information about characters or settings in a story; or to locate a specific piece of information from a table in an expository text. Locate and Recall items can also require readers to form connections across text segments that are near one another in the text, such as fairly straightforward inferences about the relationships between ideas presented in adjacent sentences (e.g., A caused B or A occurred before B). Finally, readers may be asked to infer the meanings of unfamiliar words using information in the sentences immediately surrounding that word.

Integrate and Interpret. The second ~~set of~~ Comprehension ~~Targets~~Target describes what students do as they Integrate and Interpret information from one or more texts. These processes can involve making connections across sentences, paragraphs, or sections within or across texts to synthesize ideas under a common theme (e.g., justice or loss) or idea (e.g., how food goes from the farm to tables in people's houses). In making these connections, readers rely 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 communicate ideas. In order to engage in these processes, readers may be required to navigate complex hyperlinks or multimodal elements, such as video or interactive graphics.

Test items that gauge readers' ability to Integrate and Interpret may ask readers to 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 ask readers to explain or predict a character's behavior by relying on multiple pieces of text information about that character's history and dispositions, or they might ask readers to describe how the setting of a story contributes to the theme. 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 them or to rely on signal phrases (e.g., "to the contrary") to understand the connections among ideas.

Analyze and Evaluate. The third Comprehension Target, Analyze and Evaluate, describes the processes associated with examining and assessing one or more texts during and after reading. Readers may analyze by closely examining the choices an author makes about content and form and how those choices affect meaning. ~~The reader~~Readers may then use those analyses to evaluate a text by judging various aspects of the text as well as its overall effectiveness. In order to engage in Analyze and Evaluate processes, readers must view texts in relation to knowledge from other sources. Sources may include their existing knowledge base (Alexander, 2012; Lee, 2011) or common tools and criteria used in literary analysis, historical reasoning, or scientific argumentation (Lee & Spratley, 2010; Greenleaf et al., 2016; van Drie & van Boxtel, 2008). Readers also draw on their knowledge about and preferences for particular rhetorical strategies, such as the use of language, organization of text, or articulation of claims and evidence.

In items associated with the Analyze and Evaluate target, readers might be asked to evaluate the coherence, credibility, or quality of one or more texts. Readers may be asked to make judgments about the effectiveness of an author's use of figurative language, the degree to which the author provides sufficient evidence to support a claim, or the trustworthiness of the source (e.g., venue and author) (Bråten, Stadtler, & Salmerón, 2018; Meola, 2004; Ostenson, 2014; Wineburg, 1991; Wineberg & McGrew, 2017). For example, readers might use information appearing in one text as the basis for evaluating the ideas or the use of language in a second text.

Use and Apply. The final ~~set of~~ Comprehension ~~Targets~~Target, Use and Apply, reflects the culmination of comprehension, in which understandings acquired during reading are used in new situations or applied in the development of novel ideas and products (Goldman, Greenleaf, & Yukhymenko-Lescroart, 2019; Pearson, Palincsar, Biancarosa, and Berman, 2020). This set of targets reflects contemporary understandings that comprehension may involve a series of processes that culminate in readers taking some kind of action in the world outside of text. As they engage in Use and Apply processes, readers must consider how to reframe ideas from their reading and experiences to create a new product for a specific purpose and audience (Marzano, 1988). As readers reflect on how to respond to items that require such processes, they take into account their purposes, norms established by 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).

Items designed to assess Use and Apply processes will ask readers to use information they acquire through reading to solve a problem or create a new text. For example, after reading a set of commentaries, readers might be asked to produce a blog-type message for a public audience that captures the most relevant information or offers an argument about an issue. Readers might also be asked to use one or more texts as a model for developing a new text or graphic representation. In a literature context, readers might be asked to rewrite an aspect of a story in accordance with a particular, specified goal.

Comprehension Targets and the NAEP Definition of Reading Comprehension. The Comprehension Targets reflect the understanding that the extent to which a reader succeeds at particular reading tasks is dependent on many factors related to the reader's experiences,

knowledge, language development, motivations, and perceptions of self. The Comprehension Targets also reflect the centrality of readers' use of reading processes, including a range of different kinds of inferential reasoning, in the meaning they construct. In developing items that target a range of knowledge and skills under conditions that replicate many aspects of authentic reading, the NAEP Reading Assessment provides a more precise and ecologically valid measure of students' reading comprehension.

Contexts and Purposes

As stated earlier in this framework, a central principle of the NAEP Definition of Reading Comprehension is that, as a human meaning-making activity, reading comprehension is a purpose-driven activity, situated within contexts that shape every aspect of readers' engagement with text and that influence how readers respond to and learn from the experience of reading. As a result of this principle, the 2026 NAEP Reading Assessment contextualizes almost every component of reading comprehension. This section describes how two expanded components of the 2026 NAEP Reading Assessment, Disciplinary Contexts and Purposes, contribute to this contextualization.

Disciplinary Contexts. Given recent advances in theory, research, and practice about reading within disciplines, NAEP has elevated the importance of disciplinary reading in literature, science, and social studies to reflect the increased importance of disciplinary reading in schools, state standards, and large-scale reading comprehension assessments. Students will read in each context, and their reading performance on test items will be reported by disciplinary contexts, along with an aggregate score for performance across all three. Reading in such contexts involves reading texts that are drawn from the range that students encounter when reading about literature, science, and social studies. It involves engaging in tasks that yield new understanding, enable problem-solving common to such contexts, and focus on historical and contemporary social issues.

Literature Contexts. Perhaps more than in any other disciplinary domain, reading is the center of literary study and enjoyment. Themes of human experience pervade works of literature—nature and humanity, struggle and survival, love and friendship, loss and betrayal, victory and defeat, mortality and meaningfulness. Reading literary texts, such as poetry, fictional and nonfiction narratives, and criticism, provides opportunities for enjoyment and for reflection and analysis around these themes, including how they shed light on their own experiences and social worlds. Literature also often provides opportunities to connect with cultures and experiences similar to or different from one's own, extending readers' understandings about the world. Literature also invites its readers to examine text as a repository of language, rhetorical moves, and structure; to connect its ideas to those in other texts, authors and those of otherauthors and literary traditions; and to situate problems in contemporary and historical contexts.

Science Contexts. Science contexts are primarily focused on observing and explaining the natural world. Although these scientific activities do not depend exclusively on reading, texts play an important role in learning about and communicating science ideas in school and non-school settings. Learning the concepts and processes of science in school involves the use of varied texts to describe, report, and articulate claims about the natural world (e.g., textbooks) and to record systematic efforts to act upon it (e.g., observation protocols, lab notes, experimental descriptions, journal articles). Outside of schools, individuals often access

scientific information (e.g., in newspapers and on internet sites) needed to understand issues and solve problems. Moreover, the application of reading to understanding and acting upon the natural world calls on an array of reading strategies, as well as understandings about how scientists determine findings and what constitutes credible evidence for those findings.

Social Studies Contexts. Social studies includes history, geography, cultural studies, civics, and government, with less common ~~forays into~~coverage of disciplines such as sociology and anthropology. These fields offer unique ways of thinking and organizing knowledge and investigating social systems and events, current and past. In schools, 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). Social studies explores how humans organize societies and governments, how societies make use of available resources, and how cultures develop and change over time. In order to understand social studies texts, readers bring both conceptual tools needed to understand patterns in the social world (e.g., trade-offs, how perspective impacts representation) and understandings about how claims are developed and supported. Reading in social studies also requires the application of a broad range of the reading processes described in the comprehension targets.

Purposes. Purposes are a key component of the 2026 NAEP Reading Assessment. Purposes reflect a commitment on the part of NAEP to ensure that readers know why they are engaging in every part of the assessment, and to reflect the fact that all reading is done in relation to specific purposes. Within the disciplinary contexts described above, the assessment will be oriented toward purposes for reading, and these purposes will be communicated to students throughout the assessment.

Broad Purposes. When students take the 2026 NAEP Reading Assessment, each set of readings and activities they encounter will be situated in one of two broad purposes for reading that reflect standards and curriculum frameworks across the United States—reading to develop understanding and reading to solve a problem.

Reading to Develop Understanding requires students to read texts carefully and respond to comprehension test items generated from the four Comprehension Targets. These items may assess students' understanding of concepts described in a science text or the development of a literary theme, for example. These purposes tend to resemble those associated with items on widely- used reading comprehension tests. Readers might read with the purpose of understanding the motives of a particular character in a literary text or read scientific texts to understand the significance of a public health threat.

Reading to Solve a Problem requires that students work across multiple texts and perspectives while solving a problem. These activities entail using information gained during text comprehension in the service of a specific action or to create a product. For example, readers might be asked to use information across four different short texts to develop an argument for or against a city ordinance requiring bicycle lanes on all city streets with a certain traffic load.

Specific Purposes. In addition to these broad purposes, more specific purposes for reading particular texts or engaging in particular tasks will also be communicated to students. For example, within a Literature Context, students may be assigned a role and given a goal, such as working with task characters (avatar collaborators) in a book group to prepare a presentation about which character in a narrative behaved heroically. Or they might be asked to read a

brochure for a new bicycle to evaluate how well the claims about the bicycle's qualities are supported with evidence.

        Contexts and Purposes and the NAEP Definition of Reading Comprehension. The NAEP Definition of Reading Comprehension describes the role of contexts and purposes in shaping texts and activities related to reading comprehension. This definition relies on research documenting that, when readers taking the assessment know what they are doing, why they are doing it, and what role they are expected to play, the assessment is more likely to serve as a valid proxy for their reading in authentic reading contexts (O'Reilly et al, 2018). Efforts to make contexts and purposes available to students stand in contrast to the practices of many widely used standardized tests of reading comprehension. In some assessments, readers are presented with individual passages and directed to read and answer questions following each passage, with little guidance about the purpose for reading and comprehending the passage. Such tests imply a purpose, namely reading to demonstrate how well one can perform on the test. But they do not explicitly connect with any activity readers might engage with outside of a testing situation. The aim of these components is to reflect the purposes, texts, activities, and resources that influence students' reading in school, home, and community settings.

Texts

        Because texts are central to the NAEP Definition of Reading Comprehension, the 2026 NAEP Reading Framework recommends sampling from the large domain of texts that fourth, eighth, and twelfth graders are likely to encounter in school and non-school settings, as is described in more detail in the ~~Design~~ chapter 3. This portfolio of texts ranges from classic to contemporary text forms that characterize reading within and across varied disciplinary contexts. Texts will be selected with multiple and diverse criteria in mind: cultural diversity, disciplinary representation, and developmental appropriateness with regard to complexity, topic, and modality.

        Disciplinary Texts. NAEP will sample texts that are used within the three broad disciplinary contexts described above: literature, science, and social studies. The features of these texts will vary by disciplinary context and include the genres, text types, and discursive, rhetorical, and syntactic structural characteristics specific to texts in those disciplines. Sampling will also consider that such text features are normative rather than absolute, developed to address disciplinary purposes ~~for their use.~~. This means that there is overlap across disciplines regarding the kinds of texts used within disciplines.

        Literature Texts. NAEP will draw on literary texts to reflect the range of classic and contemporary genres, text structures, literary language, and cultural traditions that students experience in their classrooms and communities. Literary texts may reflect long--standing cultural traditions, like myths, short stories, novels, drama, and poetry. They can also include current evolving forms, such as fan fiction, author interviews, book reviews, and graphic novels. The challenge of reading literature is also reflected in specific discourse patterns, including word choice, sentence structure, and figurative language. Language used in literature also situates narratives in time and cultural traditions and draws on archetypal characters typical of those traditions. Literature texts may also be ironic, satirical, or narrated from a certain point of view to cue non-literal interpretations (Appleman, 2017; Lee, Goldman, Levine, & Magliano, 2016; Rabinowitz, 1987).

Science Texts. Science texts sampled for NAEP will reflect the formats, language, and structural elements germane to pedagogical, public, and professional science discourse whose purpose is to convey information, findings, and varied applications of scientific ideas. Science texts include technical information, such as raw data, bench notes, journals, personal communications, handbooks, refereed journal articles, and review articles (Goldman & Bisanz, 2002), as well as more general texts, including press releases, news briefs, websites, and blogs. Such texts draw on varied text structures, such as cause and effect, correlation, problem and solution, sequence, comparison, exemplification, descriptive classification, extended definition, and analogy. Science texts also include many kinds of visuals, including tables, graphs, equations, diagrams, models, and flowcharts, as well as description, exposition, and narrative text (Cromley et al., 2010; Lemke, 1998; van den Broek, 2010). Several challenging language constructions are also common to these texts, including nominalized verbs (e.g., *digest* becomes *digestion*), passive voice (e.g., a liter of hydrochloric acid is added to the solution), and technical and specialized words (e.g., transpiration or metamorphic) (Fang & Schleppegrell, 2010; O'Hallaron, Palincsar & Schleppegrell, 2015).

Social Studies Texts. NAEP will also sample from the varied forms of texts common to the social studies. Selection ~~should~~will represent a wide array of text types, forms of representation, sources of information, and perspectives. These texts document human activity across cultures, societies, and time periods. They include newspaper articles, diaries, letters, speeches, records of sale, advertisements, official government documents, photographs, cartoons, maps, artwork, music, and video and audio recordings. They also include interpretive books and articles about events, time periods, or people, and classroom textbooks. Social studies texts may organize ideas chronologically or thematically to represent time periods, social structures, continuity and change, cause and consequence, and varied social or historical perspectives to consider how the past influences the present (Charap, 2015; Seixas, 2010; Seixas, et al., 2015; Schreiner, 2014). Varied text structures use linguistic frames to mark arguments, persuasion, chronology, cause and effect, perspective, or comparison and contrast. Texts from long ago may even require readers to consider language and the policy contexts within which ~~they~~the texts were generated.

Digital Platform. Like the 2019 NAEP Reading Assessment, the 2026 Assessment will be entirely based in a digital platform. The widespread presence of computers and smart devices in modern society has changed ideas about what counts as text. Students in school are frequently required to read literary, science, and social studies texts that reflect the digital environment, an environment that is different from the world of print on paper. ~~On line~~Online newspapers and magazines are replete with graphs that allow readers to simulate different scenarios and see possible outcomes when a causal factor is altered. Digital science texts now in use in schools include simulations that dynamically illustrate what happens to one human body system when variables in the other body systems change.

Digital texts may be static, with no movement of the text on-~~-~~screen (Barron, 2015) and require readers to make sense of ideas using print and images (e.g., photographs, diagrams, tables) very much like those in a print-on-paper world. Dynamic texts require readers to follow movement across modes (e.g., between print and video or static image) or across nonlinear locations (e.g., clicking a hypertext link that moves you to another section) to construct meaning (Beach & Castek, 2016; Giroux & Moje, 2017; Kinzer & Leander, 2003; Kress, 2013; Manderino, 2012). Reading within and across multiple texts that contain both static and dynamic

textual elements makes reading more complex, especially when texts contain conflicting ideas and varying stylistic features that further contribute to complexity. Readers must work actively within and across these text arrangements to construct meaning and create a situation model for a particular reading purpose.

Like the 2019 NAEP Reading Assessment, many state assessments have recently migrated to online digital testing platforms. Widespread use of digital texts was acknowledged by the Common Core State Standards (CCSS) in English Language Arts (NGA-CCSSO, 2010) and by multiple state consortia assessments (including SBACSmarter Balanced and PARCC). Like reading in many of today's classrooms, these assessments include print texts paired with audio clips, podcasts, infographics, and video segments. Even states that moved away from the CCSS and consortium assessments have retained standards and assessments that acknowledge widespread use of digital texts in homes, schools, and communities. Digital platforms offer a range of affordances, including increased attention to principles of Universal Design of Assessment to increase ecological validity and precision in measuring reading comprehension (Coiro, 2020; Fitzgerald, Higgs, & Palincsar, 2020).

Text Complexity. NAEP has long taken a multifaceted approach to assessing the complexity and accessibility of texts to determine which features of text to emphasize in selecting texts. The 2026 NAEP Reading Framework continues this approach, evaluating quantitative and qualitative features of texts, along with reader-text considerations.

Quantitative text complexity measures consider long-standing indicators of complexity, such as the type and number of features that make a text more difficult to read, including such features as familiarity of vocabulary, sentence length and complexity (e.g., Stenner, 1996; Kincaid et al, 1975), and more recent developments, such as the degree of cohesion of ideas across parts of the text, and even the degree to which a given story, for example, exemplifies the classic characteristics of a story (e.g., Graesser, et al., 2014; Sheehan, et al., 2014)

Qualitative tools include careful examination of additional discourse features and conceptual load. Examples might include evaluating the transparency of the relationships between paragraphs or sections (problem-solution, cause-effect), or assessing the quality of a definition and examples provided in a text to help students understand an unfamiliar concept. In reader-text considerations (NGA-CCSSO, 2010), NAEP considers the representativeness of texts for various subgroups by addressing the questions, "For whom, in what specific contexts, and with what levels of support are specific texts harder or easier to comprehend?" (Pearson & Hiebert, 2014). With added use of interconnected digital texts, the 2026 NAEP Reading Assessment will also capture navigational complexity (such as the number of links traversed to answer a question) to evaluate the number and nature of moves readers must make within and across digital texts (Coiro, 2020).

Text and the NAEP Definition of Reading Comprehension. Texts are used in the NAEP assessment in ways that tie to all other aspects of the NAEP Definition of Reading Comprehension. The assessment's texts reflect disciplinary contexts, as well as the multiple genres and modalities, used in both school and non-school settings, as well as the many kinds of digital and multimodal texts that make up the textual diets of most students. Broad sampling increases the likelihood that all readers will encounter texts that connect to their experiences and identities, as well as to those texts that are more distant.

Universal Design Elements

The purpose of the 2026 NAEP Reading Assessment is to measure students' reading comprehension across a diverse range of test-takers. To help accomplish this purpose, the 2026 NAEP Reading Assessment employs principles of Universal Design of Assessments (UDA). Universal Design of Assessments (UDA) calls for the purposeful design of assessments that are accessible to the greatest number of students possible in order to accurately measure the same construct—in this case, reading comprehension—across the diversity of test takers (Thompson, Johnstone, & Thurlow, 2002; Thompson, Thurlow, & Malouf, 2004). To do this, assessments draw on design features, available to all test takers, called Universal Design Elements (UDEs).

UDEs are design elements of the assessment environment intended to help all test-takers access, organize, analyze, and express ideas when engaging in complex tasks, such as reading comprehension. As such, UDEs aid students' ability to engage with the content that is being tested by reducing the noise (what measurement scholars call *construct-irrelevant variance*) introduced when students lack familiarity with other aspects of assessment. For example, students might not know what the term *synopsis* means when it appears in a test item but could construct one if they knew it was like a summary. Or they might not initially be able to answer questions about the details of an obscure article but would be able to if they knew that the topic was motorcycle design. Or they might not be able to answer a vocabulary question on page 3 of a passage not because they did not know the word, but because scroll bars are a challenge for them.

Importantly, UDEs are designed to improve measurement for students across the performance spectrum rather than for only some students (Johnstone, Altman, & Thurlow, 2006). UDEs minimize but do not eliminate needs for some students' special accommodations, much like access ramps to increase building access may not enable all individuals to enter without added support. Designers validate UDEs before widespread use to ensure that purposes are reliably accomplished, enhancing precise measurement (Johnstone, 2003; Johnstone, Altman, & Thurlow, 2006).

Use of UDEs means that difficult tasks are difficult because they offer rigorous assessment of the construct being measured and not because they introduce unnecessary complexity or other construct-irrelevant sources of variance. For instance, digital test features were employed in the 2019 NAEP, including a look-back button to link test items to points in passages where relevant information was provided to avoid unnecessary searching, scrolling, and page turning; specific directions for approaching the reading of a text; a resetting feature that provided a correct response to a previously answered item so readers could continue without carrying misconceptions from one item to the next; and task partners (e.g., avatar classmates or teachers) to complete tasks in simulation of many classroom assignments. Informed by the use of these features in the 2019 assessment, the 2026 NAEP Reading Assessment uses three expanded categories of UDEs: task-based, motivational, and knowledge-based.

Task-based UDEs. Task-based UDEs are designed to clarify requirements and guide readers in their use of available resources. They increase access and sustain readers' attention as they take the assessment. They clarify the expectations for readers and help them examine and use available resources within the assessment blocks (CAST, 2020; Dejong, 2006; Zhang & Quintana, 2012). They maximize the likelihood that readers are able to cognitively engage with complex NAEP-designed reading experiences within the compressed time frame of an assessment. They might include a sequential set of directions to 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). They might also include graphic organizers that allow readers to record and revisit their ideas, reduce time spent on searching and scrolling, and, thus, provide more time for students to read, evaluate, and engage with text content. These UDEs might also include simulated student work examples or mentor texts that offer models of approaches to tasks before students complete similar tasks independently (e.g., Sparks & Deane, 2014).

Motivational UDEs. Motivational UDEs are intentionally embedded into reading activities to encourage and support readers' interest, engagement, and persistence, especially when they encounter challenging tasks. These UDEs are informed by the substantial body of research that describes the beneficial influence of motivation on reading comprehension (Alton & Proctor, 2008; Buehl, 2017; CAST, 2020; Guthrie & Klauda, 2015). They may also maintain readers' interest by communicating explicit connections between the broader purpose for completing a task and the sub-tasks that need to be completed along the way. UDEs in the form of task characters provide written and/or oral directions or serve as experts or peers to provide information or moral support. Task characters may also serve as a simulated target audience with whom readers can communicate new understandings about what they have read and learned (e.g., Use and Apply). Motivational UDEs may also include the kind of resetting feature, described earlier, which has been part of NAEP since 2019.

Knowledge-based UDEs. Knowledge-based UDEs are designed to provide relevant information about topics, concepts, or vocabulary that students may need to make meaning from text as they read. Contemporary models of reading comprehension (Kintsch, 1998; McNamara, 2021; van den Broek & Helder, 2017) describe the significant, positive impact of readers' existing, text-relevant knowledge (especially topic knowledge) on their text comprehension. Wide variations in students' knowledge result in reading comprehension performance scores that reflect, not readers' comprehension skill, but instead their differences in background knowledge about specific topics, in addition to differences in comprehension skill. A reader who happens to have knowledge related to the text presented in the assessment will be better able to use the processes described in the comprehension targets as they read and respond to questions. For instance, in comprehending a text called Patagonia Glaciers, a reader who happens to have knowledge about glaciers is likely to be better able to successfully answer the comprehension questions than one who might be a skilled reader but has no relevant topic knowledge. Knowledge-based UDEs for the 2026 NAEP Reading assessment expand the use of brief passage introductions that offer topic previews in the form of brief text, videos, or photographs.. The 2026 assessment continues using vocabulary pop-ups to offer on-demand definitions of untested vocabulary. Such knowledge-based UDEs, will help to address this long-standing potential source of bias in assessment, resulting in more accurate measurement of text comprehension across readers (Johnston, 1981). Within the 2026 NAEP Reading Assessment, knowledge-based UDEs, as well as task-based and motivational UDEs, are not necessarily a part of every assessment block. Additionally, many current NAEP blocks do not include knowledge-based UDEs, and tasks that do not include knowledge-based UDEs will continue to be a part of the 2026 NAEP Reading Assessment.

UDEs and the NAEP Definition of Reading Comprehension. Universal Design Elements in the 2026 NAEP Reading Assessment reflect the NAEP Definition of Reading Comprehension in several ways. UDEs enable readers to engage with topics to be read about by

providing brief previews and offering instructions on how to complete assessment tasks. They also include lookback buttons and definitions of some words (only those not measured on the assessment), thus reflecting the kinds of navigational aids and tools available in typical reading situations. In addition, UDEs clarify the nature and order of tasks and expected responses. The provision of knowledge-based UDEs reflects the fact that the 2026 NAEP Reading Framework is directly addressing the decades-old concern about many reading comprehension assessments: that they assume ~~all readers possess the same~~that sampling a wide variety of texts can sufficiently account for inevitable variation in readers' text-related background knowledge. Including these UDEs helps the NAEP assessment to better reflect the conditions of everyday reading situations.

Contextual Variables

In addition to the responses to comprehension items, NAEP also uses questionnaires to gather information about schools and students' interests and experiences. NAEP reports reading achievement to reflect these data, collectively called contextual variables. These include race/ethnicity, English language proficiency, socio-economic status, and region of the country~~, and, for special NAEP initiatives, large cities and districts.~~.. There are many links between these contextual variables and the NAEP Definition of Reading Comprehension. For example, NAEP has issued special reports that summarize performance according to students' habits and attitudes (e.g., How much do students like school~~; how~~? How often do they read for pleasure, go to the library, and/or read or write on a digital device?).

NAEP collects data to gain insight into contextual variables via questionnaires that are completed by students and school personnel. The questionnaire items offer many opportunities to gather information about students and their reading. Besides their demographic characteristics and language experiences, questionnaire items can also provide information about students' perceptions of the texts they read, their reading activities in school and community settings, and the encouragement and instructional support they receive from peers, teachers, or community agency leaders. Such information provides insights into the knowledge, interest, motivation, engagement, habits, attitudes, language competence, skills, and strategies that students bring to their reading comprehension. Reporting results solely by students' demographic characteristics might contribute to a perception that all students within each demographic group are the same. For example, reporting results by students' race/ethnicity might lead the public to infer that the achievement differences between racial groups are attributable only to students themselves rather than to the opportunities to learn which have been presented to them. These ideas are described more fully in Chapter 4. ~~Therefore, additional information is important for contextualizing and better understanding the circumstances in which low-performing readers learn.~~

By providing more nuanced reports that display variability within groups, and by measuring disparities in resources and opportunities to learn, the 2026 NAEP Reading Assessment seeks to make variability within groups and explanatory variables associated with reading performance more visible. Instead of portraying student groups as unitary and homogeneous, this approach will yield a more nuanced and complete measure to better understand reading disparities as the result of a ~~complexity~~complex of factors. (For more information about ~~reporting~~how contextual variables are reported, see Chapter 4.)

The digital format, which has been implemented starting in 2017, also allows NAEP to capture students' time on tasks and navigational moves as they complete the assessment. The

process data now available because of the data-gathering assets of the digital platform can provide information about student journeys through the texts, directions, UDEs, and items students traverse during the assessment. From these data, NAEP can construct indicators about how students direct their attention (including moment-by-moment shifts in focus),) and how long (or how little) they linger on different segments of the texts, the items, the UDEs, or the directions. These indicators can be used to help interpret performance differencedifferences in a richer context (Guthrie & Humenick, 2004; Guthrie & Klauda, 2015).

Contextual Variables and the NAEP Definition of Reading Comprehension. There are many links between the NAEP Definition of Reading Comprehension and the contextual variables. In general, the questionnaire items allow NAEP to better understand the relationship between performance and different student variables: (a) demographic data (race/ethnicity, SESsocioeconomic status, or community type), (b) perceptions about themselves as readers, or (c) their experiences in school and community contexts. The process data allow NAEP to connect performance to cognitive activities such as attention. Using this information to contextualize results allows for more accurate interpretations of student performancesperformance.

Summarizing the Relationship Between the Definition and Assessment Components

This chapter has described the NAEP Definition of Reading Comprehension and the NAEP Reading Assessment, and the relationship between them. Exhibit 2.1 summarizes these relationships, demonstrating how current understanding of reading comprehension, as embodied in the Definition of Reading Comprehension that opens this chapter, is represented in NAEP through the components of the assessment.

Chapter 3 takes the next step by describing the structure of the assessment and illustrating the use of key design principles and practices that will allow NAEP test developers to create an assessment that includes the components described here.

Exhibit 2.1. Relationships Between the NAEP Definition of Reading Comprehension Definition and the NAEP Reading Assessment

Assessment Components	Features of the NAEP Definition of Reading Comprehension			
	Contexts	Readers	Texts	Activities
Comprehension Items	Reflect a view of the outcomes of reading as influenced by factors within and outside of the assessment.	Address an array of skills and strategies related to comprehension, including literal, inferential, analytical, and critical responses along with items that ask students to apply ideas in the texts.	Query different types of comprehension within and across texts and different aspects of the texts, including local and global features and meanings.	Attend to disciplinary contexts, purposes, and text challenges to determine how items will reflect the four comprehension targets.
Contexts and Purposes	Invoke rich contexts (discipline-related and otherwise) as a way of situating reading in settings that involve reading comprehension.	Communicate purposes for reading, introduce social elements, such as a digital "guide" or peers, and enhance engagement by focusing on contemporary issues.	Include varied texts that align with disciplinary contexts and purposes.	Establish authentic contexts, structures, and purposes for reading and formulate tasks that are aligned with those purposes.
~~Disciplinary Contexts Purposes~~				

28

	Features of the NAEP Definition of Reading Comprehension			
Assessment Components	Contexts	Readers	Texts	Activities
Texts *y*	Include a variety of texts that represent a range of cultural traditions, disciplinary contexts, and reading purposes.	Select texts that are broadly representative of varied cultural traditions, backgrounds, experiences, and identities.	Include texts from a wide range of genres, modalities, formats, and disciplinary traditions.	Include varied texts that align with the disciplinary contexts, broad purposes, and genres appropriate for the block.
Disciplinary Texts *Digital Texts* *Text Complexity*				
Universal Design Elements	Reflect the kinds of resources that are commonly available during reading in school, workplace, and community contexts.	Provide previews of the topics, information about unknown words that are not the focus of the assessment items, and instructions on how to complete assessment tasks, allowing readers to engage in more challenging reading tasks.	Increase broad access to texts, such as providing definitions of key words not measured on the assessment and offering lookback buttons.	Provide information that clarifies the nature and order of tasks and expected responses.
Contextual Variables	Gather information	Gather information about	Gather information	Gather information

Assessment Components	Features of the NAEP Definition of Reading Comprehension			
	Contexts	Readers	Texts	Activities
Questionnaire Items	about the contexts of readers' lives and experiences in and out of school.	demographics, motivation, and in- and out-of-school reading practices.	about the amount and kinds of texts that readers encounter in and out of school settings.	about reading activities that readers commonly engage in at school and outside of school.
Process variables	Compare pathways when reading in different disciplinary contexts and for different purposes.	Track each participant's navigation through the assessment— reading texts and responding to items.	Compare pathways through the assessment when employing different sorts of texts.	Compare pathways for different sorts of items, both format and Comprehension Targets.

This chapter describes the assessment design components that contribute to best educational measurement practices, as outlined by the National Research Council (2001; AERA/APA/NCME, 2014) and used in previous NAEP Reading assessments (National Assessment Governing Board, 2019). These practices include incrementally augmenting current assessment design with features that are carefully tested and refined over time: a hallmark of NAEP development practices since the inception of the assessment.

The chapter is divided into three sections. The first section provides an overview of considerations related to developing block components of the 2026 NAEP Reading Assessment. This involves situating readers within a disciplinary context, a broad purpose, and a specific purpose and role for each block. The second section discusses the task components and how they can be used to expand the ways in which readers are asked to demonstrate their ability to engage in the comprehension processes outlined in Chapter 2. Task components include texts and comprehension items. The third section details considerations for ~~using~~leveraging digital assessment features, including item response formats, Universal Design Elements (UDEs), and process data, ~~and item formats~~ in line with principles of validity, fairness, and inclusivity (Thompson, Johnstone, & Thurlow, 2002). Overall, the design considerations outlined in this chapter are intended to enable the 2026 NAEP Reading Assessment to allow the greatest number of students to participate in ways that result in more valid inferences about their comprehension performance as situated in purposeful, disciplinary contexts.

Situating Readers Within Assessment Blocks

A block is the largest organizational unit for the 2026 NAEP Reading Assessment. In a typical NAEP reading session, test-takers engage in two grade appropriate blocks. The design of every block involves situating readers within a *disciplinary context*, a *broad purpose for reading*, and a *specific purpose* and *role* for the reader working through the block. See Exhibit 2 in Appendix C, which illustrates a range of design features that should be considered when designing assessment components. These features vary along a continuum within a block, from less to more dynamic and cumulative.

Designating Disciplinary Context

All blocks will sample from a range of grade-appropriate texts within one of three disciplinary contexts, including literature, science, or social studies contexts. The primary context for each block will be identified according to one of these contexts so that NAEP can report reading performance scales for each of these disciplinary contexts, along with an aggregate scale for performance across all three contexts. In some cases, a block may contain texts associated with more than one disciplinary context. In these cases, the block is designed as both a primary reading context that shapes the overall reading purpose and a secondary context identified by one or more interdisciplinary or cross-disciplinary topics or genres. The distribution of disciplinary contexts by grade level varies according to the approximate amount of time that students in the U.S. are engaged in the respective contexts at grade levels 4, 8 and 12. Exhibit 3.1 shows the design principle and provisional distribution targets for sampling disciplinary contexts at each grade level.

Exhibit 3.1. Principle and Provisional Distribution Targets for Sampling Disciplinary Contexts by Grade Level

Principle for Sampling Disciplinary Contexts: The percentage of Literature decreases across grades as the percentage of Science and Social Studies increases				
Grade Level		4	8	12
Disciplinary Context	Literature	50	40	33
	Science	25	30	33
	Social Studies	25	30	33

Designating a Broad Reading Purpose

In addition to situating readers in one of the three disciplinary contexts, each assessment block is also designated as having one of two broad purposes: Reading to Develop Understanding or Reading to Solve a Problem. Situating reading in purpose-driven tasks has demonstrated potential for promoting student readers' interest and engagement in existing NAEP reading assessments (Educational Testing Service, 2019).

Reading to Develop Understanding (RDU) blocks are designed to measure what readers do when asked to deeply read and comprehend—literally, inferentially, interpretively, and critically—in or across disciplinary contexts. Reading to Solve a Problem (RSP) blocks are designed primarily to assess what readers do when asked to demonstrate understanding across multiple texts and related perspectives while solving a problem. Reading to Solve a Problem activities entail developing understanding, or comprehending text, but in the service of using this understanding to take a specific action or create a product, such as a written explanation or a classroom presentation.

In both types of blocks, these broad purposes are intended to help readers prepare for reading in order to develop understanding or to solve a problem. The design principle and provisional distribution targets for sampling broad purposes by grade level are depicted in Exhibit 3.2.

Exhibit 3.2. Principle and Provisional Distribution Targets for Sampling Broad Reading Purposes by Grade Level

Principle for Sampling Broad Purposes. The percentage of Reading to Develop Understanding (RDU) blocks decreases across grades as the percentage of Reading to Solve a Problem (RSP) blocks increases				
Grade Level		4	8	12
Broad Reading Purpose	RDU	60	50	40
	RSP	40	50	60

Identifying Specific Purposes and a Reader Role

Both RDU and RSP blocks also have specific purposes with reader roles that shape how and why readers engage with the tasks, texts, and comprehension items in one of the three disciplinary contexts. These specific purposes differ from the broad block purposes (i.e., RDU or RSP) because the duration of their guidance is limited to the text or texts within a given task in the assessment block. Test developers for the 2026 NAEP Reading Assessment will craft these purpose-driven statements with an eye toward reflecting the real-world contexts and purposes for which readers engage with and make sense of a diverse range of texts.

Reader roles are designed to reflect how readers typically engage with texts and each other in different contexts (e.g., fourth grade classmates and a teacher in a literature circle discussion at school, or a group of friends at home reacting to news about a local event in their town). Some blocks may ask readers to take on a simpler, less immersive role that offers fewer specifications for the kinds of tasks with which readers will engage. Other blocks may assign readers to take on more immersive roles that offer more specifications for how readers should engage with the reading purpose, tasks, and expected outcomes.

Specific purposes and reader roles are explicitly shared with test-takers as part of the directions at one or more locations in the block. Exhibit 3.3 depicts an example of what readers might see when they begin the Grade 4 Reading to Develop Understanding sample block in a literature context. (see Appendix C). In this block, readers are invited to participate in a book discussion group about the short story *Hana Hashimoto, Sixth Violin* by Chieri Uegaki and Qin Leng with three other fourth grade student task characters (simulated avatar classmates).

). In addition to reading directions about the discussion goal, students are told they will read parts of the story and respond to items situated in threetwo purpose-driven tasks. Because test-takers encounter additional texts and items in different parts of the block, more specific purposes may be given to situate their work on particular comprehension items in the context of each new text. Note, in this example, each additional text is an excerpt from the same story.

The goal of the 2026 NAEP Reading Framework is to immerse readers in discipline-specific blocks for which both reading purpose and reader role are transparent to better simulate

the situations in which most readers find themselves in school, workplace, and community situations.

Exhibit 3.3. **Task-specific purposes presented at the beginning of a Grade 4 Reading to Develop Understanding block using the text** *Hana Hashimoto, Sixth Violin* **(a short story) by Chieri Uegaki and Qin Leng**

Introduction

You will read a book called *Hana Hashimoto, Sixth Violin* by Chieri Uegaki. After you read the story, you will participate in a small discussion group with three classmates to understand how Hana changes over the course of the story as a result of events involving her family.

To prepare for the book discussion, you will read the story while completing three tasks.

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.
3. Use your understanding of the characters to predict what might happen after the story ends.

Your teacher for this project will be Mr. Beadle

You will work with three classmates in your discussion group: Brian, Diana, and Michael

Welcome

You will read the story, *Hana Hashimoto, Sixth Violin*, by Chieri Uegaki and Qin Leng to prepare for a book discussion.

First, you will learn about **important events** in the story and **characters' thoughts, feelings, and actions.**

Then, you will **write about what the main character, Hana, is like as a person so that you are ready to discuss the book** with three peers.

Your teacher for this project will be Mr. Obas:

You will work with three classmates in your discussion group:

Gia   Gabe   Luisa

NEXT

Developing Assessment Tasks: Texts and Items

After readers are situated in the assessment block, they encounter two or more tasks, each with its own specific purpose. A task is a subunit within each block on the 2026 NAEP Reading Assessment. Each NAEP reading block has 2-3 tasks, one or more texts, and related comprehension items. Developers take into consideration time, total passage length, and grade appropriateness when determining the number of texts in each assessment block. Extended pieces of literature or a full argumentative essay might result in only one text with one or two tasks. Shorter texts such as a haiku poem, photograph, search engine result, or Twitter post might result in more than one text for a particular task.

For example, Exhibit 3.4 from an ePIRLS Grade 4 assessment block illustrates how several texts are embedded into one screen to authentically represent the array of texts young readers encounter when reading on the Internet; these texts include a webpage with two tabs and a navigational menu, an embedded hyperlink (which is the source of the answer as displayed in the blue pop-up box when the link is selected), a photo of a rocket, a photo of Mars' surface, a dynamic image of two planets spinning around the sun, and an advertisement with a hyperlink button that leads readers away from the relevant information. The item is intended to assess fourth graders' understanding of how to use embedded hyperlinks to locate and recall important information about the passage.

Exhibit 3.4. Example of multiple texts readers encounter as part of one task on the ePIRLS (2016) Grade 4 reading assessment

All grade-appropriate blocks will sample from a variety of task-specific purposes and a range of texts, including reading materials that students might use in their everyday lives, in and out of school (see, for example, Creer, 2018; Dobler & Azwel, 2007). The texts can represent one or more genres, modalities, or disciplines. See Exhibit 1 in Appendix A for additional

considerations for sampling text formats and modes. See Exhibit 2 in Appendix A for examples of different kinds of text formats and modes.

Selecting Texts

Text Selection Criteria. Passages in the 2026 NAEP Reading Assessment are selected using rigorous criteria that include:

- *Authenticity*. Do texts represent the types of texts that students encounter in their reading in and out of school?

- *Diversity*. Do texts reflect an appropriate range of perspectives, geographical regions, gender, and social and cultural traditions characteristic of the diverse U.S. population, and are they written by diverse authors?

- *Engagement*. Will texts encourage and maintain student interest?

- *Developmental appropriateness*. Do the texts reflect grade level expectations of the students assessed at grades 4, 8, and 12?

- *Disciplinary appropriateness*. Do the texts represent the range of genres/text types and text features in the disciplinary contexts of Literature, Science, or Social Studies?

- *Quality and ~~coherence~~cohesion*. Are the texts well-written and ~~considerate,~~ organized in ways that promote comprehension and learning? Do non-fiction texts, and especially those in a modality other than print, include brief and purposeful topic introductions where appropriate?

- *Complexity*. Are the language features (vocabulary, syntax, discourse and rhetorical structures) representative of the specific grade and disciplinary context?

Several of these text selection criteria are elaborated below with a number of principles and design considerations.

Authenticity. Most texts included in NAEP Reading will be presented in their entirety, as students would typically encounter them. However, some texts may be excerpted 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. Texts will be selected to evoke the range of reading comprehension processes, or targets. Only in exceptional cases, NCES and its contractors may consider commissioning authors to write a text that satisfies the needs of a particular assessment block. For example, it might become highly challenging to find a text of a particular length that is suitable for a specific grade level for a RSP purpose. In the exceptional cases in which commissioned writing may be required, it should follow the text selection criteria applied to authentic texts. In very rare cases, then, commissioned texts may be used as part of a set of texts. Thus, such commissioned texts will not serve as the main, or anchor, text for a text set, nor will students be asked items focused on evaluating the credibility or accuracy of such texts. See Exhibit 3 of Appendix A for more detail.

Developmental Appropriateness of Texts. Texts included in the assessment will be of different lengths. In grade 4, passage lengths will range from 200-800 words, in grade 8 from 400-1000 words and in grade 12 from 500-1500 words See Exhibit ~~1~~4 in Appendix A. Differing passage lengths are employed for several reasons, including the total time readers have to complete the block. To gain valid information about students' reading comprehension, stimulus

material should be as similar as possible to what students use in their in-school and out-of-school reading. Unlike many common reading tests that use short passages, the 2026 NAEP Reading Assessment will include complete texts of greater length. Such texts require students to use a broader and more complex array of reading strategies, reflecting student reading in authentic in- and out-of-school situations (Goldman, 2018; Paris, Wasik, and Turner 1991).

Reflecting classroom practice, students in earlier grades generally read shorter texts while older students read longer texts. 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 compare, synthesize, and critique texts in terms of their content, themes, and stylistic features. In these cases, the total number of words will reflect the recommended passage length range for each grade.

Because videos may be used in NAEP assessments built from the 2026 NAEP Reading Framework, some attention should be given to video length. The length of a video segment will vary in relation to its purpose and to overall block time. Video length may also increase across grade levels. However, because students have greater engagement and perceived retention rates for shorter as compared to longer videos (Slemmons et al., 2018), video length should generally be kept relatively short, especially compared to the length of other written texts within the task.

Disciplinary Appropriateness of Texts. Selected texts must be representative of the discipline in both content and structure, reflecting the range of genres and discourse features detailed in Chapter 2. Because reporting prompted by the 2026 NAEP Reading Framework will feature scales for the three disciplinary contexts, it is also important to specify both the variability of student reading within contexts and the commonalities across each context. Based on the account provided in Chapter 2 of the range of text types, text structures, and text features, Exhibit 25 in Appendix A shows important ~~textual~~text elements that characterize texts in each of the disciplinary contexts, while acknowledging that many text features are common across disciplines. A responsibility of test developers, as they build the portfolio of test blocks and tasks at each grade level, is to try to incorporate the entire array of text types and features in the blocks for each grade level. See *Assessment and Item Specifications for the 2026 NAEP Reading Framework* for more details.

Standards for Cohesion and Complexity of Texts. Efforts should also be made to promote the strategic balance and selection of texts across blocks. This process should be informed by general standards of quality, ~~coherence~~cohesion, complexity and "considerateness" (including both qualitative and quantitative measures; e.g., conventional readability criteria, reader-text connections, language structures and vocabulary considerations; Anderson & Armbruster, 1984) and reflect contemporary standards applied to digital texts and other contemporary media forms. Because readers use specific knowledge to identify important information in different types of texts, developers attend to variations in organization and cohesion in line with ~~common~~ text structures and text features that are found in common across disciplinary contexts (see Exhibit 36 in Appendix A). Test developers should strive to select texts with features that cue readers' attention to structure and influence the recall of information (Wixson & Peters, 1987).

The extent to which readers' background knowledge, experiences, and interests connect to a text and its topic will also be considered when evaluating a text's complexity, 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). Textual ideas in disciplinary contexts 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). Unfamiliar concepts should be defined with examples provided. Designers should aim for a flexible and diverse representation of language and structures across the blocks.

There is also wide variance in the nature and quality of graphical or multimodal displays of ideas in today's texts. Therefore, in selecting texts, it is important to create a sample that represents the grade-appropriate array of graphical and structural representations (e.g., static, dynamic, multimodal, nonlinear) found in print and digital reading materials. As well, texts often appear, and are used in sets. Thus, it is important to determine grade-appropriate numbers of texts, and the 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.

A potential difference between traditional and digital texts is the nature of text arrangement and the means with which readers navigate through and across texts (Cho, 2014). In selecting digital texts, it is important to attend to the features that allow for navigating ~~complex textual~~multilayered digital text environments (Afflerbach & Cho, 2017; e.g., search engines, dynamic hypertexts linked within and across documents) to reflect what readers do when they use the Internet. Further, digital texts represent diverse combinations of the information contained in text and the media used to present that information. For example, a digital text may include short (e.g., 30 second), embedded video and links to other sources of information. Thus, it is important to determine that the ideas, perspectives and modes presented in digital media reflect what readers encounter in their academic and everyday lives.

Engaging experts in selecting texts that reflect authentic social and cultural traditions in a range of disciplinary contexts. The text selection process is best conducted by experts with disciplinary, educational, and cultural knowledge about the nature and structure of texts that are representative of particular disciplinary contexts and cultural traditions in specific grade levels. Such experts should represent diverse cultures and languages in order to identify texts that reflect the broad range of student readers' knowledge and experiences.

Developing Comprehension Items

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 use 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, as described in Chapter 2, should be proportionally distributed across dimensions of the block (see ~~Exhibit 4 in Appendix A):~~Exhibit 7 in Appendix A).
  - across grade levels.
  - across the disciplinary contexts of literature, science, and social studies.

○ across broad purposes of blocks.

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.

● *Clarity and transparency.* Items should be accessible and transparent. They should be written in accessible, straightforward language, and accompanied by directions that clearly explain what steps readers should take during the activities (e.g., which texts to read and for what purpose) and how their responses will be evaluated.

● *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~~multilayered digital text environments (e.g., static and dynamic) and to make different kinds of inferences, from local bridging inferences to more complex inferences across texts and applications of knowledge to a new situation (e.g., Use and Apply). As such, audio and visual ~~texts will~~features may have items associated with them.

● *Varied knowledge sources*. Items should invoke a variety of knowledge sources in accordance with the comprehension targets in a given assessment block. Across items, readers should be called upon to employ certain kinds of background knowledge (e.g., knowledge of vocabulary and language structures, knowledge of text structures and features) and to draw information from different sources in the texts (including information at various types of representation [e.g. directly stated in prose, embedded in a visual representation, or implied through symbolism] and across different locations in the text). On the other hand, items should not assess knowledge sources irrelevant to the items and associated comprehension targets in a given block. For example, items should not ~~ask~~be answerable by readers ~~to draw~~only drawing upon text-independent domain knowledge, topic knowledge, knowledge of technical vocabulary or idiomatic expressions, or conceptual or domain knowledge in particular subject areas. Knowledge-based UDEs are ~~therefore~~ incorporated into given blocks to maximize students' ability to engage with the content that is being tested. Thus, knowledge-based UDEs are designed to reduce the noise associated with knowledge sources not being assessed in a given block and also provide orientations to the topical knowledge addressed in the text(s).

Planning the Distribution and Characteristics of Comprehension Items. The four comprehension targets do not represent a hierarchy of strategies or skills. The difficulty of any particular item, regardless of which comprehension target it is designed to elicit, should be shaped by the content of text(s) (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. The single most important standard that the 2026 NAEP Reading Assessment will meet is asking questions about matters of substance in the texts. Chapter 2 contains examples of what test items might ask readers to do with respect to each of the four comprehension targets.

~~Exhibit 4 in Appendix A~~Exhibit 7 in Appendix A presents guidelines for distributing items mapped to comprehension targets across grade level and blocks. These flexible

distributions allow for the possibility of varying the number of items for each target depending on block type. One broad principle is that the percentage of items designed to assess Integrate and Interpret or Analyze and Evaluate ideas increases across grades. In addition, in Reading to Solve a Problem (RSP) blocks, the percentage of items designed to assess Locate and Recall ideas decreases across grades as the percentage of Use and Apply ideas increases. Finally, 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.

Considering Navigational Complexity of Texts, Tasks, and Items. Developers 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; that is, the *type* of inference (a local, straightforward inference within a paragraph 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, test developers will follow guidelines from the *Assessment and Item Specifications for the 2026 NAEP Reading Framework* 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 the development of future NAEP reading tasks as NAEP learns more about how reader attributes interact with various task demands to influence comprehension performance.

Language Structures and Vocabulary in the Comprehension Items. Language structures and vocabulary in the 2026 NAEP Reading Framework refers to the application of the reader's understanding of individual words, grammatical structures, and discourse structures characteristic of grade-appropriate texts to text comprehension. Specifically, the 2026 NAEP Reading Assessment will include items designed to evaluate readers' application of their knowledge of useful grade-appropriate words and language structures to their understanding of a text or a set of texts (see Exhibit 58 in Appendix A). Because these items target readers' application of the meaning of highly useful language found across grade-appropriate texts to text comprehension, testing items will exclude language known to be part of students' everyday oral proficiency, rare words of limited application across grade-appropriate texts, discipline-specific concepts, and idiomatic expressions characteristic of particular cultural and idiosyncratic discourse practices.

A maximum of 15-20 percent of items in any assessment block will assess readers' application of passage-relevant Language Structures and Vocabulary to text comprehension, while concurrently measuring a specific comprehension process. Due to the intricate relation between language understanding and text comprehension, language structures and vocabulary will not be measured independently from comprehension targets. Instead, they will be doubly

coded for Comprehension Target (e.g., Locate and Recall; or Integrate & Interpret) and Language Structures and Vocabulary.

A note on open-ended responses. Whereas measuring students' understanding of passage-relevant grade-appropriate language is crucial, it is also important not to confuse language dexterity with the demonstration of text understanding in open-ended responses. Thus, consistent with the 2009-2019 NAEP Reading Assessments, the 2026 NAEP Reading Assessment will generate scoring rubrics and training for scorers that are language-conscious so that students are not erroneously penalized for language features irrelevant to the comprehension processes being assessed (for example, a student's written answer that displays accurate comprehension should not be negatively affected by uses of unconventional grammar or misspelled words).

Digital Assessment Features: The Role of Item Response Options, UDEs, and Process Data

An essential goal of the 2026 NAEP Reading Framework is establishing valid assessment tasks that can reliably measure diverse students' real-world reading comprehension. In the 2026 NAEP Reading Assessment, this goal is accomplished in two ways. First, all test components are designed to support ecological validity, which refers to the extent to which assessment elicits students' reading performance as it would be demonstrated in real-world settings. Newer, digital tools in particular allow assessments to situate cognitive acts of reading, to the extent possible, in complex but authentic home, school, and work reading contexts and to do so in ways that are ecologically valid (Mislevy, 2016). Second, by employing newer, digital tools, the 2026 NAEP Reading Assessment supports construct validity by providing more accurate interpretations of test results, thereby increasing the ~~potential~~ validity of scores across the diversity of test takers (c.f., Mislevy, 2016; Thompson et al., 2002).

To undertake these aims, the 2026 NAEP Reading Assessment is grounded in Universal Design of Assessments (UDA). As described in Chapter 2, UDA calls for the purposeful design of assessments that are accessible to the greatest number of students possible in order to accurately measure the same construct across the diversity of test takers (Thompson, Johnstone, & Thurlow, 2002; Thompson, Thurlow, & Malouf, 2004). See Exhibit 3.5 for an overview of UDA principles. The NAEP 2026 Reading Assessment employs UDA (Johnstone et al., 2006; Thompson et al., 2002) to select from a broad range of digital assessment features in order to design an assessment from which stakeholders can make more valid interpretations of assessment scores for all test-takers. Such digital assessment features include the purposeful selection of item response formats, universal design elements, and process data, as described in each of the next three sections. See Exhibit 3.6 for an overview of how these digital features, as well as other aspects of the 2026 NAEP Reading Assessment, align with principles of UDA.

Exhibit 3.5. Seven Principles of Universal Design of Assessments (UDA)

Principle Number and Name*	Description of Principle
1. Inclusive Assessment Population	This principle supports equitable participation in, and use of, assessments. Assessments should measure the performance of a wide range of students reflective of the population the assessment aims to represent. The assessment should do so in a way that ensures that students with diverse characteristics have opportunities to "demonstrate competence on the

	same content" (Johnstone et al., 2002, p. 6). This does not mean that the test will be less rigorous or that content should be altered. Rather, this is achieved through accessibility of content using diverse formats (e.g., item formats), technological tools (e.g., Universal Design Elements, or UDEs), and designs that include diverse test-takers.
2. Precisely Defined Constructs	Precisely defined constructs help to ensure that an assessment measures the construct it intends to measure rather than aspects not part of that construct, which creates construct-irrelevant variance. Without a precisely defined construct, it is hard to know whether items and other design features work towards measuring the intended construct or whether they might, in fact, be measuring something else.
3. Accessible, Non-biased Items	The purpose of this principle is to ensure that all test takers can access the content being assessed so that items measure the same construct for all students who take the assessment (i.e., items are "non-biased"). For example, if a passage contains a highly culturally-situated term that might be more familiar to some sub-populations of test takers (e.g., to boys more than to girls), this might unfairly advantage these students, resulting in inaccurate measurement across these subpopulations. Bias is measured statistically by comparing the difficulty of items across subpopulations of students.
4. Amenable to Accommodations	This principle refers to the physical design of the test (e.g., font, colors, graphics) being easily accessible for students' sensory abilities or easily modified (e.g., avoiding vertical text allows for the easier modification of written text into Braille).
5. Simple, Clear, and Intuitive Instructions and Procedures	In accordance with this principle, instructions and procedures of an assessment should be easily understandable regardless of a student's background (e.g., experience, knowledge, language use, concentration level). Instructions that use clear, simple language that is consistent across the assessment serve to maximize the ability of the assessment to measure the intended construct.
6. Maximum Readability and Comprehensibility	This principle refers to the ability of a text to be understood by all test takers so that readability does not interfere with the measurement of other content (e.g., on a math test, a student's ability to read an item stem does not make it harder for them to complete the task).
7. Maximum Legibility	This principle refers to test elements (e.g., text, tables, figures, illustrations, and response formats) being easily understood. Developers should consider elements such as contrast, type size, spacing, and typeface when developing a test that is as understandable as possible.

*These UDA principles are drawn from Thompson et al., 2002, where they are referred to as "elements" (see page 6).."

Exhibit 3.6 Alignment of the 2026 NAEP Reading Assessment ~~With~~with Principles of Universal Design of Assessments (UDA)

UDA Principle*	Alignment of Aspects of the 2026 NAEP Reading Assessment with UDA Principles
1. Inclusive Assessment Population	*Inclusive Population Assessed in NAEP Reading:* NAEP Reading aims to measure *reading comprehension* in a way that represents *all* students within the U.S. population at grades 4, 8, and 12 by not excluding any groups from sampling. *UDEs* UDEs minimize bias while supporting construct validity by activating students' knowledge, interest, and understanding of tasks across the diverse range of test-takers, helping to ensure that all students can access and understand the items. This supports the ability of the assessment to measure the same construct for all students, aligning with UDA Principles 1, 2 and 3. • Task-based UDEs facilitate students' ability to focus ~~limited~~ cognitive resources on the assessment tasks and items by providing clear instructions about what to do during the task (but not how to do it). • Motivational UDEs activate interest in the topics of texts and tasks, eliciting motivational processes that typically occur in out-of-test reading situations and thus improving validity of assessment items. • Knowledge-based UDEs preview untested topic knowledge and provide definitions for vocabulary not intended to be assessed ~~(e.g., a term not assumed to be possessed by *all* students).~~. This maximizes the extent to which the assessment can measure the same, intended construct for all, diverse test-takers by minimizing the possibility that one group is advantaged over another and facilitating better measurement for all test-takers.
2. Precisely Defined Constructs	*Definition of Reading Comprehension:* Chapter 2 of the framework defines the construct of *reading comprehension* and explains how this construct is operationalized using the comprehension targets as situated within the disciplinary contexts and broad purposes. This clearly defined construct helps to ensure that the assessment is measuring what it intends to measure (i.e., construct validity) by outlining exactly what is included and not included, helping to ensure that items can capture this construct and not elements outside of this construct. *Reader Roles Support Ecological and Construct Validity:*

	Reader roles are designed to situate the reader within a disciplinary context and broad purpose, as readers would be during out-of-test reading activities. While assessments can never perfectly measure the constructs they intend to measure as those constructs exist in reality, assessments aim to do so to the extent possible (i.e., what is referred to as ecological validity). In so doing, this also supports construct validity, in alignment with the "precisely defined constructs" called for in UDA Principle 2. Situating the reader within a disciplinary context and broad purpose also allows the reader to access the content being measured because it activates the reader's prior understandings relevant to those disciplinary contexts and purposes, allowing for more precise measurement of the construct. *Specific Purposes:* Situating readers within specific purposes (e.g., a reader is asked to read a story and participate in a book discussion) activates readers' prior understanding of what it means to read within a given task purpose and in so doing facilitates their ability to engage in the items and tasks. Specific purposes also help make clear to the reader what they are supposed to do with the texts and why. This aligns with "precisely defined constructs" because the specified purposes enable the assessment to do a better job of measuring the student's ability to engage with the construct and not, for example, their ability to figure out what they are supposed to do. *Item Formats:* Thoughtful selection of item formats to measure particular comprehension targets within the context of the texts and specific purposes supports students' access to the test construct because they are able to focus limited cognitive resources on tasks aimed to measure the construct. This supports the assessment's ability to measure the construct it intends to measure (Principle 2) by facilitating *all* students' ability to access the construct (Principle 3).
3. Accessible, Non-biased Items	*Regular NAEP Reading Research and Development Process:* Item bias is tested through NAEP's regular item review and pilot testing procedures to ensure that items are not more or less difficult for students from particular subpopulations. To test item bias, the difficulty of items across different subpopulations of students (e.g., boys and girls) is compared to ensure that items measure the same construct across groups. Biased items are revised until they no longer demonstrate bias. *Disciplinary Contexts & Purposes:* Because all students being tested are familiar with the school-based disciplinary contexts of literature, science, and social studies, and with the Reading to Develop Understanding and Reading to Solve a Problem purposes as they are situated within these contexts, sampling texts and tasks from these disciplines and using these purposes helps to minimize bias, since all students can be presumed to be familiar with the kinds of texts used within these three disciplines. *Range of Texts and Tasks Represented:*

44

	Selection of a diverse range of texts and tasks representing different student identities, interests, knowledge, and other backgrounds helps to ensure equity across diverse subpopulations of test-takers. Such broad sampling facilitates equitable test items and scales.
4. Amenable to Accommodations	*UDEs and Item Formats:* UDEs and thoughtful use of item formats limit the need for special accommodations. For example, task-based UDEs and item formats such as "drag and drop" can limit the need for accommodations such as extended time because they facilitate students' thoughtful use of time and focus on the texts and tasks being measured rather than on unrelated organizational skills.
5. Simple, Clear, and Intuitive Instructions and Procedures	*Instructions:* Instructions, in simple language, facilitate measurement of the intended construct (in this case, reading comprehension) because they allow readers to focus limited cognitive attention on the items rather than on the instructions. *Clear Comprehension Items and Tasks:* Similarly, items written using simple, clear language that is easily understandable regardless of a student's background (e.g., experience, knowledge, language use, interest) support the student's ability to engage in the items that are measuring reading comprehension ability aligned to the comprehension targets. Both of these aspects help to ensure that the items are measuring the intended construct (e.g., the student's ability to make meaning from literature) rather than aspects unrelated to the construct (e.g., the student's ability to understand written instructions or to understand the item stem).
6. Maximum Readability and Comprehensibility	*Selection of Grade-Appropriate Texts:* Texts are selected based on readability and text cohesion elements relevant to the grade levels in which they are tested. This helps to ensure that students taking the test can be presumed to be able to read and understand texts at these particular levels.
7. Maximum Legibility	*Visual Layout:* The 2026 NAEP Reading Assessment layout considers elements such as contrast, font type and size, and spacing within the digital environment to facilitate the validity of items because it supports' students' ability to focus limited cognitive resources on the items rather than on visual features. For example, layout should be easily accessible for different students' sensory abilities. Careful consideration of these elements also allows the assessment to be amenable to accommodations (Principle 4) because the layout is easily modified when accommodations do need to be made (e.g., translating the assessment into Braille).

* These UDA principles are drawn from Thompson et al., 2002, where they are referred to as "elements." UDEs are "Universal Design Elements."

Item Response Formats

Central to the development of 2026 NAEP Reading Assessment is the careful selection of the ways in which students respond to items. From 1992 through 2016, 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. Selected-response items for use on the 2026 NAEP Reading Assessment include a variety of formats. The 2026 NAEP Reading Assessment thus employs Selected Response and Constructed Response options. Additionally, NAEP will be exploring additional kinds of Dynamic Response options. Some examples of item response formats are presented in the next sections. See Appendix D for additional examples.

Selected Response Options. These kinds of responses allow the student to select one or more choices from provided options and include the following types:

- **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); see Exhibit 3.7.

- **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; see Exhibit 3.8.

- **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; in some cases, they also drag them into the target fields.

46

Exhibit 3.7. Example of Matching Response Format from PARCC Grade 8 Literature

Exhibit 3.8 Example of Grid Response Format from PISA

Chicken Forum Released Item #3

Constructed Response Options. These kinds of responses allow the student to develop their own response within a given parameter (e.g., a certain number of characters) and include:

- **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~~one 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 response type across grade level are presented in Exhibit 3.9.

Exhibit 3.9. 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%

48

Grade 12	40-50%	40-45%	10-15%

Dynamic Response Options. NAEP is currently exploring the use of dynamic response options to assess comprehension (e.g., graphic organizers and drop-down menus). NAEP 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. Many existing state assessments, as well as PARCC and ~~SBAC~~Smarter Balanced, use these kinds of item response formats. 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 2026 NAEP Reading Assessment, when it is indicated that such item types bring value to the assessment. For example, dynamic item formats introduce opportunities to assess how readers:

- Search and locate information (e.g., dynamic search engines); (see Exhibit 3.10).

- Select and identify information (e.g., multiple choice items with new media distractors);

- Reorder or rearrange information (e.g., ranking, categorizing, and sequencing items);

- Substitute or correct information (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 information (e.g., tiling, select, and order);

- Construct relationships among information (e.g., dynamic concept maps, multimodal representations); or

- Construct spoken responses (e.g., recorded spoken language in open-ended responses).

When selecting the format of any particular item, developers 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.

Exhibit 3.10 Example of a Dynamic Search Engine Item from ePIRLS 2016 for Grade 4 Students
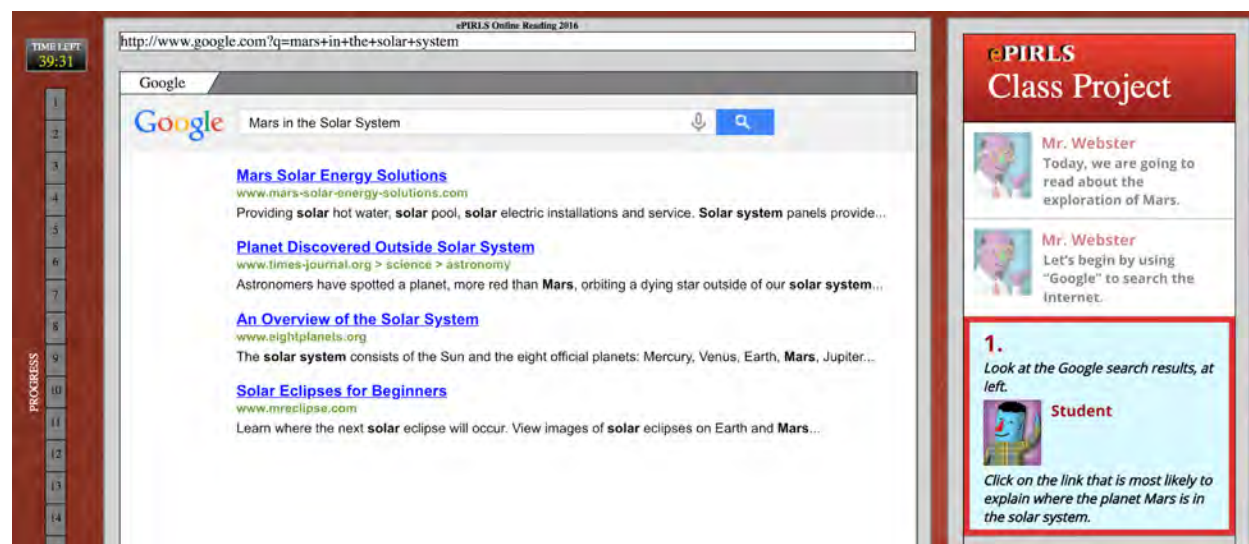
Universal Design Elements (UDEs)

Grounded in Universal Design of Assessments (Johnstone et al., 2006; Thompson et al., 2002), the NAEP 2026 Reading Assessment employs design features known as Universal Design Elements (UDEs). UDEs provide orientation, guidance, and motivation to sustain readers' journeys through the block. They are designed to mirror typical (non-testing) reading situations to improve the validity of the assessment. UDEs also offer a way for NAEP to develop fair and inclusive assessment tasks. The *fairness* of an assessment refers to a judgment about the appropriateness of decisions based on test scores (AERA, APA, & NCME, 2014). Research has shown that a student's background, language, and experience is important in how they interpret assessments (Solano-Flores & Nelson-Barber, 2001). Because these influences shape student thinking, they must be taken into account when trying to reduce bias in assessment items and support validity (Lee, 2020; Siegel, Markey, and Swann, 2005).

All readers have access to UDEs. UDEs, or the "built-in features of computer-based assessments," have been increasingly included in NAEP since the introduction of the digital platform in 2017, and are available for *all* students (NCES, 2017). Importantly, UDEs are not the same as legally mandated accommodations. While the use of UDEs might minimize the need for special accommodations, UDEs are not designed to fully address accessibility needs for the full population of students who take the 2026 NAEP Reading Assessment. Other assessment features, called *accommodations*, are legally mandated for *some* but not all students with additional testing needs (see NAEP Accommodations, last updated Oct. 2019). Examples of accommodations available on some assessments include extended time, options for responses in Braille or Sign Language, or having test-items read aloud. Universal Design of Assessments and the inclusion of UDEs are the means to enable *all* readers to validly demonstrate what they know and are able to do.

Types of UDEs. Examples of UDEs already exist in operational NAEP Reading (e.g., highlighters and look-back buttons) to reflect real-world experiences and how readers use

50

technology. Amidst the use of these digital supports by all test-takers, NAEP has effectively maintained the ability to capture trends over time (NCES, 2017). Increasingly complex reading purposes and more dynamic texts in today's society demand a broad collection of UDEs to enable test-takers to fully engage with the assessment (Mislevy, 2016). Consequently, the 2026 NAEP Reading Framework includes three broad categories: task-based UDEs, motivational UDEs, and knowledge-based UDEs. The three categories of UDEs are designed to accomplish three different, yet sometimes overlapping, functions as described next. The next section clarifies the role of each UDE and offers some hypothetical examples of how these might appear in the 2026 NAEP Reading Assessment. Additional details are provided in the item specifications. Some examples of UDEs are presented in the next sections. See Appendix E for additional examples of UDEs.

Task-based UDEs. In the 2026 NAEP Reading Assessment, task-based UDEs are used to clarify requirements and guide readers in their use of available resources in the testing space. These UDEs are designed to increase access to test content and to sustain readers' attention. A task-based UDE at the beginning of an activity (e.g., a sequential set of directions) might clearly communicate expectations for how and why readers should engage with a collection of texts. Such UDEs might also help readers plan and monitor their work across multiple texts and tasks (de Jong, 2006) by providing guidance on how to move among the texts. As readers move through the block, task-based UDEs might include graphic organizers that allow readers to record and revisit their ideas; these types of UDEs aim to reduce time spent on low-level activities (scrolling to find the location) while providing students more time for higher order activity—reading, evaluating, and engaging with text content (Sparks & Deane, 2014).

Exhibit 3.11 illustrates an example of an Integrate and Interpret item with a task-based UDE that is aligned with UDA principles calling for "assessment instructions and procedures~~...~~…to be easy to understand, regardless of a student's experience, knowledge, language skills, or current concentration level" (Thompson et al., 2002, p. 13). The item is designed to measure the student's ability to describe, in depth, a character, drawing on specific details in the text. To demonstrate this skill, the student needs to identify a character trait that is relevant, but selecting an accurate trait is insufficient to meet the construct measured. The student needs to be able to connect the selected character trait with a deeper interpretation of the character and the details of the text. In providing the word bank as a task-based UDE ~~(in this case, a word bank) is provided for,~~ all students ~~to enable them to select from an assortment of character traits and select the one most in line with~~ have an equivalent opportunity to focus more of their ~~reasoning about the main character based on her actions in the story. More~~ time and attention on the use and apply construct to be measured, rather than ~~one~~ on trying to generate a character trait word ~~choice could be an acceptable answer, but some selections are better than others, and the appropriateness of any word is linked to the reader's ability to provide a reasonable justification for their choice.~~. This type of task-based UDE is an example of one that aims to assess more challenging comprehension processes while allowing readers to access the ~~new~~ item in the relatively short period of time allotted by the assessment. ~~Such~~This clarity of expectations also maximizes the likelihood that readers ~~are able to~~will cognitively engage with complex NAEP-designed reading experiences within the short time frame allotted to each block.

The use of a word bank as a task-based UDE also aligns with principles calling for "accessible, non-biased items" and the removal of "non-construct oriented...barriers" to the assessment content (Thompson et al., p. 9). In this case, the word bank ~~is designed to~~

51

decrease decreases construct-irrelevance by providing a set of words from which test-takers can *select*, rather than *generate*, a relevant character trait. That is, the The provided words allow all readers, and especially English learners, to access the test and validly engage with the item designed to measure their ability to make inferences about character traits and not their ability to generate unfamiliar words in a timed assessment context. Similarly, this task-based UDE aims to reduce testing bias so that all students, regardless of their native language, have an opportunity to make sense of the story and demonstrate how to make inferences about characters and support their answers with evidence from the text.

Exhibit 3.11. A Grade 4 Integrate Use and Interpret Apply item illustrating a task-based UDE in the form of a word bank providing a set of character traits from which readers can select their choice and then use it as part of their constructed response.

Motivational UDEs. In the 2026 NAEP Reading Assessment, motivational UDEs are designed to facilitate students' interest in assessment content and persistence with challenging tasks (Alton & Proctor, 2008; Buehl, 2017; CAST, 2020; Guthrie & Klauda, 2015). Motivational UDEs might, for example, provide an engaging pre-reading preview or video that helps to generate a minimal amount of interest in an assessment block. See Exhibit 3.12, where a pre-reading preview and accompanying ~~30~~15 second video of children playing ~~the violin~~ string instruments serves to pique students' interest in the topic of ~~a~~the reading passage. The passage is about a girl who ~~learns~~enters a talent show contest to perform the violin she has just learned how to play~~ the violin~~. Such UDEs can increase the test's ability to measure the intended construct for all students, regardless of their prior interest and motivation.

Exhibit 3.12. A Motivational UDE in the form of a ~~30~~15 second video clip of students playing stringed instruments for the Grade 4 ~~text~~short story Hana Hashimoto, Sixth Violin by Chieri Uegaki~~.~~ and Qin Leng

As with task-based UDEs, these kinds of motivational UDEs align with UDA principles calling for "accessible, non-biased items" as well as "precisely defined constructs" (Thompson et

53

al., 2002, p. 10) by stimulating prior interest and motivation and thus removing some construct-irrelevant variance for students who might come to an assessment task with no prior interest in the topic or activity that is the focus of the assessment itemblock.

Motivational UDEs may also maintain readers' interest by communicating explicit connections between the broader purpose for completing a block and the sub-tasks that need to be completed along the way. UDEs in the form of task characters may provide written and/or oral directions, or interact directly with readers as experts, teachers, or peers to provide information (see Exhibit 3.13). Task characters may also represent members of an authentic target audience to whom readers can represent and communicate new understandings about what they have read and learned (e.g., Use and Apply). 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 engage with or react to the reading activity as a whole (Guthrie & Klauda, 2015; van den Broek, Bon-Gettler, Kendeou, & Carlson, 2011).

Exhibit 3.13. Teacher and student task characters remind the reader of the task goal <u>for the second task</u>.

Now, to prepare for the discussion, you will write about **what Hana is like as a person.**

Here are some of my notes about **Hana. Can you add some more?** Be sure to use specific details from the story about her thoughts, feelings, and actions.

Hana Hashimoto, Sixth Violin
By Chieri Uegaki & Qin Leng

Perhaps next year Hana would be able to perform one of Ojiichan's favorite pieces. But for now, Hana played a little melody she had been practicing, remembered from night lit by dancing fireflies. She imagined that the notes would drift out through the window, past the bright rabbit moon and beyond, and Ojiichan would hear them and smile.

OUR NOTES

Note 1: Hana's brothers made fun of her. She practiced anyway. The text says, "Hana practiced every day."

Note 3:

Note 2: When Hana gets on stage, she is feeling nervous. The texts says, "Hana swallowed her nerves like medicine."

Note 4:

NEXT

Knowledge-based UDEs. In the 2026 NAEP Reading Assessment, knowledge-based UDEs will provide two types of information: (a) topic previews in the form of short introductions to either the entire block or to a specific task and text, and (b) definitions or examples for unfamiliar vocabulary unless a word is explicitly tested in a comprehension test item)... Topic previews may take the form of short videos, images,written texts only, unless video, image, or a previewother kinds of specific concepts addressed in theintroductions are already part of an authentic source text. Topic previews should be offered as appropriate any time that access to information that is not part of the items being assessed could differentially advantage or disadvantage particular readers. DeterminationA determination must be made by assessment developers about whether a UDE is construct relevant. Other digital media (e.g., dynamic animations, glossary hyperlinks to related images—with or without language translation,translations—and simulations of interesting or challenging phenomena) can provide visual and multimedia cues to support readers' understanding of unfamiliar vocabulary or challenging concepts.words and phrases likely to pose construct irrelevant barriers to comprehension. Please see Exhibit 3.14 for the kinds of knowledge that will and will not be assessed. Finally, as noted in chapter 2, blocks without UDEs, including those without knowledge-based UDEs, are part of the current assessment and will continue to exist in the 2026 NAEP Reading Assessment.

Exhibit 3.14 Reading Knowledge ~~Assessed and Not~~to Be Assessed in the 2026 NAEP Reading Assessment

Knowledge ~~Not Intentionally~~Inherent to Reading Comprehension (to Be Assessed)	Knowledge Not Intentionally Assessed
Knowledge of: • Text structures (descriptive, causal, compare and contrast, problem-solution, etc.) • Vocabulary and language structures • Genres and rhetorical structures • Authors' craft	• Text-independent domain knowledge • Topic knowledge • Knowledge of technical vocabulary or idiomatic expressions Conceptual or domain knowledge in particular subject areas

What is Measured on the Assessment Through Comprehension Targets
~~Knowledge of:~~ • ~~text structures (descriptive, causal, compare and contrast, problem-solution, etc.)~~ • ~~vocabulary and language structures~~ • ~~genres and rhetorical structures~~ • ~~authors' craft~~ ~~That enables students to demonstrate their ability to:~~ ~~use~~Students' Ability to: • Recall specific text information • Use text features to derive meaning • ~~discern authors' rhetorical strategies and purposes~~ • ~~draw~~Draw inferences based on information in text • ~~synthesize information across text or multiple texts~~ • ~~analyze~~Integrate information within and across texts • ~~critically evaluate~~Analyze information presented in text • Analyze authors' rhetorical strategies and purposes • Evaluate sources of information in text • ~~use~~Use and apply ~~knowledge~~information from texts

Importantly, knowledge-based UDEs never provide answers to comprehension test items. Instead, they preview untested topic information, activate readers' knowledge, and pique interest in ways that permit readers to engage in the types of literal, interpretive, evaluative, and

application processes (i.e., the four comprehension targets described in Chapter 2) required to demonstrate their comprehension of challenging text (Alexander & Jetton, 2000; Buehl, 2017).

Exhibit 3.15 offers one example of a multiple choice Integrate and Interpret item with a Knowledge-Based UDE that aligns with UDA principles calling for "accessible, non-biased items" (Thompson et al., 2002, p. 9). The knowledge-based UDE (a pop-up box defining "talent show") is used appropriately to provide students with background information that does not overlap with the content being assessed. In this case, the multiple-choice item is not intended to measure students' understanding of the phrase "talent show." Rather, the item is intended to measure students' ability to make an inference about ~~why~~how Hana's brothers ~~flee~~first respond to her decision to play the ~~house every day~~violin in the talent show, based on ~~other character's words and~~ their actions and words (Hana's brothers ~~cover their ears~~"nearly fell out of a tree" and ~~complain about the "horrible noise" from Hana's violin practicing).~~they tell her, "you'll be a disaster!"). Since the whole story is situated in the context of a talent show, the lack of topic knowledge about what a "talent show" is might unfairly disadvantage readers who are not familiar with this term. Biases such as this in tests can result in imprecise, inaccurate, and unfair assessments of students' ability to engage in the construct being measured. The NAEP Reading Assessment does not assess what students know about different topics and disciplines; that is the job of disciplinary assessments such as social studies or science. Instead, the NAEP Reading Assessment measures how well students can reason about the information provided in texts as that reasoning is reflected in the comprehension targets used to create comprehension items. Therefore, knowledge-based UDEs ~~helps to~~like this one orient readers to the topic of the text ~~in an effort to~~, without impact on constructs being measured, and reduce testing bias so that all students have an equitable opportunity to make sense of the story and demonstrate how to make ~~inference~~inferences about characters.

Because the meaning or use of the phrase "talent show" is not directly assessed in this block, this Knowledge-based UDE also aligns with UDA principles calling for "precisely defined constructs" and the removal of "non-construct oriented...barriers" to the assessment content (Thompson et al., p. 9). In this case, the pop-up box defining a talent show is designed to decrease construct-irrelevant variance. That is, the definition allows all readers (and especially those with little knowledge about the kind of show a "talent" show is) to access the text and validly engage with an item designed to measure the reader's ability to make an inference about character actions and words rather than the reader's understanding of what a talent show is.

Exhibit 3.15. A knowledge-based vocabulary UDE in the form of a pop-up box defining the term "talent show." The pop-up appears when a test-taker clicks on the highlighted term.

Selecting appropriate locations for UDEs. Developers decide on appropriate locations in which to insert UDEs into each block of the assessment. Because some NAEP Reading 2026 tasks involve 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 task characters, visual flow charts, or simple written statements may be used to guide readers through the task sequence and structure in any given block.

A major question for block developers is how to decide when to employ and when to forego the deployment of a specific UDE as the potential for added support is weighed against the potential for increased cognitive burden on the reader. Developers will also consider how to

populate the grade-appropriate assessment space with UDEs while recognizing that readers have time limits within which to accomplish expected outcomes.

Process Data

Because 2026 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. While these data are not reported for individual students, aggregations of 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 influence reading comprehension performance at one or more locations in the NAEP assessment space. Examples of process data developers use to account for reader variations include:

- Timing data (e.g., time on passages and items),

- Navigation data (e.g., navigating among passages, pages within passages, hyperlinks, using the next button to move through a block); see Exhibit 3.16,

- Data on using other affordances (e.g., the "Look Back Button," glossing), and

- Item response process data (e.g., which answers readers choose, order of selections, answer changes, response mode, use of eliminating options in multiple choice items).

Exhibit 3.16 Example of a Constructed Response Item from ePIRLS 2016 for Grade 4 that Collects Navigational Process Data. The Space Camp image and blast off button serve as a type of distractor item designed to capture process data about readers who click on irrelevant details (i.e., advertisements) on a webpage rather than attending to the comprehension item at hand.

Overall, the strategic use of UDEs and determination of process data collected in each block enables the 2026 NAEP Reading Assessment to fully engage test-takers with complex comprehension tasks while also generating information to better account for the reading performance of fourth, eighth, and twelfth grade students. As knowledge about the use of UDEs becomes more robust and precise, more of these features should be operationalized in the NAEP Reading Assessment in the years ahead.

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 4 illustrates how these new design features allow the 2026 NAEP Reading Assessment to report the reading achievement of the nation's children in new ways that enhance the interpretive capacity of NAEP results.

The purpose of Chapter 4 is to describe how the results of the NAEP Reading Assessment will be communicated to the nation from the year 2026 onward. The chapter 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 4 outlines how the 2026 NAEP Reading Assessment will collect information that can help contextualize and explain the results it reports and serve as a useful resource for informing educational policy related to teaching reading and learning to read.

Reporting Results

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-). Results of the NAEP Reading 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*) discussed below. By design, the assessment reports results of overall achievement; it is not a tool for diagnosing the needs of individuals or groups of students. Reported scores are at the aggregate level; by law, scores are not produced for individual schools or students.

In addition to reporting aggregate results for the nation, states, and TUDA school districts, the Nation's Report Card allows for examination of results by school characteristics (urban, suburban, rural; public and nonpublic),) and socio-demographic student characteristics (race/ethnicity, gender, English learner status, socioeconomic level, and disability status (, i.e., supported by an individualized educational program), and English learner status.Individualized Education Program). The NAEP Data Explorer is a publicly accessible tool that allows users to customize reports and to investigate specific aspects of student reading achievement, such as performance on different comprehension targets or by selected contextual variables. Also, reports of the results of survey questionnaires are produced each year on various topics (e.g., students' Internet access and digital technology at home, instructional emphasis on reading activities, confidence in reading knowledge and skills, teachers' satisfaction and views of school resources).

Legislative Provisions for NAEP Reporting

Under the provisions of the Every Student Succeeds Act (ESSA) legislation, 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. Their results are included in state and national reporting. Participation in NAEP does not substitute for the mandated state-level assessments in reading and mathematics at grades 3 to 8.

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). Ten large districts

participated in 2003 and 2005. The number of districts participating in TUDA has grown over time to a total of 27 beginning in 2017. With student performance results by district, participating TUDA districts can use results for evaluating their achievement trends and for comparative purposes.

Through ESSA and the NAEP TUDA program, the NAEP Reading results report student achievement for the nation, states, and select large urban districts, enabling comparisons between states, large urban districts, and various student demographic groups.

Achievement Levels

Since 1990, the National Assessment Governing Board has used student achievement levels for reporting results on NAEP assessments. Generic policy definitions for achievement at the *NAEP Basic, NAEP Proficient,* and *NAEP Advanced* levels describe in general terms what students at each grade level should know and be able to do on the assessment. Reading achievement levels specific to the NAEP Reading Framework were developed to elaborate on the generic definitions. New reading-specific achievement level descriptors replaced those aligned to the previous framework (NAGB 2009). Exhibit 4.1 presents the generic achievement level descriptors. See Appendix A for the final achievement level descriptions.

Exhibit 4.1. Generic NAEP achievement levels

Achievement Level	Policy 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.

Reporting Results of the Updated NAEP Reading Assessment

While satisfying legislative requirements and maintaining the scale score and achievement level reporting structures, the 2026 NAEP Reading Framework updates and enhances the assessment and its reporting system to accomplish the following broad goals:

● Emphasize equity, rigor, precision, and validity throughout the assessment design and the reporting system.

● Revise 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 to increase knowledge about factors that can expand opportunities to learn.

● Transform the navigational data (sometimes called process data [Ho, 2017]), referring to how students make their way through the texts and test items) into measures that help explain test performance, as well as student interest and metacognition.

- Increase the capacity of NAEP Reading databases (including enhancements for the NAEP Data Explorer) in ways that encourage educators, policymakers, and researchers to conduct more nuanced analyses of NAEP Reading performance.

To achieve broader equity goals— with particular attention to providing within an integrated system that provides more nuanced reports and useful data to key stakeholders on research-based contextual variables focused on opportunities to learn —, the NAEP reporting system will:

1. Disaggregate scores for demographic subgroups in greater detail to provide a more accurate and dynamic description of student performance.

2. Expand the number of categories for reporting the achievement of English learners to better reflect the variability of English language proficiency within this population.

3. Reconceptualize reporting and contextual variables as an integrated system to explain student performance in ways that make the data collected more useful for policy makers and educators. Provide information on research-based contextual variables (derived from demographic, questionnaire, and process data) focused on opportunities to learn.

Reporting Categories

The framework reporting system described below provides opportunities to interpret findings from NAEP Reading results by amplifying the demographic and descriptive student categories. The reporting system expands use of the data derived from the assessment to afford deeper understanding of how socioeconomic status (SES) and race/ethnicity intersect with opportunities to learn in schools and communities (e.g., the availability of libraries or access to challenging curricula). This disaggregation of SES within race/ethnicity allows for examination of diversity within groups. To support productive interpretations of results, the reporting of achievement results for the NAEP Reading Assessment will also disaggregate reporting by current and former English learner status.

NAEP Reading Assessment results have provided indispensable information on students' performance with traditional reporting variables parsing results into subgroups to portray how students perform within specific contexts—state, region, access to technology, socioeconomic level, and many more. By expanding reporting categories and adding more contextual variables, NAEP will now be able to point the way to plausible hypotheses for policy makers to consider in crafting reforms. Thus, the 2026 NAEP Reading Framework builds on the strengths of the prior NAEP reporting system by including enhancements to the reporting and explanatory capacity of NAEP through reporting by disciplinary contexts; disaggregating results within demographic categories; and expanding reporting categories for English learners.

Reporting by Disciplinary Contexts

The 2009–2019 framework had two subscales: reading for literary experience and reading for information. The 2026 NAEP Reading Framework uses three subscales to report on 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. In addition to continued reporting of outcomes as a point on a scale from 0-500 and as the percentage of students who score within different achievement level bands (*NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*), the 2026 NAEP Reading will report additionally on each of the Disciplinary

Context scales. This enhancement is informed by increased attention to reading in the content areas in state standards across the nation.

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 status. In addition, results are reported by school characteristics, such as public/private, urban/rural, and region of the country.

Because the 2026 NAEP Reading Framework seeks to capture the dynamic variability within student groups, NAEP disaggregates student group data to show, at a minimum, differences of socioeconomic status within the student subgroup of race/ethnicity. 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, 2019) and PISA (OECD, 2019), socioeconomic status (SES) does not predict achievement in reading comprehension as accurately as it does in the U.S. Consequently, it seems likely that SES alone does not offer a direct or sufficient explanation for reading performance and that additional contextual variables are crucial to better understand variability in reading. Enhanced reporting can help policy makers and stakeholders better understand reading performances in context. For example, these data may allow policy makers to consider how access to resources that support rich literacy opportunities (e.g., high quality teaching, rigorous curriculum, community-based institutional structures such as libraries) may serve as an underlying driver of achievement.

Additional parsing of the results in this way could be important because the results might suggest that what is, on the surface, presumed to be a cohesive and static category may indeed include significant differences in access to resources. Examining SES and race/ethnicity with a more nuanced lens can surface factors that are highly amenable to change, e.g., resource allocation. When the data are disaggregated by states and TUDA districts as described in the 2026 NAEP Reading Framework, they should thus be more helpful to stakeholders for addressing the needs revealed by the assessment.

Expanding Reporting Categories for English Learners

English learners (ELs) 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 state-established standards for grade-level English proficiency and so are at the beginning or intermediate phases of acquiring English. In the prior NAEP reporting system, students were designated either as *not English learners* or *English learners* at the time of the assessment. The results for students who had been classified as ELs but who were no longer classified as such were reported along with students who had never been identified as ELs; hence, there was no way to disaggregate data to observe or track the successes and increases in achievement of former ELs.

The 2026 NAEP Reading Assessment results expand reporting categories in order to present data that is more attuned to the complex composition 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 keeping with the latest research and current requirements for state-level reporting under ESEA, Section 3121(a), the reporting system for the 2026 NAEP Reading Assessment disaggregates scores by three English proficiency categories for which school systems that participate in NAEP already collect data:

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 and who have formally exited that status;

3. *Non-English learners* – Monolingual students who speak only English; bilingual students who speak English and another language and who were never previously identified as English learners; bilingual students who reached grade-level standards of English proficiency more than two years ago.

Reporting NAEP results for these three categories will allow more nuanced interpretation of data for students who are designated as current or former ELs and highlight challenges these students may face. 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 reading achievement. This expansion of EL reporting categories will shed light on any progress—or lack thereof—that might be detectable in the group of Former ELs. With states increasingly able to collect this information about English learners' histories, and the likelihood that a majority of states will have these data available by 2026, the 2026 NAEP Reading Framework expands reporting categories for English learners in order to more accurately represent the descriptive data states and districts are already using to understand the performance of these students.

Contextual Variables

Students participating in the NAEP assessments respond to survey questionnaires that gather information on variables important to understanding reading achievement nationwide. Teachers and school administrators also complete questionnaires. ~~To the extent possible, information is also gathered from non-NAEP sources such as state, district, or school records to minimize the burden on those asked to complete the questionnaires.~~ Questions are intended to be non-intrusive; free from bias; secular, neutral, and non-ideological; and do not elicit personal values or beliefs. To the extent possible and to minimize the burden on those asked to complete the questionnaires, demographic information regarding school and student characteristics is also gathered from non-NAEP sources such as state, district, or school records.

As stated in Governing Board policy, the collection of contextual data on students, teachers, and schools is necessary to fulfill the statutory requirement that NAEP include information whenever feasible that is disaggregated by race or ethnicity, socioeconomic status, gender, disability, and English learner status. Contextual information serves the additional purpose of enriching the reporting of NAEP results by examining factors related to academic achievement in the specific subjects assessed. To satisfy the goal of enriching reports on student achievement in reading, contextual variables are selected to be of topical interest, timely, and directly related to academic achievement. In addition to questionnaires, information on

contextual variables is also obtained by analyzing process data derived from computer monitoring of students' navigation within the assessment tasks completed.

The 2026 NAEP Reading Assessment uses an expanded set of research-based contextual variables (Guthrie & Klauda, 2015; Guthrie, Wigfield & Von Secker, 2000) to understand reading achievement. Contextual variables are measurable, and some are also malleable (that is, they can be influenced). These include *reader characteristics* (e.g., students' self-perceptions about engagement and motivation, knowledge, self-efficacy, agency, effort, and interest) and *environmental characteristics* (students' perceptions about facets of home, community, or school settings, including their perceptions about classrooms, sense of belonging, and support).

The current NAEP Reading Framework collects and reports data on contextual variables, factors that shape students' opportunities to learn, including time, content, instructional strategies, and instructional resources. Contextual variables are used to predict or account for variance in ~~an~~the outcome of interest,~~ such as~~ reading comprehension scores on NAEP. The 2026 NAEP Reading Framework's emphasis on the cultural assets of individuals and the power of context to shape learning and development leads naturally to the need to identify and expand research-based contextual variables for reading. By taking into account students' differential engagement with reading and their access to home and community resources such as libraries, tutoring, and out-of-school programs, the expanded contextual variable data are intended to help contextualize and explain students' differential performance on the NAEP Reading Assessment.

The 2026 NAEP Reading Framework expands the scope of contextual variable data collected in conjunction with the NAEP Reading Assessment to reflect expanded knowledge in the field regarding cultural validity in assessment (Solano-Flores, 2010). 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, 2010; Solano-Flores & Nelson-Barber, 2001, p. 555). Attention to cultural validity in assessments can guide the development of instruments to capture the proposed contextual variables by anticipating how students with different background experiences will interpret what is being asked of them. This approach to assessment acknowledges that reading as a social and cultural practice influences 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 shaped by their everyday experiences (Lee, 2007, ~~2016~~2016a). 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; 2019). ~~From the multitude of potential contextual variables, the 2026 NAEP Reading Framework expands the questionnaires by adding a manageable, selected set of research-based and malleable factors. By taking into account students' differential engagement with reading and their access to home and community resources such as libraries, tutoring, and out-of-school programs, the expanded contextual variable data may help contextualize and explain students' differential performance on the NAEP Reading Assessment, and thereby support policymakers and stakeholders in identifying potential means to shift policy and education practice to better serve our nation's students. Guided by the latest research, the 2026 NAEP Reading Assessment includes contextual variables that seek to capture both reader characteristics and environmental characteristics.~~

With the aim of supporting policymakers and stakeholders to understand student performance and craft effective policy and education practice, the 2026 NAEP Reading

Framework envisions an integrated and coherent system of reporting. Research-based contextual variables form an interrelated network intended to capture reader and environmental characteristics. Information on each variable is collected from student, teacher, and administrator questionnaires and process data. Across the different questionnaires, information is collected on school characteristics, socio-demographic student characteristics, and student interests and experiences. Taken together, the network of contextual variables is intended to 1) predict performance on the outcome measure of reading comprehension; 2) be malleable (that is, it can be influenced in school and community settings); and 3) avoid unnecessary or inappropriate intrusions into the lives of students' families. Specific questionnaire items and process data queries are selected or created to address the variables in light of each one's potential contribution to the whole.

Reader Characteristics

Research demonstrates that when students do not see an assessment as meaningful or relevant, it may not adequately capture what they know and are able to do (Valencia, Wixson, & Pearson, 2014). With respect to reader characteristics, the 2026 NAEP Reading Framework seeks to describe the role of students' perception of the interest, difficulty, and familiarity of texts, tasks, and contexts on their performances (Pintrich and Schrauben 1992; Eccles, O'Neil et al. 2005; Valencia, Wixson et al. 2014). The assessment construct (reading comprehension) calls for better understanding 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). Reader characteristic data to be collected from questionnaires and process data include the following:

Cognition and Metacognition

1. **Cognitive strategies** in reading comprehension refer to skills used to understand a text, such as drawing inferences to connect sentences together and checking to be certain that text information is fully understood (OECD, 2011).

2. **Metacognitive strategies** in reading comprehension refer to, for example, a student's use of a mental guidance system to perform such operations as deciding which sections of text are most relevant to an assigned reading goal, how to link two sections, and/or when to reread to seek more information or clarify understanding (Cho & Afflerbach, 2017).

3. **Topical knowledge** refers to students' use of their pre-existing knowledge of the reading topic to enable them to understand text information and construct new knowledge (O'Reilly &Wang, 2019).

Engagement and Motivation

1. **Volume of reading** refers to the amount of reading a student does for personal interest, pleasure or learning (Schaffner, Schiefele, Ulferts, 2013).

2. **Reading for enjoyment** refers to the goals, uses, purposes, reasons and benefits students have for reading in school and out of school (Pitzer, & Skinner, 2017).

3. **Motivations for reading** refer to students' attention, effort, sense of self-efficacy, interest, and value for reading a particular text with a unique set of tasks and questions related to it (NAEP Reading Special Study, 2019).

Environmental Characteristics

Environmental characteristics are equally important in accounting for student performance. For example, students vary in their participation in cultural communities that may value reading in varied ways and integrate reading into their lives for different purposes (Skerrett, in press). Students' histories of engagement and participation constitute resources readers accumulate across their lifetimes and bring to bear on reading tasks, including those on 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. Understanding students' differential access to community resources that support literacy development (i.e., libraries, tutoring, out-of-school programs) is important, since as these environmental contexts shift, so do the roles of reading and texts in students' lives. The degree to which schools and communities offer access to out-of-school resources influences, to some degree, students' opportunities to learn (OTL), including their own self-initiated learning, which may vary considerably. These characteristics are surveyed with regard to students' perceptions of them. Environmental characteristic data to be collected from questionnaires and process data include the following:

Perceptions of School and Community Resources

1. **School social support** refers to the extent to which students perceive that their teachers and peers believe they contribute positively to classroom reading (through listening, speaking and interacting well with others) (Vaux, Phillips, Holly, Thompson, Williams, & Steward, 1986).

2. **Belonging in school** refers to the extent to which students perceive themselves to be accepted members of the school community (Faircloth, & Hamm, 2005).

3. **Participation in out-of-school reading/literacy activities** refers to the degree to which students have access to resources (i.e., books, computers, media centers, camps, and community organizations) that utilize literacy for enjoyment, communication, learning, and pursuing a variety of activities (Bowen, Bowen & Ware, 2002).

Perceptions of Teacher, Instructional, and Classroom Supports

1. **Teacher support for reading engagement** refers to the extent to which students perceive their teacher(s) as providing materials and tasks that encourage the development of their reading competence and engagement (Afflerbach, Hurt, & Cho, 2020).

2. **Teacher support for motivation** refers to the degree to which students perceive their teacher(s) to support their interests, self-efficacy, and reading goals (Wigfield & Wentzel, 2007).

3. **Teacher support for students' background experiences** refers to the students' perceptions that their teacher recognizes and uses students' cultural, language, and social knowledge during reading instruction (Shin, Daly & Vera, 2007).

4. **Program and curricular support for reading development** refers to the extent to which teachers and administrators perceive that the school's reading program and curriculum enables them to support students' development of effective reading practices.

The NAEP 2026 Reading Framework expands collecting and reporting of contextual variables via use of refined survey item design, thereby allowing policy makers and stakeholders to gain more actionable insights regarding the variables' influences on students' efforts and their

performances. For example, students' reported sense of reading engagement and motivation could be positively related to higher levels of NAEP Reading performance (Guthrie, Wigfield & You, 2012). Students' positive perceptions of their teachers' support and classroom climate could also be associated with higher NAEP Reading performance (Pitzer & Skinner, 2017). If relations such as these emerge from NAEP, they could have meaningful implications for the need to attend to perceptions, identity, and affect to support reading comprehension and achievement. Consideration of such factors is consistent with research on the importance of social and emotional well-being to learning (~~Damasio 1995; Markus and Kitayama 1991; Weiner 1985~~Durlak et al., 2015; Elias, 2019; Guthrie~~, J. T.,~~ & Klauda, 2016; Guthrie, Wigfield~~,~~ & You, 2012; Mahoney et al., 2019; Taylor et al., 2017), the incorporation of social-emotional learning into the design of classroom and school climate (Farmer et al., 2019; 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 (Katz~~, Brynelson & Edlund,~~ et al., 2019; Shin~~, Daly & Vera,~~ et al., 2007).

These 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 compared, controlling for reading time, the gender performance gap disappears (Torppa, Eklund, Sulkunen, Niemi & Ahonen, 2018). Since both reading time and motivation are malleable factors that 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. Availability of such contextual variables disaggregated within race/ethnicity and SES also provide opportunities to understand malleable factors that can be impacted by the organization of instruction.

Data Sources

Beyond expanding the coverage of contextual variables, the 2026 NAEP Reading Framework also updates the method for collecting such information. In addition to items in the *questionnaires* that are routinely completed by students, teachers, and administrators from participating schools or drawn from available state, district, or school records, information about some variables will be obtained from the *process data* (computer-generated records of navigational data collected automatically as students engage with the assessment) (Ho, 2017; Bergner & Davier, 2018). Exhibit 4.2 provides a list of variables, along with their source in the revised contextual variable plan.

Exhibit 4.2. Contextual Variables

Variables	Source		
	Student Questionnaire	Teacher/ Administrator Questionnaires	Process Data
Reader Characteristics			
Cognition and Metacognition			
Cognitive strategies	√	√	√
Metacognitive strategies	√		√
Topical knowledge	√	√	

Engagement and Motivation			
Volume of reading	√	√	√
Reading for enjoyment	√	√	
Motivations for reading	√	√	
Environmental Characteristics			
Perceptions of School and Community Resources			
School social support	√	√	
Belonging in school	√	√	
Participation in out-of-school reading/literacy activities	√		
Perceptions of Teacher, Instructional, and Classroom Supports			
Teacher support for reading engagement	√	√	
Teacher support for motivation	√	√	
Teacher support for students' background experiences	√	√	
Program and curricular support for reading development	√	√	

Enhancing NAEP's Explanatory Reporting Capacity

This chapter provides evidence for the potential of NAEP's reporting system to both report on and offer insights into relations between reading outcomes, students' cognitive processes and perceptions about factors that contribute to reading comprehension. The importance and visibility of NAEP results are unquestioned 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. Yet, NAEP results have also been subject to misinterpretation (Linn and Dunbar 1992; Jaeger 2003; National Research Council 2017). Because results are reported in broad categories (Race by Grade or Language Status by School Setting – Urban/Rural), they can be inappropriately interpreted. In addition, in the past, achievement results have seldom been reported as a function of malleable factors, either for reader characteristics (e.g., student motivation) or environmental characteristics (e.g., opportunity to learn factors), yet it is the understanding and attention to malleable factors that are most likely to lead to improved policies and practices that can shift student outcomes. Implementing the changes summarized below can mitigate potential misinterpretations and increase the usefulness of NAEP data.

1. **Reframe the Reporting System Within the Larger Assessment Construct.** As discussed in preceding chapters, the 2026 NAEP Reading Assessment is guided by a commitment to equity, rigor, precision, and validity while grounded in scholarship about the nature of all learning and human development. The assessment reflects the field's evolving understanding of reading comprehension, cognitive processes, and the changing nature of reading demands in today's society. Importantly, it optimizes readers' opportunities to demonstrate reading comprehension that reflect the changing demands of

our increasingly complex world (Mislevy, 2016; National Research Council, 2018). Reframing and expanding the reporting system is as important as the assessment construct itself in enhancing NAEP's explanatory power and its key role in promoting equity in the nation's education.

2. **Revise Questionnaires**. To increase the capacity to examine the impact of contextual variables related to readers and their environments, NAEP seeks to revise and refresh questions to better reflect current research. A thorough review of current surveys—both the reading-specific and core questionnaires for the three categories of participants (students, teachers, and administrators)—will determine questions that need to be revised, replaced, or discarded. While continuing its history of ensuring the appropriateness and sensitivity of all NAEP questionnaire items, this review also enables development of questions that reflect improvements in survey item design and that will allow for better data (i.e., the data reflect the constructs outlined for questionnaires in Exhibit 4.2).

3. **Disaggregate Scores to Achieve More Nuanced and Explanatory Reporting.** Just as international, state, and formative/benchmark assessments have increased disaggregation of data in reporting, it is essential to add nuance to the reporting of performance for the major demographic categories (e.g., SES within race/ethnicity) to keep NAEP reporting structures current and useful.

4. **Expand Reporting Categories for English Learners**. Expanding the number of categories for reporting the achievement of ELs enables NAEP to track the progress of different subgroups, importantly for the added category of former ELs. By reporting the performance of non-ELs and former ELs separately, it will be possible to determine whether the two groups perform at similar levels on the NAEP Reading Assessment.

5. **Mine Process Data for Evidence of Cognitive and Metacognitive Processing**. Initial forays evaluating the utility of the process (logfile) data for NAEP (Bergner & 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 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.

6. **Enhance the Visibility and Utility of the NAEP Reporting Portfolio.** An effort to expand, energize, and advertise the untapped resources of the NAEP reporting portfolio would allow for more nuanced data analyses. The NAEP Data Explorer, for example, permits users to go online and generate more sophisticated analyses than typically appear in the Report Card, which, by its nature, can only provide foundational reporting. In the NAEP Data Explorer for the 2019 Reading Assessment, a user can query the database to obtain a report which, for fourth 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; regardless, NAEP data can be used to obtain more in-depth,

statistically reliable reports beyond the standard ones offered by the Nation's Report Card.

Moreover, 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 users can query with the Data Explorer tool. Increasing the funding for these initiatives would dramatically increase the portfolio of the more nuanced explanatory analysis suggested by this framework. It would be useful to 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.

Implementing these steps, including a systematic study of the NAEP reporting portfolio, could serve to create an integrated system designed to better explain student performance. Such a process would use reporting variables, contextual variables, and the all-important outcome variable of comprehension, to create and evaluate the efficacy and utility of just such a system, including consideration of its costs, benefits, and feasibility.

Conclusion

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 2026 NAEP Reading Framework nonetheless offer increased opportunities to understand the validity, efficacy, and utility of students' assets and needs as readers.

The NAEP Reading Assessment attempts to address 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 contextual variables that can help explain comprehension scores. The identification of malleable factors by the 2026 NAEP Reading Assessment reporting system also provides information that educators and policy makers can use to guide the improvement 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. 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 and help the field and the nation to avoid some common misinterpretations of data (e.g., overgeneralizing about groups).

The enhanced reporting system for NAEP will provide a wealth of new data sources for policymakers at state and district levels. Having access to reporting by states and networks of districts, such as TUDA, can inform state- and district-level initiatives about 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 updated reporting system offers opportunities for researchers who will have access to a wider

range of data for exploring foundational questions around the dynamic nature of reading comprehension.

Ultimately, the focus on equity, rigor, precision, validity, and the definition of reading comprehension informing the NAEP 2026 Reading Framework can shape future investments in expanding student access to robust opportunities for reading and literacy engagement in and beyond schools.

~~**Glossary terms placeholder**~~

Accessibility: Designed or made available so all test-takers can participate or be engaged with the texts and/or assessment.

Accommodations: Modifications to the administration of an assessment that allow students with special needs or English Learners to meaningfully participate in the assessment without conveying any test advantages.

Achievement Level Descriptors: Descriptions of student performance at official NAEP achievement levels (NAEP Basic, NAEP Proficient, and NAEP Advanced), detailing what students should know and be able to do in terms of reading comprehension on the NAEP Reading Assessment

Activity (reading): Everything that readers do when they comprehend, apply and communicate their understanding of texts.

Agency: Individuals' power or control over their performance or efforts.

Assessment blocks: Largest organizational unit of the NAEP Reading Assessment, which includes a disciplinary context, broad reading purpose, 2 or more tasks, 1 or more texts, and 9-12 comprehension items.

Authentic text: Communication or composition produced by an author for publication purposes.

Avatar: Assessment task character acting as a simulated task partner.

Background knowledge: Previously acquired information and understanding about a concept, event, procedure, process, or topic. See prior knowledge.

Cognitive model (of reading comprehension): Theoretical construct that identifies mental operations to show the relationship between knowledge and reading comprehension.

Component: The parts of the reading comprehension assessment, specifically comprehension items, disciplinary contexts, broad purposes, texts, universal design elements, and contextual variables.

Comprehension item: Question or task that test-takers answer or complete to demonstrate how well they understand and can use what they read.

Constructed response: An open-ended response (short or long) to a comprehension item; includes a scoring guide to evaluate students' answers.

Construction-integration model: Theoretical account that depicts the multiple models of meaning that readers create and employ to comprehend: surface level (accurate decoding or literal meaning); text-based (key ideas and inferences within the text); situation model (the links that readers make between their knowledge and text ideas).

Context: The physical, temporal, historical, cultural, or linguistic setting for an event, performance, statement, or idea; latter fully understood and assessed in terms of context.

Contextual variables: Factors in the home, school, community, or workplace setting that shape students' opportunities to learn, including time, content, instructional strategies, and instructional resources.

Cultural assets: The strengths students bring with them to the classroom or to the assessment, including academic and personal background knowledge, life experiences, skills and knowledge used to navigate everyday social contexts, and world views.

Cultural validity: Effectiveness with which an assessment addresses the sociocultural influences that shape student thinking and how students make sense of assessment items and respond to them.

Decoding: Applying letter sound knowledge to a letter or string of letters to translate it into a sound representation.

Design principle: Guideline for how the assessment is structured or created (e.g., guidelines for the distribution of disciplinary contexts or purposes for 4th, 8th, and 12th grades).

Developmental appropriateness: Items, tasks, or texts that are suitable for readers at certain ages, grade levels or maturity stages in terms of content, how they are written, and cognitive or academic demands.

Digital assessment feature: A characteristic of an electronic, online, or computerized evaluation.

Digital platform: Electronic location or environment on the internet or computer where a technologically enabled assessment is operated.

Digital text: Electronic print, communication (e.g., audio, visual, images) or composition on a computer.

Digitally-based assessment: Electronic, computer-based, or online evaluation of individuals' performance.

Disaggregation: Separated into parts or elements. In the 2026 Framework, considering the effects of one variable, such as income, within another, such as race/ethnicity.

Discipline/ Disciplinary Context: Specialized academic domain (e.g., Literature, science, social studies) with specific purposes, tasks, ways of thinking, vocabulary, rhetoric, and discourse conventions.

Discrete tasks: Stand-alone text passages and related questions.

Distribution: How an item is divided, spread or organized.

Domain knowledge: Information or understanding about a particular academic field (e.g., geography) or discipline or concept (e.g, rock formation).

Dynamic text: Non-static digital format. Involves movement or navigation across modes (e.g., print, images, or video) or nonlinear locations (e.g., a hypertext link).

Ecological validity: The extent to which an assessment elicits students' reading performance as demonstrated in real-world settings, such as school, home, community or workplace.

English Learner: Second-language learner of English who speaks minority language at home, but enrolled in a bilingual education or English-as-a-second-language (ESL) program at school to develop grade-level English proficiency.

English-language proficiency: An English Learner's assessed level of speaking, writing, listening, and reading in English. Includes the use of English in academic and social settings.

Equity: The state of being fair, just, and free from bias or favoritism.

Expository text (exposition): Nonfiction composition or classification of discourse. Presents information or ideas, instructs.

Figurative language: Employed by authors of literature to create images or associations that extend beyond literal meaning of words (e.g., metaphors, hyperbole, personification, and simile).

Fluency: Quick and accurate oral reading with expression or prosody that reflects the meaning of the text.

Former English Learners: Second-language learners of English exited from bilingual education or ESL programs within the last two years and participants in all-English classrooms.

Foundational reading skills: The basic competences needed for English reading comprehension, such as word recognition (decoding and vocabulary knowledge), sight word reading, and fluency.

Global inference: Reader's assumption or conclusion based on ideas or evidence drawn from prior knowledge and across the text.

Historical reasoning: Critical thinking about the past that involves evaluating the credibility of primary sources. May be assessed by the Analyze and Evaluate Comprehension Target when students read texts in the disciplinary context of social studies.

Hypertext: Interconnected documents or sources of information that readers can immediately access on the internet through diverse actions (clicking on a word, a link, etc.)

Inferential reasoning: Act or process of deriving logical conclusions from premises known or assumed to be true; the conclusions drawn from this process. In 2026 NAEP reading assessment, involved in all four Comprehension Targets.

Foreshadowing: Use of hints or clues in a narrative to suggest future action.

Knowledge-based UDE: A type of Universal Design Element (UDE) that includes topic previews/introductions and vocabulary pop-up definitions.

Linguistic knowledge: Native-speakers' unconscious understanding of the language(s) (vocabulary, syntax, etc.) spoken in their homes and communities. What is taught to students about English in school.

Malleable factors: Conditions, items or issues that can be changed or modified in students' schools or communities.

Metacognition: Awareness and analysis of one's own learning, reading, or thinking processes.

Modality: Different ways that information is presented (e.g., auditory, visual, tactile, kinesthetic).

Motivational UDE: A type of Universal Design Element (UDE) that encourages and supports readers' interest, engagement and persistence, especially when encountering challenging tasks.

Multimodal text: Meaning conveyed through still and moving images, animations, color, words, music, and sound.

Navigational complexity: The difficulty of progressing through assessment components and modalities to demonstrate comprehension based on what test takers encounter and have to do. Includes the number and types of texts to read, inferences to make, tasks to complete, items to answer, responses to provide, and modes (print, visual, images, audio, etc.).

Operationalization: To put into action or to realize.

Opportunities to learn (OTL): 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.

Prior knowledge: Previously acquired information and understanding about a concept, event, procedure, process, or topic. See background knowledge.

Process data: Information collected as students navigate the digital assessment, including the time taken to read texts and respond to questions, how often they return to the text to answer questions, and their use of optional digital tools.

Reader self-efficacy: An individual's belief in his or her capacity to read effectively to accomplish reading tasks.

Scenario-based tasks: Simulated settings in which students read passages while following steps to accomplish a particular purpose, especially to solve a problem.

Selected response: Answers in which a student selects one or more options from a given, limited set of answer choices.

Situation model: Part of the Construction-Integration model of reading comprehension (Kintsch, 1988). The level where readers make links between text ideas and their own knowledge.

Social Emotional Learning (SEL): How humans "develop healthy identities, manage emotions and achieve personal and collective goals, feel and show empathy for others, establish and maintain supportive relationships, and make responsible and caring decisions" (CASEL, https://casel.org/what-is-sel/).

Sociocultural context: The environments and experiences that shape individuals' thinking, learning, and development, including reading comprehension. Diverse communities' values, beliefs, experiences, communication patterns, and styles of teaching and learning.

Static text: Non-moving print, graphics, or images.

Student identity: A student's evolving view of self in a given social context influenced by his or her experiences, personal history, and other events.

Syntax: The organization of words or phrases into sentences in a text, composition, or speech.

Task-based UDE: A type of Universal Design Element that clarifies requirements and guides readers in their use of available resources; increases readers' access and sustains their attention as they take an assessment.

Text complexity: The conceptual, structural and linguistic features that create comprehension challenges for readers. Includes density and nuance of ideas and language structures, word frequency, passage length, syntactic complexity, and stylistic features. Typically monitored by research-based quantitative measures of readability and qualitative analyses of semantic, syntactic, and discourse elements.

Text genre: Category used to classify literary and other works by form, technique, or content.

Text structure: Organization of ideas in a composition. In narrative compositions, according to a sequential, event-driven story grammar; in expository compositions, according to rhetorical structures (e.g., description, comparison-contrast, sequence, problem-solution, or conflict-resolution).

Text-based inference: Act or process of deriving logical conclusions or assumptions based on information stated in the composition.

Topic knowledge: Understanding or information about the specific subject of a text or text segment, such as dinosaurs or river formation. Tends to be more specific than domain knowledge or world knowledge or prior/background knowledge.

Trait: A distinguishing feature or quality.

Universal Design Element (UDE): A feature of the assessment environment provided to help all test takers access, organize, analyze, and express ideas when engaged in complex tasks.

Universal Design for Assessment: Principles for creating and administering evaluations or tests so accessible, include as many types of students as possible, and result in valid inferences or scores in terms of grade-level performance.

Validity: How accurately a method measures what it is intended to measure.

Variance: A statistical measurement of the spread between numbers in a data set.

Vocabulary pop-up: A knowledge-based UDE in NAEP that a test taker can access to obtain the meaning of a word important for understanding the overall text but not assessed in the comprehension items.

World knowledge: Global information about other cultures, countries, and people. See background and prior knowledge.

Exhibit 1. Principle and Provisional Distribution Targets for Sampling Assessment Design Elements: Text Formats and Modes

For All Grade Levels

~~Exhibit 1.~~

Principle: The percentage of different text formats (static or dynamic) and modalities (print, sound, image, and multimodal) should reflect their distribution in the population of texts that students encounter in and out of school at different grade levels.
• As dynamic and multimodal texts increase in our society and schools, NAEP should aim to keep pace with those shifts.
• Current NAEP: 80% print, 20% other modalities

Exhibit 1 provides guidance to developers about sampling different kinds of texts (where texts include multimodal forms of representation). The underlying assumption in the exhibit is that there exists a continuum of forms of representation. That continuum is bounded at the one end by more static, print texts and at the other end by a complex and variable range of text types, features, and purposes. The exhibit provides advice about sampling for the present (80/20 static/dynamic and multimodal) and the future (to reflect the distributions in school and society).

Exhibit 2. Illustrative Examples of Texts and Other Media Across Single Static and Dynamic Texts and Multilayered Digital Text 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 2 in this appendix.	*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 |

MULTILAYERED DIGITAL TEXT ENVIRONMENT

• Augmented reality text
• Blog
• Database
• Digital creation/composition tool
• Dynamic simulation
• Email
• Interactive model

• Google document or Google folder
• Role play simulation
• Search engine
• Social media (e.g., Facebook, Instagram, Twitter)
• Threaded discussion
• Webpage or website

Exhibit 2 provides examples of the types of texts/media that designers should consider for the three text environments (single static, single dynamic, and multilayered digital) in NAEP blocks.

Exhibit 3. Commissioned Texts: Parameters and Constraints

Guidelines for Using Commissioned Texts
The following guidelines seek to provide clarity about the circumstances under which commissioned texts might be used and the criteria with which developers should use such commissioned texts: • Rare, never to exceed more than 5-10% of all texts included in NAEP at any grade level; 5% limit at 12th grade unless permission issues are encountered • Only used when an appropriate authentic text cannot be located to include within a text set for a block, but never as an "anchor" text for a block • Authored by writers within the discipline in which the block is situated and using specific criteria to meet strict guidance re: form and purpose • Vetted for accuracy, authenticity, and appropriateness by experts in the discipline, NCES's text selection panel, and the ADC • No items asking students to evaluate source credibility of such commissioned texts will be used • Will meet the same complexity and other criteria for text selection as all texts for NAEP Reading

Exhibit 3 summarizes the guidelines that developers will use to determine if, when, and how texts will be commissioned to meet particular needs that cannot be met by sampling already published (i.e., authentic) texts.

Exhibit 4. **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

Exhibit 2.Exhibit 4 provides ranges for the total number of words in the text(s) within a given block. This total might be distributed across 1-4 texts depending on the broad purpose (Reading to Develop Understanding or Reading to Solve a Problem) of a block.

Exhibit 5. **Typical Text Elements Across Disciplinary Contexts**

Context	Genres and Text Types	Discourse, Language Structures, and Text Elements
Literature	Fiction (Short stories, novels, plays) • Myths, legends, and fables • ~~Short stories~~ • Coming of age stories • ~~Novels~~ • ~~Dramas~~ • ~~Poetic traditions~~ • Satires • Science fiction • ~~Satires~~ • Magical realism • ~~Biographies~~ • Fantasy • ~~Memoirs~~ • Comic books • Graphic novels • Manga • Fanfiction Poetry • Haiku, sonnet, ballad, dirge, epic, etc. Related Nonfiction • Memoirs • (Auto)biographies • Literary analyses • ~~Literature reviews~~Reviews and recommendations • Author profiles ~~and biographies~~	• Plot ~~and~~ types • Character types • Narrative elements (character ~~structures~~, setting, plot, conflict, rising action, climax, resolution) • Figurative language (symbolism, imagery, simile, metaphor, personification, satire) • Point of view • ~~Dialogue~~ • Theme • Soliloquy, dialogue, and monologue • Diction ~~and~~, word choice • Repetition, exaggeration • ~~Exaggeration~~ • ~~Theme and message~~ • Flashback • Foreshadowing • Mood, tone, irony, paradox, and sarcasm • Visual and graphical elements such as illustrations and photographs • Multimodal elements such as narrative soundscapes • Description, ~~exposition~~ • Narrative and ~~narrative elements and~~expository text structures
Science	• ~~Reports~~ • Science reports • Press releases • ~~News briefs~~ • Science news and features • Science magazine articles • Reference materials and field guides • Discovery narratives, ~~biographies,~~	• Linguistic frames and signals for organizing arguments, comparisons, sequences and/or causal chains • Abstraction and nominalization (e.g., use of technical terms like transpiration to represent a sequence of events in an explanation ~~sequence)~~)

	- Biographies and first-person accounts - Blogs and other forms of public engagement in science - Science websites, such as those of universities, federal and state agencies, formal research groups, hospitals, etc. - Raw data - Bench notes and science journals - JournalProcedures - Published research articles - Personal communications	• Epistemological qualificationEmbedded definitions (science specific words explained in the text) - Science-specific definitions for polysemous words (e.g., heat, energy) - Qualification of claims: may, probably, indicates, suggests, etc. • Spatial (place, location) and temporal indicators (era, time, sequence, and tense) - Linguistic and numeric indicators of magnitude and scale - Visual and graphical elements such as charts, tables, graphs, equations, diagrams, schematics, models, photographs, digital scans and images • Multimodal elements such as simulations orsimulation, time lapse photography and animations
Social Studies	• Primary, secondary, • Historical and tertiary text traditions (mainly in history) - Primary:contemporary documents such as newspaper articles, editorials, political cartoons, broadsides, blogs, census data, diaries, letters, speeches, inventories and records of sale, advertisements, archival documents, cultural artifacts • Secondary: interpretive - Biographies and autobiographies - Historical and contemporary photographs and video - Data (tables, charts, graphs, infographics) conveying information such as demographic, employment and education levels, voter registration and turnout statistics, Gross Domestic Product and other economic measurements, etc.	- Linguistic frames and signals for organizing arguments, comparisons, and/or causal chains - Lexical expressions that mark chronology or argument - 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 expressions and ideologicalterminology - Ideological markers of language and rhetorical devices (word choices, emotional appeals, hyperbole) - Visual and graphical elements such as maps, timelines, political cartoons, photographs - Multimodal elements such as digital stories, procedural texts, public service announcements

Interpretive explanations ~~of~~**or arguments about** historical, social, and cultural phenomena and trends. **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. All texts should include information about their sources and authors. In general, NAEP applies a standard of accuracy and trustworthiness to the texts it selects, especially in matters of scientific inquiry. For certain tasks, however, it is necessary to use texts with questionable, or at least different, levels of accuracy and trustworthiness if the purpose of a block, or a task within a block, is to engage students in analysis and critique of texts. It is even more likely that NAEP will employ texts that represent different perspectives on an issue when students are asked to compare the multiple perspectives that texts/authors bring to a social or scientific issue.

Exhibit 5 provides a list of the text types and elements that test developers will consider as they sample texts within the three disciplinary contexts of literature, science, and social studies. Examples are provided for both broad organizational structures (genre and text type) and highly specific features that define the nature and flow of discourse at more specific levels of text (sections, paragraphs, sentences, and even words). While it is impossible in NAEP to represent the entire range, these elements define the portfolio of possibilities that developers will consult in selecting specific texts, making sure that a range of broad organizational structures and specific features are represented in the sample for each discipline and each grade level.

Exhibit 6. (Note: Many text types and elements are common across disciplines.)

Text Structures and Features Within and Across Single Static and DynamicTexts and Multilayered Digital Text Environments

SINGLE STATIC TEXT	SINGLE DYNAMIC TEXT
~~Textual~~Text **structures** are comparable to those in a printed format for texts designed to inform, entertain and/or persuade. ~~Textual~~Text **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~~Text **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~~Text **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~~MULTILAYERED DIGITAL TEXT ENVIRONMENT

~~Text~~**In multilayered digital text environments (Cho & Afflerbach, 2017), 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.

Exhibit 46 describes the possible relationships among important factors in shaping the distribution of texts, especially now that many of the texts within NAEP will bring digital affordances along with those of print texts. It provides an overview for developers about what they should expect in blocks built in accordance with the 2026 NAEP Reading Framework. Ideas within each cell are likely to change and expand as new kinds of texts and technologies continue to emerge.

Exhibit 7. Distribution of Cognitive Comprehension Targets Across Grade Level and ~~Blocks~~Broad Purposes

~~Grade Level and Comprehension Targets~~Rules of Thumb	
· **The distribution of items for the comprehension targets should be monitored at the pool level (across the two broad purposes—Reading to Develop Understanding and Reading to Solve a Problem) at each grade level** · **All Comprehension Targets are employed at each grade level.** · **All Comprehension Targets require students to consult the text in order to select or construct responses. What changes across targets (from Locate and Recall, to Interpret and Integrate, to Analyze and Evaluate, to Use and Apply) is the sophistication of the text-based reasoning and the inferences involved.** · **Moving across grades, the proportion of higher-level Comprehension Targets increases** · **RDU blocks, by definition, do not require the application of ideas to a new task. Hence the bulk of Use and Apply items will be in RSP blocks; however, NAEP should be open to the possibility that an RDU block might merit an item based on the Use and Apply Comprehension Target.**	
Grade	**Combined Block Pool: both Reading to Develop Understanding and Reading to Solve a Problem Blocks (% Target Ranges per Block)**
Grade 4	
Locate and Recall	15 - 30 ~~- 40~~%
Integrate and Interpret	10 - 30 ~~- 40~~%

90

Analyze and Evaluate	10 - 20%
Use and Apply	~~10~~0 - 20%
Grade	**Combined Block Pool: both Reading to Develop Understanding and Reading to Solve a Problem Blocks (% Target Ranges per Block)**
Grade 8	
Locate and Recall	10 - 20%
Integrate and Interpret	20 - 30%
Analyze and Evaluate	20 - 30%
Use and Apply	~~10~~0 - 20%
Grade 12	
Locate and Recall	10 - 20%
Integrate and Interpret	~~20~~25 - 30%
Analyze and Evaluate	~~30 - 40~~25- 35%

Use and Apply	~~20~~0 - 30%

Exhibit 7 provides both the principles and ranges anticipated for the distribution of items for each comprehension target within blocks developed for each broad purpose (RDU and RSP) at grades 4, 8, and 12. Because item development is so greatly influenced by the affordances of the texts selected, the ranges for item types will vary from block to block, even within each broad purpose. Hence, as with previous frameworks, NAEP monitors the range of comprehension targets by looking at the total distribution across all of the blocks within a grade level for each disciplinary context.

Exhibit 8. Inclusion and Exclusion Criteria for Connected Language and Vocabulary

Language Structures & Vocabulary Included / Excluded ~~From~~from Testing	Criteria
Included	• 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 linked to central ideas in the passage • Words and language structures found in grade-appropriate 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*)
Excluded	• Rare words of limited application across grade-appropriate texts and discipline-specific concepts (e.g., fiduciary, photosynthesis) • Idiomatic expressions (e.g., spill the beans, up in the air) • Words and language structures that are already likely to be part of students' oral proficiency at a specific grade level.

*Note: A total of 30 percent of items in any assessment block will assess passage-relevant Language Structures and Vocabulary knowledge while concurrently measuring a specific comprehension process.

Exhibit 6.Exhibit 8 describes the types of words and structures that developers may and may not include when developing the set of vocabulary items for a given block. Vocabulary items are doubly categorized: (a) by the language structures and features in this table; and, (b) by the comprehension targets. In terms of reporting, scores on vocabulary items are aggregated with other comprehension items to create an overall comprehension block score for each student.

Principle and Provisional Distribution Targets for Sampling Assessment Design Elements: Text Formats and Modes

For All Grade Levels
Principle: The percentage of different text formats (static or dynamic) and modalities (print, sound, image, and multimodal) should reflect their distribution in the population of texts that students encounter in and out of school at different grade levels. • As dynamic and multimodal texts increase in our society and schools, NAEP should aim to keep pace with those shifts. • Current NAEP: 80% print, 20% other modalities

Exhibit 7. Range of Design Features for Assessment Components with which Students Might Engage in A Block

Assessment Component	More constrained and conventional assessment features		More complex, dynamic, and iterative assessment features
Block	Less involved specific reading purposes that focus students' attention on a theme, question, or problem to be explored during the block (e.g., consider how a character changes throughout a story). Not, all tasks within the block necessarily work directly toward this theme.	⟷	More involved specific reading purposes 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, 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 block will help readers work towards this theme, question, or problem.
Role of readers	Reader is less constrained (assigned less of a role) by specific reading purposes that contextualize expectations for how to engage with provided texts and tasks.	⟷	Reader is more constrained by specific reading purposes and role expectations about 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 reading purpose(s) and expected outcome(s).
Tasks	*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 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.	⟷	*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 interdependent; more need for resetting. *Culminating elements:* More involved culminating task at the end of an activity that directly addresses the question or problem; major driver of the block.

94

Assessment Component	More constrained and conventional assessment features		More complex, dynamic, and iterative assessment features
Assessment Component	More constrained and conventional assessment features		More complex, dynamic, and iterative assessment features
Texts	*Number*: 1-3 related texts; excerpts rather than entire texts from some texts may be included rather than in their entirety *Dynamism*: More static texts with minimal dynamic features *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*: More texts with dynamic or multimodal text features *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.
Universal Design Elements (UDEs)	Less complex reading purposes that may involve UDEs for knowledge or motivation but lesser need for task-based UDEs.	↔	More complex and inter-related reading purposes that may involve UDEs for knowledge or motivation but greater need for task-based UDEs.

Exhibit 8. 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 XXX.	*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~~
- ~~Blog~~
- ~~Database~~
- ~~Digital creation/composition tool~~
- ~~Dynamic simulation~~
- ~~Email~~
- ~~Interactive model~~

- ~~Google document or Google folder~~
- ~~Role play simulation~~
- ~~Search engine~~
- ~~Social media (e.g., Facebook, Instagram, Twitter)~~
- ~~Threaded discussion~~
- ~~Webpage or website~~

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~~2026 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**~~

Organizational Features and Structures of the Reading Construct: Contexts, Purposes, Comprehension Targets, and Text Complexity

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~~2026 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 within an assessment block, 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.

Broad and Specific Reading Purposes

~~According to the sociocultural model, reading~~ Reading activities in an assessment block are situated within not only a disciplinary context but also a broad reading purpose. ~~This section describes~~Each assessment block is designated as having one of two *broad* purposes: Reading to Develop Understanding or Reading to Solve a Problem. Reading to Develop Understanding (RDU) blocks ask students to *read and comprehend deeply* (analyzing, inferencing, interpreting, and critiquing) in or across disciplinary contexts. By contrast, Reading to Solve a Problem (RSP) blocks ask students to demonstrate understanding across multiple texts and related perspectives in order to solve a problem. Reading to Solve a Problem activities do involve comprehending text, but in the service of a specific action or product, such as a classroom presentation.

Both RDU and RSP blocks also have *specific* purposes with reader roles that shape how and why readers engage with the ~~mapping of~~tasks, texts, and items in each block. Unlike the broad purposes, these specific purposes are applicable only to the texts in a given task in the assessment block. The purpose-driven statements will reflect the contexts and scenarios in which reading ~~purposes~~in the real world occurs. The subsections below describe how specific reading purposes map 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, maps, data, 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~~in static, dynamic, and ~~between~~multimodal 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~~should be able to 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~~should be able to 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~~should be able to 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 within and across ~~a text or~~static, dynamic, and multimodal 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~~should be able to 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 static, dynamic, and multimodal 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 static, 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~~should be able to 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~~should be able to 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~~should be able to 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~~.~~ when reading static, dynamic, and multimodal texts. 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~~should be able to analyze how word choice impacts a text's meaning and tone. Readers ~~can~~should be able

to 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~~should be able to 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 static, dynamic, and multimodal 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~~should be able to 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~~should be able to 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 static, 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~~should be able to 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~~should be able to 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~~.~~ when reading static, dynamic, and multimodal texts. 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~~should be able to analyze how word choice impacts a text's meaning and tone. Readers ~~can~~should be able to 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~~should be able to 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~~should be able to 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 static, dynamic, and multimodal 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 canshould be able to 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 canshould be able to 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 canshould be able to 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).

___

This ~~last section presents two hypothetical examples~~appendix is provided to describe design considerations, based on the principles outlined in the framework, that assessment developers might weigh as they develop blocks. Each design decision requires tradeoffs, and assessment developers must consider which tradeoffs to make and why. Such decisions are guided by the components of the assessment—the disciplinary context, broad purpose, tasks and texts, and comprehension targets. Moreover, developers must consider whether and how different design features (item response formats, UDEs, and process data) will be used so that a broad array of features are included, in purposeful ways, across the multiple blocks that are sampled.

Employing the 2026 NAEP Reading Assessment ~~blocks. The first example illustrates a~~Framework Principles: Assessment Components

The 2026 NAEP Reading Assessment Framework describes three areas of design considerations about which developers will make decisions: the block components (disciplinary context, broad reading purpose, specific reading purpose, and reader role); the task components (tasks, texts, and items); and the design features (item response formats, UDEs, and process data). See Exhibit 1 for an illustration of how these areas relate to one another.

It is important to note that developers do not necessarily make decisions about these three areas in this order; rather, some of these decisions might be iterative and mutually informative. For example, in developing a literature block for a certain grade level, the developer might first choose a text and broad reading purpose and then determine the reader's role and a specific purpose appropriate to the text. Thus, the areas are only used to illustrate the relationship of these considerations to one another and how students might experience the block.

First, students learn what disciplinary context and broad purpose they are working in, and then they learn the specific purpose and their role. Second, students are given a text or texts to read and tasks to work on as they read that text. As students engage with the texts and tasks, they complete comprehension items, which are situated within the tasks, as illustrated in Exhibit 1. Third, design features such as item formats, UDEs, and process data are used to leverage the digital assessment environment to measure, as precisely as possible, how well students perform on the blocks. The relationships among all of these features of the assessment are synergistic. The disciplinary context and broad reading purpose drive the specific reading purpose, reader role, selection of texts, and the tasks; all of which, in turn, inform the comprehension items. Items are created in relation to item response formats, as different formats are used to collect different kinds of information. Similarly, all assessment components inform the use of UDEs because UDEs are used to help ensure that all students can gain access to the tasks required of them to complete the assessment and that the assessment measures students' reading comprehension of the texts and not something else (e.g., how well they can read or follow test directions). In this manner, a well-integrated block results, with all of the parts working in tandem.

Exhibit 1 illustrates the assessment components and their relationship to one another. Each block defines a disciplinary context, broad purpose, block-specific purpose, and reader role. Each block also outlines 2-3 tasks, which are explicitly stated to the reader and which might include sub-tasks, for readers to complete as they read one or more texts. For each task, there might be one or more comprehension items. UDEs are only employed as needed to bolster construct validity and ensure better measurement of the reading comprehension construct. Similarly, process data are only collected in places where developers think it might be useful for understanding why students perform the way that they do or for informing revision or future research and development.

As developers develop a block, they make decisions about each of the components described in Exhibit 1. In the following section, we describe some of the different considerations developers might think about as they make decisions about the assessment components illustrated.

Exhibit 1. Design Components of a 2026 NAEP Reading Assessment Block
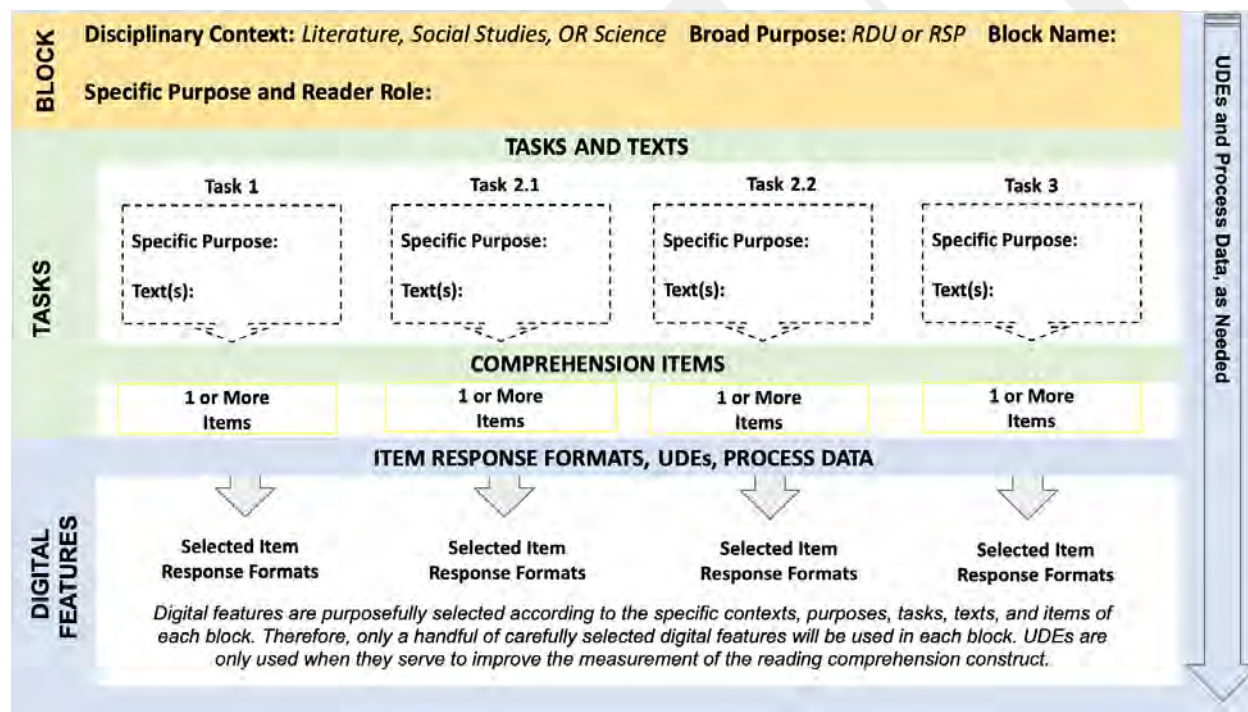
Considering the Range of Variations Within Assessment Components and Across a Block

When blocks are developed in accordance with the 2026 NAEP Reading Framework, the expectation, as outlined in Chapters 2 and 3, is that any of the components in a block (i.e., rows in the exhibit) can vary along a continuum, as depicted in Exhibit 2. That is, some blocks are more likely to include static texts and less cumulative tasks, items, and/or UDEs from one item to the next (left of center on the continuum), while other blocks are more likely to include dynamic/multilayered texts and more cumulative tasks, items, and/or UDEs from one item to the next (right of center on the continuum).

Exhibit 2 illustrates the continuum of design features from which developers might choose for each assessment component in the testing block. Note that within a given block, one component may have features that fall more on the left end of the continuum while features of

another component fall more on the right. Further, the complexity of different design features, and therefore of assessment components, may vary within a task. For example, for one task/text, the features might be less complex, but for a second task/text, they might be more complex. Or, for a single task/text, the purpose might be straightforward but the UDEs might be more complex. In all blocks, formats and features will continue to provide opportunities for readers to engage with an array of texts and tasks made possible in the digital platform used for all NAEP assessments.

Exhibit 2. Continuum of Variation in Features of Assessment Components Within a Block

Assessment Component	Less Dynamic and Cumulative Across Content and Format		More Dynamic and Cumulative Across Content and Format
Specific Reading Purposes	Purposes allow readers to focus attention on developing a deep understanding of a theme, question, or issue to be explored during the block. Not all tasks or items within the block necessarily work directly toward this theme, and there are opportunities for items to be less related to the specific purpose.	⟷	Purposes are paired with an essential inquiry question or problem to be examined throughout the task. All tasks and items within the block help readers work towards this theme, question, or problem.
Reader Role	Fewer parameters are specified for the reader's role. The reader is placed in a situation that provides fewer pieces of information about how to engage with the provided tasks and texts. The reader might be placed within a situation that contextualizes expectations for how to engage with provided texts and tasks. However, this situation provides less information about that role.	⟷	More parameters are specified for the reader's role within the block. The reader is placed in a situation that provides multiple pieces of information about how to engage with the provided tasks and texts. Readers may be assigned a particular role, and their role may be more specified, particularly in relation to reading purpose(s) and expected outcome(s).
Tasks	Purpose-driven tasks and items are situated in line with disciplinary context, but tasks are less related to one another	⟷	Purpose-driven tasks are situated in line with disciplinary context but tasks are more tightly structured so that one task builds on the previous;

111

	with less probability of readers moving back and forth across items within tasks; less need for resetting. Less involved culminating task, or no culminating task. Task not necessarily a determinant of all items in block.		more probability that tasks are interdependent; may have more need for resetting. More involved culminating task at the end of an activity that directly addresses the question or problem; major driver of the block.
Texts	*Number*: 1-3 topically related texts; excerpts may be included. *Dynamism*: More static texts with minimal dynamic features. *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.	⬌	*Number*: 2-4 topically related and interconnected texts may be included. Readers may be asked to choose only some texts to engage with and in line with task purposes. *Dynamism*: More texts with dynamic and/or or multimodal text features. *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.
Items	Items are less connected to the overall specific reading purpose for the block and there are more opportunities for items to be related, but less connected, to this specific purpose and to the related tasks; Less dynamic item formats to support less complex tasks and items.	⬌	Items are more connected to the overall specific reading purpose for the block. There are more opportunities for items to be more directly related to the specific reading purpose for the block and to the related tasks; More dynamic item formats to support more complex/multilayered tasks and items.
Universal Design Elements (UDEs)	Fewer cumulative reading purposes that may require UDEs for knowledge or motivation and potentially	⬌	More cumulative reading purposes that may require UDEs for knowledge or motivation and potentially greater need for task-based UDEs.

	lesser need for task-based UDEs.		
Process Data	Potentially fewer locations where process data involving reading actions could provide additional information about comprehension performance; sources may include, but not be limited to, timing data, navigation data (use of look back buttons), and use of varied item response formats.	⟷	Potentially more locations where process data involving reading actions could provide additional information about comprehension performance; sources might include, but not be limited to, timing data, more complex navigational practices across multiple sources and/or use of more dynamic item response formats.

Specific Guidelines for Block Development

Despite the range of variations in assessment components described above, as developers consider the different decisions they must make when designing a block, it is useful to keep the following points in mind:

1. Students deserve to know the tasks that lie ahead of them in the block. Guidance in the form of task-based UDEs is essential.
    a. Both purpose and reader role need to be made apparent at the outset of a block.
    b. Students should be reminded of purpose and role at the outset of each specific task within a block.
2. Since directions can be a source of construct irrelevant variance, they should always be conveyed in as accessible and straightforward a register as possible.
3. There is always a button available to allow students to listen to directions (or listen and read at the same time).
4. Just as expectations that students will be able to handle more complex text across the grades, so the expectations that they will be able to handle more complex guidance and activities also increases.
5. Cognitive labs, block tryouts, and pilot testing should ultimately guide NAEP in determining the optimal balance among these principles, especially when they come into conflict with one another. The experience in GISA and in the current 2019 operational NAEP SBT blocks offer an existence proof that these guidance features are manageable by 4th, 8th, and 12th graders. When these sorts of guidance features were included along with other UDEs in the 2017 special study, the enhanced blocks provided an overall comprehension performance advantage and resulted in higher motivational ratings by students, especially in the earlier grades. NAEP needs to monitor these matters with great vigilance.

Block Sketches
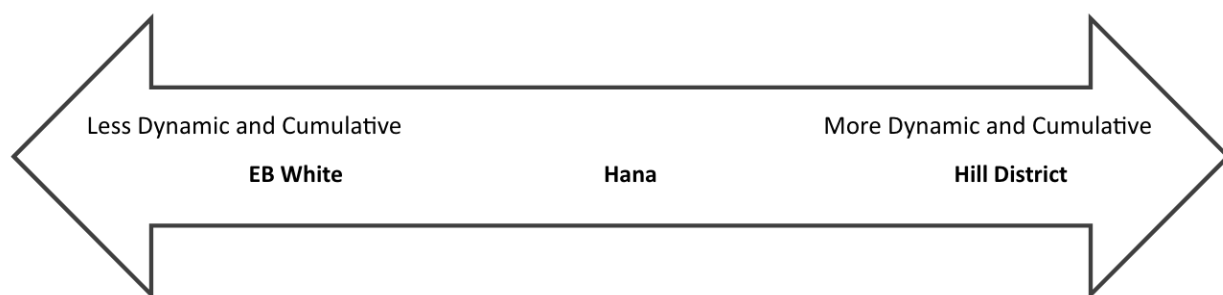
Sketches of three different blocks are provided to illustrate a range activity within assessment blocks that students might encounter when they participate in the 2026 NAEP

Reading Assessment. To accomplish this goal, the Appendix offers three hypothetical sketches of blocks (showing only a sampling of items from each) that might be developed using the components (from Chapter 2) and the design principles (from Chapter 3) of the 2026 NAEP Reading Framework. Importantly, these sketches are designed to exemplify key concepts from the framework and do not represent blocks or items that will be used on future NAEP assessments. Tasks presented with multiple sample items are provided to help readers of the framework envision how theoretical ideas in the framework might guide assessment design. However, these sketches do not represent fully expectations for enacting the NAEP style guide and other test specifications.

The first example (labeled *Hana* because it is built upon a short story text entitled *Hana Hashimoto, Sixth Violin* by Chieri Uegaki and Qin Leng) illustrates a block developed for the broad purpose of Reading to Develop Understanding (RDU~~) block, and the second example illustrates a~~). The second example (labeled *Hill District* because it is built upon a set of activities surrounding an authentic civic issue in the Hill District neighborhood of Pittsburgh, PA) illustrates a block developed for the broad purpose of Reading to Solve a Problem (RSP~~) block.~~

~~The first example outlines components in a Reading to Develop Understanding ()~~. And the third (labeled *EB White* because it is built upon a pair of texts, one *about* and one *by* the author E. B. White) illustrates a second, but more traditional, RDU~~) block in which fourth graders read to engage with texts in a literature context. In this block, fourth- grade~~ block. Referring to the underlying continuum of variation for assessment components within blocks as detailed in Exhibit 2 above, these three block sketches are situated on three hypothetical points along that continuum, as illustrated in Exhibit 3.

Exhibit 3. Underlying Continuum of Variation in Assessment Components in the Block Design for E.B. White, Hana, and Hill District Block Sketches

An overview of the three block sketches. As suggested, *Hana* exemplifies what features of assessment components in RDU blocks might look like at the center of the continuum. In this block, grade 4 readers preview a short video of young children playing in an orchestra ~~and then~~ (a motivational UDE) to pique their interest in playing the violin, the topic of the text. Then, they read and interpret story excerpts from the short story, Hana Hashimoto, ~~*Sixth Violin,*~~ by Chieri Uegaki ~~as they~~ in preparation for a book discussion with three peers. First, students are asked to read to develop an understanding of the ~~main~~ characters, key events, and author's craft ~~and~~. Second, they apply their insights to ~~predict events beyond the story.~~ describe what Hana is like as

114

a person. so that they are ready to contribute to the discussion.

~~The second example illustrates what eighth~~ The *Hill District* block includes features of assessment components more characteristic of those toward the right of the continuum that 12th graders might encounter in a ~~Reading to Solve a Problem (~~RSP~~)~~ block with texts situated in a social studies context. In this block, students engage in more ~~complex~~cumulative reading tasks that might include two to four more dynamic or multilayered texts and involve greater integration across texts and items, all of which contribute to a generative opportunity to use and apply meaning from ~~the text to solve a problem. While both assessment blocks include tasks, texts, items, and UDEs, differences in what readers experience illustrate just a sampling of the range of possible design features from which developers might choose in their creation of purpose-driven tasks embedded in any single block.~~ multiple texts to solve a problem.

~~**Grade 4 Assessment Block. Reading to Develop Understanding in a Literature Context: Hana Hashimoto, Sixth Violin**~~

EB White illustrates a second RDU block, but for an 8th grade literature context and with a more traditional look and feel than the *Hana* block. It retains many of the features students might encounter in commercially available standardized tests of reading comprehension, on state reading examinations, or on blocks characteristic of NAEP tasks developed from earlier frameworks. In fact, this example was created by using the two texts from a released 8th grade NAEP Block drawn from the 2011 NAEP Assessment.

When viewing these examples, it is important to keep in mind the following points:

- The purpose of these block sketches is to help readers of this 2026 Reading Framework develop an understanding of the range of comprehension activity and assessment components students might experience when they participate in the NAEP Reading Assessment.

- None of the examples is complete in the sense that all of the components and features are fully developed in the exact form in which they would appear on a finished test booklet. These examples are more like elaborated sketches that provide a preview of what each

115

block might look like, recognizing that not all of the actual items, UDEs, and other features are fully developed. Sometimes, for example, the type of UDE needed is specified but not actually provided (e.g., a particular word might make a plausible pop up vocabulary definition), or the type of comprehension item is indicated but not actually developed (e.g., an analyze/evaluate item is needed here to test students' understanding of the author's use of irony). In some cases (e.g., the Hill District block), two exemplars with different formats are provided to illustrate alternative ways to design task and item features in any particular block.

- While all three exemplar blocks include purposes, contexts, tasks, texts, items, and UDEs, differences in what readers experience illustrate just a sampling of the range of possible design features from which developers might choose in creating purpose-driven tasks embedded in any single block.

- Any given block, even a block that is situated toward one or the other end of the continuum (from Exhibit 7 in Appendix A), may have some features that lean more toward the center or even in the other direction. In other words, a given block might lean toward the traditional end of the continuum on texts (as does the Hana block) but toward the innovative end on item formats (as does Hana). The *EB White* block lends is otherwise classic RDU block, but lends itself to a Use/Apply culminating task (which is more characteristic of RSP blocks).

- The inclusion of the *EB White* exemplar has been included intentionally to reflect NAEP's commitment to maintain a healthy sample of tasks that feature print-based texts, RDU purposes, relatively few UDEs, and items that reflect the entire array of comprehension targets. As in all aspects of development, NAEP builds on its current strengths as it incorporates important developments in the nature of texts and tasks that students encounter in the ever-changing world of literacy.

Hana Hashimoto, Sixth Violin, Grade 4

The following example (not intended to be a complete block or to represent an actual NAEP Reading assessment) offers a sketch of what a Grade 4 Reading to Develop Understanding in a Literature Context block might look like. In the sketch, we walk through the assessment components described in the framework and illustrated in the block design visual (see Exhibit 4). These include the block components (context, purpose, grade level), the tasks (the tasks as well as the texts and items that students use to accomplish those tasks), and the digital features (item response formats, UDEs, and process data). In so doing, we describe how these components might be used by assessment developers when creating blocks to achieve some of the aims described in the framework.

Exhibit 4. Block Design for Hana

BLOCK

Disciplinary Context: *Literature, Grade 4* Broad Purpose: *Reading to Develop Understanding (RDU)* Block Name: *Hana*

Specific Purpose: Read to learn about what Hana is like as a person so you can participate in a book discussion with classmates
Reader Role: To work with three classmates to prepare for the book discussion

TASKS

TASKS AND TEXTS

Task 1
Specific Purpose: Learn about important events in the story and characters' thoughts, feelings, and actions
Text(s): *Hana Hashimoto, Sixth Violin*

Task 2
Specific Purpose: Write about what Hana is like as a person so that you are ready to discuss the book with peers
Text(s): *Hana Hashimoto, Sixth Violin*

COMPREHENSION ITEMS

6-7 Items

6-7 Items

DIGITAL FEATURES

ITEM RESPONSE FORMATS, UDEs, PROCESS DATA

Selected Item Response Formats:
Single and multiple selection multiple choice; fill in the blank; short constructed response

Selected Item Response Formats:
Short constructed response; matching; zones; short constructed response

Task-based UDEs: Teacher and student task characters; Task-reminder; Word bank
Motivational UDEs: Teacher and student task characters; String Instruments Video

UDEs and Process Data, as Needed

Block Components (Disciplinary* Context*, Purposes,* and *Tasks

Reader Role). This block is designed to assess how ~~4th-grade~~Grade 4 readers develop understanding within a single~~,~~ print text ~~by forming an~~ in a literary context. In this block, readers identify important events in the story and analyze how characters' thoughts, feelings, and actions describe the kind of people they are. Then, readers use and apply what they have learned to form an overall interpretation ~~about a story's~~of the main character, Hana. They choose a character trait from a word bank and then ~~applying that understanding to consider what might happen after the story ends. More specifically, readers are invited to engage with a group of fourth-grade students (represented by task characters in the assessment) 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. While many items give students opportunities to demonstrate their understanding and develop their thinking across the story, the texts (video and story) and items are relatively independent of one another. The test block also includes opportunities to develop understanding around other aspects of the story that may, or may not, contribute to that characterization. Throughout the block, 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~~explain how Hana fits that character trait based on the thoughts, feelings, and actions they have already interpreted.

Specific Reading Purpose(s) and Reader Role. At the beginning of the assessment (see Exhibit ~~1~~5), readers are ~~invited~~told that they will read the story *Hana Hashimoto, Sixth Violin*, by Chieri Uegaki and Qin Leng. Then, they are introduced to the specific purpose and reader role of reading to participate in a small book discussion group ~~about the story *Hana Hashimoto, Sixth Violin*~~ with ~~the~~ three ~~other 4th~~fourth grade ~~student~~ classmates (represented in the assessment by task characters~~. A~~ Gia, Gabe, and Luisa). They are also introduced to their teacher for the

117

project (represented by the task character ~~joins them to explain the discussion goal, which focuses on understanding how Hana grows and changes over the course of the story as a result~~Mr. Obas).

Then, a **task-based UDE** in the form of ~~events involving her classmates and her family. To~~two statements informs students what tasks will be expected of them. Here, students are told that, to prepare for the book discussion, ~~students are told~~they will read ~~parts of~~the story and ~~respond to items situated in three purpose-driven tasks to:~~ 1) ~~identify~~learn about 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~~characters' thoughts, feelings, and actions ~~in the story; and 3) apply their understanding of the characters in order to predict what might happen after the story ends.~~; and, 2) use what they have learned about Hana's to describe what she is like as a person. **Motivational UDEs** (here, student ~~task character classmates~~ and ~~a~~teacher ~~task character; see also bottom of Exhibit 1~~avatars and an introductory video) serve to ~~situate and~~ motivate readers to engage with the block.

Exhibit ~~1. Task-specific purposes~~5. Specific purpose, reader role, and ~~student~~ task characters serve to situate readers in a Grade 4 Reading to Develop Understanding block involving the short story *Hana Hashimoto, Sixth Violin* by Chieri Uegaki and Qin Leng

Introduction

You will read a book called *Hana Hashimoto, Sixth Violin* by Chieri Uegaki. After you read the story, you will participate in a small discussion group with three classmates to understand how Hana changes over the course of the story as a result of events involving her family.

To prepare for the book discussion, you will read the story while completing three tasks.

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.
3. Use your understanding of the characters to predict what might happen after the story ends.

Your teacher for this project will be Mr. Beadle

You will work with three classmates in your discussion group: Brian, Diana, and Michael

Welcome

You will read the story, **Hana Hashimoto, Sixth Violin**, by Chieri Uegaki and Qin Leng to prepare for a book discussion.

First, you will learn about **important events** in the story and **characters' thoughts, feelings, and actions.**

Then, you will **write about what the main character, Hana, is like as a person so that you are ready to discuss the book** with three peers.

Your teacher for this project will be Mr. Obas:

You will work with three classmates in your discussion group:

Gia   Gabe   Luisa

NEXT

Next, test-takers ~~are invited to~~ view a ~~30~~15 second video of young children playing in an orchestra found at a website designed for young children ~~(https://www.wonderopolis.org/wonder/why-do-orchestras-need-so-many-people~~ (see Exhibit ~~A~~26). This short video is an example of a **motivational** ~~and knowledge-based~~ **UDE designed to introduce readers** to the sounds and emotions one might experience when playing in an orchestra~~, while providing minimal background knowledge to students who may be less familiar with stringed instruments such as the violin.~~ .

119

Exhibit ~~2~~6. The teacher task character and a ~~pre-reading preview~~motivational UDE in the form of a ~~30~~15 second video clip of students playing stringed instruments serve to pique ~~students~~students' interest ~~and provide minimal background knowledge needed to make sense of the story~~

In this story, the main character Hana, decides to enter a talent show to perform the violin, a stringed instrument.
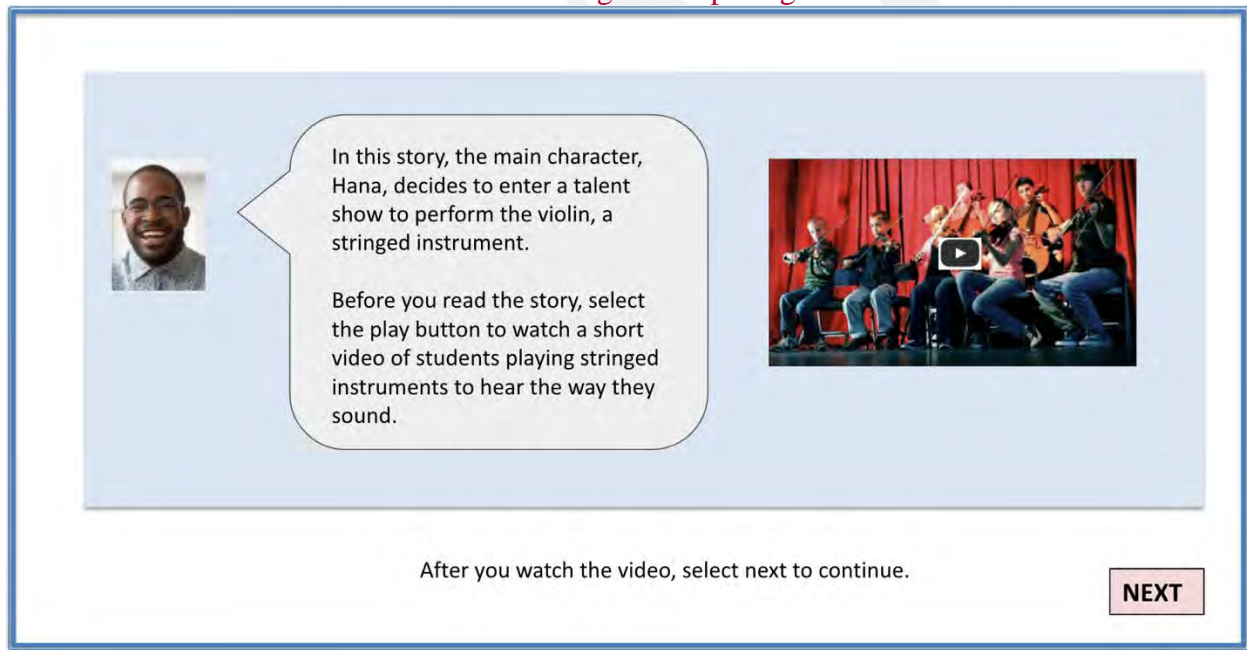
Before you read the story, clink the blue link to watch a short video of students playing stringed instruments, including the violin, to hear what it sounds like.

https://www.wonderopolis.org/wonder/why-do-orchestras-need-so-many-people

After you watch the video, click next to continue.   NEXT

In this story, the main character, Hana, decides to enter a talent show to perform the violin, a stringed instrument.

Before you read the story, select the play button to watch a short video of students playing stringed instruments to hear the way they sound.

After you watch the video, select next to continue.   NEXT

Task Components (Tasks, Text(s), and Items). After viewing the video about string instruments, students then begin reading the story and working through the tasks.

Tasks. After students are asked to read the story, the teacher reminds them of the specific reading purpose for the block (to prepare for a discussion) as well as the students' first task as they prepare for this discussion: learning about the events and characters (see Exhibit 7). In this case, the task reminder for the first task stays on the screen until students are ready to do the second task. At that point, the teacher offers a reminder of the second task, which is to write

about what Hana is like as a person. To do this, students are asked to use evidence from the story that they have already collected and interpreted on Hana's thoughts, feelings, and actions.

Text: Hana Hashimoto, Sixth Violin. In this story, a young girl named Hana signs up to play the violin in her school's talent show after having had only three lessons. Through the story, 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. When it comes time to play her violin in the talent show, Hana is at first nervous and thinks to herself, "This is going to be a disaster." However, as she looks out at the audience, she sees her friends and family. Then, Hana recalls her Ojiichan telling her to do her best and decides that is what she will do. She plays some of the everyday sounds she recalls her grandfather playing for her (e.g., a mother crow calling her chicks"). At the end of her performance, Hana takes "a great big bow." That night, her family asks her to play more of her sounds. The story ends with Hana playing her violin to herself before she goes to sleep, imagining the notes drifting out through her window and to Ojiichan in Japan while the author hints that Hana will keep practicing so that she might perform again in next year's talent show.

In the digital assessment format, readers can scroll through the story as they read, and the items appear aside the text so that readers can easily refer to the text as they complete the comprehension items. At the Grade 4 level, some illustrations from the original source text might accompany the story, as they do here (see Exhibit 7).

Comprehension Items. The array of items provides students with opportunities to develop their thinking across the story and demonstrate their understanding. Throughout the block, readers are asked to draw on textual evidence to make thoughtful interpretations of the text. The text and items are suitably independent of one another so that a student's performance on one item does not impact their performance on another item. The test block also includes opportunities to develop understanding around aspects of the story that may, or may not, contribute to the final task. Generally, however, the items help students work towards the specific purpose of the block (in this case, preparing for a book discussion), as well as the goal of each task. Exhibits 7-13 illustrate items that help students accomplish the first task of learning about the events and characters. Exhibits 14-16 illustrate items that then help students accomplish the second task of using what they have learned about the characters' thoughts, feelings, and actions to characterize Hana, in particular, by writing about what she is like as a person (see Exhibits 14-16).

Item response types ~~would~~ vary from simple multiple choice to short answer or hybrid constructed response items to give readers different kinds of opportunities to demonstrate their understanding in the block. **Sample questions** at this point ~~may~~might, for example, include single selection multiple choice items to assess readers' ability to locate and recall important events and other details (see ~~Exhibit 3) as well as~~ Exhibits 7 and 8), short constructed-response items that include fill in the blank options (see Exhibit 9), multiple selection multiple choice items (see Exhibit 10), and longer short constructed response items that ask readers to interpret and integrate ~~character traits~~details about the character's thoughts, feelings, and actions into their understanding of the story (see Exhibit ~~4~~11).

Exhibit ~~3. Example of~~ **7. A Grade 4 RDU block illustrating** a **Locate and Recall** multiple choice~~, locate and recall~~ item ~~in~~**. The teacher reminds the reader of the specific purpose (to prepare for** a ~~Grade 4 RDU block~~

How do Hana's brothers first respond to her decision to play the violin in the talent show?

A ○ They are happy for Hana. ⊖

B ○ They make fun of Hana. ⊖

C ○ They promise to help Hana. ⊖

D ○ They are angry with Hana. ⊖

Exhibit 4. ~~Example of a short constructed-response interpret~~**discussion) and** ~~integrate item in a Grade 4 RDU block~~ **the first task (to learn about events and characters)**

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 with facts from the story.

To prepare for the discussion, first read the short story and learn about the events and characters.

Hana Hashimoto, Sixth Violin
By Chieri Uegaki & Qin Leng

When Hana Hashimoto announced that she had signed up for the talent show and that she would be playing the violin, her brothers nearly fell out of a tree.

"That's just loopy," said Kenji. "You're still a beginner."

"Stop kidding," said Koji. "You can barely play a note."

"It's a *talent* show, Hana."

"You'll be a disaster!"

Hana squared her shoulders and took her violin and bow inside, leaving her brothers laughing like monkeys in the tree.

What does Hana want to do for the talent show?

A ○ Sing a song ⊖

B ○ Tell jokes ⊖

C ○ Play the violin ⊖

D ○ Climb a tree ⊖

NEXT

122

Within the block, **knowledge-based UDEs** might include pop-up boxes providing a hint about the meaning of certain ~~domain specific words or general topics of a text~~vocabulary terms that are not tested (in this case, describing what a ~~talent~~talent show is, see Exhibit ~~5) so~~8). This provides readers ~~are provided the minimal~~with some background knowledge from which to make sense of the story ~~and engage~~. This allows readers to focus on engaging with items designed to measure, in this case, their ability to make inferences about characters from their thoughts, feelings, and actions in the story.

Exhibit 8. A Grade 4 Integrate and Interpret item illustrating a knowledge-based UDE in the form of a pop-up box defining the term "talent show." The blue pop-up box appears when a test-taker clicks on the underlined term. This example also illustrates a multiple choice integrate and interpret item

Exhibit 9. A Grade 4 Locate and Recall item illustrating a fill in the blank short constructed response

Exhibit 10. A Grade 4 Locate and Recall item illustrating a multiple selection multiple choice response format

Hana Hashimoto, Sixth Violin
By Chieri Uegaki & Qin Leng

Hana practiced every day, just like Ojiichan. And every day, her brothers fled the house with covered ears, complaining about the horrible noise.
She practiced in front of her parents, who listened with care while they washed and dried the dishes.
She practiced in front of her dog, Jojo, who cocked his head and sometimes growled at the strange sounds Hana made.
And she practiced on her own, in front of an old photo of Ojiichan from his symphony days. Alone, Hana could pretend she was performing in front of an audience so appreciative they called for encore after encore.

For whom does Hana practice? Select **all** that apply.

A ○ Her friends

B ○ Her brothers

C ○ Her dog

D ○ A photo of Ojiichan

NEXT

In addition, **a look-back button (ora task-based UDE)** is embedded into theitems with excerpted quote in this item; iftext (see Exhibits 11 and 12). If readers choosewish, they can click on the underlined quote to see exactly where the excerpted text is located in the context of the original story passage in the assessment space. Multiple choice and constructed response item formats are interspersed throughout the assessment.

Exhibit ~~5~~11. A Grade 4 ~~Integrate~~Analyze and ~~Interpret~~Evaluate item illustrating a task-based UDE in the form of a look-back button that refers ~~the reader~~readers to the relevant section of text within the story and a ~~knowledge-based UDE in the form of a pop-up box defining the term "talent show" for the text _Hana Hashimoto, Sixth Violin_ by Chieri Uegaki. The blue pop-up box appears when a test-taker clicks on the highlighted term.~~

The text says, "When Hana Hashimoto announced that she had signed up for the **talent show** and that she would be playing the violin, her brothers nearly fell out of a tree...Hana practiced every day, just like Ojiichan. And every day, her brothers fled the house with covered ears, complaining about the horrible noise."

A **talent show** is a show in which different people perform a special skill.

Thinking about this part of the text, why do you think Hana's brothers flee the house every day?

    a. They are angry with Hana
    b. The noise of the violin bothers them
    c. They have somewhere important to be
    d. They like the way the violin sounds

As depicted in Exhibit 6, students could also be given a word bank (a task-based UDE) ~~from which to select relevant character traits when asked to describe the kind of person Hana is. Then, in an analyze and evaluate item with a hybrid~~ short **constructed response format**~~, students could be asked to use that word to describe Hana's actions and then explain their thinking using evidence from the story.~~

~~Exhibit 6.~~ **A Grade 4 Analyze and Evaluate item illustrating a task-based UDE in the form of a word bank providing a set of character traits from which readers can select their choice and then use it as part of their answer in the box.**

~~Students could also be given a timeline on which to drag and drop their responses about how the main character changes over the course of the story (see Exhibit 7). A collection of relevant and irrelevant notes about the character can be provided from which students can select the best answers. Once completed, students would then have access to this informational graphic as a writing support when answering the final Use and Apply item (see Exhibit 7).~~

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, lowing cows, 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 more musical notes around the dinner table. The story ends when Hana recalls the numerous songs her Ojiichan shared with her and imagines what she might play in next year's talent show.

Exhibit 12. The items for the first task help students develop an understanding of the events and characters as in this Grade 4 Integrate and Interpret short constructed response item

Exhibit 13. A Grade 4 Integrate and Interpret Item for the first task using a single response multiple choice format

The story ends when Hana recalls the songs her Ojiichan shared with her and imagines what she might play in next year's talent show. At this point, students are invited by the teacher to start the second task, which is to write what Hana is like as a person in preparation for the book discussion (see Exhibit 14).

One of the classmates (a task character in the assessment) acts as a **motivational UDE** to motivate the student to engage in collecting notes for the second task, as the classmate has already completed part of the activity. The task character also acts as a task-based UDE in reminding the student that they should use specific details from the story about Hana's thoughts, feelings, and actions. Once completed, students have access to the full set of notes, as these completed notes are transferred to the next item (see Exhibit 15).

Exhibit 14. Teacher and student task characters remind readers of the second task goal in this Integrate and Interpret item

In Exhibit 15, the other two classmates serve as **motivational and task-based UDEs** to engage students in the task while also reminding them to stay focused on the character's thoughts, feelings, and actions. The student's responses from the previous item are carried over to the next item as the completed notes, which also serves to motivate the student since they have already completed the work. These notes could also be "reset" if the student did not enter appropriate notes in the previous item so that the student's score on this item is not dependent on how they responded previously.

In Exhibit 15, the student is asked to move the notes from their notepad into the chart as they sort the notes into Hana's thoughts, feelings, and actions in preparation for writing about the kind of person she is. In the final task (see Exhibit 16), the student has access to this chart as a writing support when they answer the final use and apply item. Again, notes that are incorrect are reset so that the final item is not dependent on the way they responded to this one.

Exhibit 15. The student's responses from their completion of the previous item are carried over to the next item as the completed notes. A graphic organizer with drag and drop features offers students an efficient way to demonstrate their understanding of how the text conveys the character's thoughts, feelings, and actions in this Grade 4 Integrate and Interpret item

A longer constructed response item such as the example shown in Exhibit ~~8~~16 is designed to assess readers' ability to Use and Apply ~~their understanding to a new situation beyond~~understandings learned from the story ~~itself. In~~to form a characterization of Hana. As readers engage with this final part of the ~~assessment~~ block, ~~after listening to one of the student task characters orally~~the teacher invites them to use their chart (which they have access to) to write what Hana is like as a person in preparation for the discussion.

Then, as depicted in Exhibit 16, in a use and apply item with a hybrid constructed response format, students are given a word bank (a task-based UDE) from which to select a relevant character trait (these could be hot spots; in other words, when readers click on them, the word is highlighted and gets recorded as the student's answer to Part A) when asked to describe ~~how Hana reacted to her brothers' behavior earlier in the story, readers are invited to join the discussion group with three task character classmates and contribute their ideas.~~ the kind of person Hana is. Instead of spending time generating character trait words (which is not part of the construct this item aims to measure), the student can select from those provided. This allows the student to focus their limited time and cognitive resources on applying evidence from the text about Hana's thoughts, feelings, and actions to an analysis of the kind of person Hana is.

Exhibit 16. This final, two-part Use and Apply item illustrates the use of a task-based UDE in the form of a word bank of character traits as well as a hybrid item format where students select a choice and write about it. Students use what they have learned from the text about Hana as a person and apply that understanding to draw a conclusion about the kind of person she is.

Part A. Select a **character trait** from the word bank that best describes Hana.

WORD BANK

helpful   curious
brave   proud
smart   nervous
afraid   confident
forgetful determined

Hana Hashimoto Story

Completed Chart

Part B. Explain how Hana can be described using the character trait you selected in Part A. Be sure to use evidence from your chart about **Hana's thoughts, feelings, and actions.**
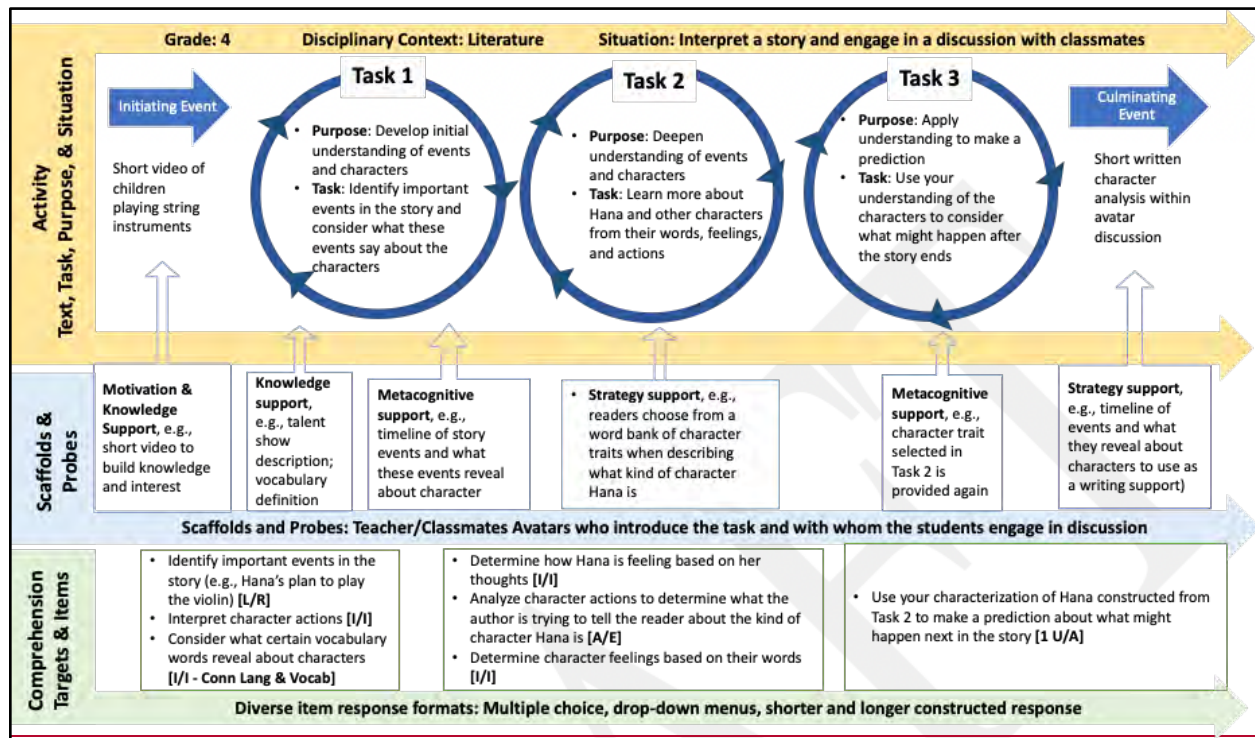
DONE

Performance Evidence and Indicators

. When interpreting reading achievement from performance on the 2026 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 block would be classified with a primary disciplinary context, grade level, and broad purpose. Scores from the Hana Hashimoto, Sixth Violin block, then, ~~can demonstrate~~describe what Grade 4 students can do in a literature context as part of a Reading to Develop Understanding block. The block is designed to measure ~~their~~students' ability to develop their understanding ~~within~~of a single text and then apply that understanding ~~to~~in a simple culminating event (in this case, ~~making a prediction,~~describing the kind of person Hana is based on her thoughts, feelings, and actions in the story~~, about what will happen after the story ends).~~).

Test developers keep ~~an elaborate~~a detailed account of all decisions that go into classifying texts and generating items from comprehension targets in each block. This process enables NAEP to 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.
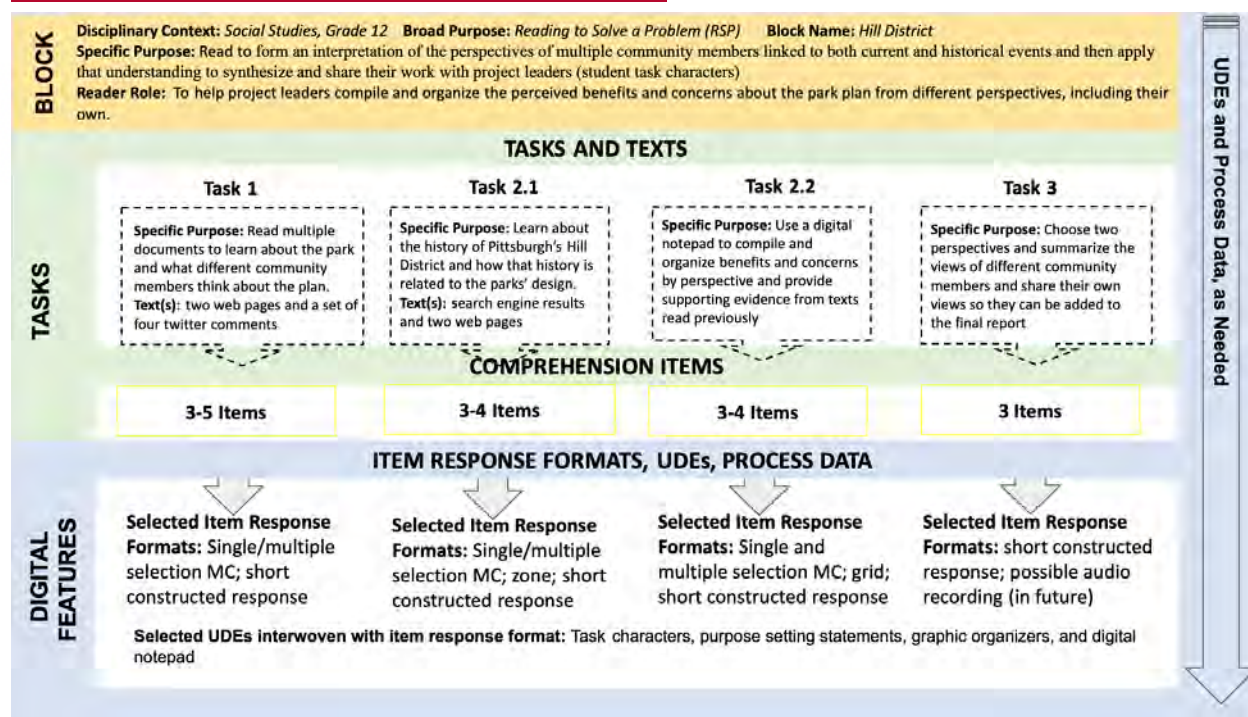
Hill District, **Grade** 8 Assessment Block: Reading to Solve a Problem in a Social Studies *12*

Block Components (Context

Context

, Purposes, and Reader Role). This block is designed to assess how 8th12th grade readers develop understanding across multiple texts in a social studies context by forming an interpretation of the perspectives of multiple community members linked to both current and historical events and then applying that understanding to solve a problem. (See Exhibit 17 for the block design and Exhibit 18 for the introduction to the block).

Exhibit 17. Block Design for Hill District Sketch

More specifically, readers are invited to engage with ~~a group of~~three students (represented by task characters in the assessment) who ~~are motivated to learn about a current civic project deeply rooted in their city's history: The City of Pittsburgh has recently announced~~ have been asked by the Mayor to compile and organize public reactions to an ambitious plan ~~for~~proposed by the City of Pittsburgh. Known as the "I-579 Cap Project," the plan involves 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.

~~Park designers at a landscape architecture firm have created a proposed park design.~~

~~Test takers are asked to learn about this project by considering the role of a key aspect~~The tasks in this Reading to Solve a Problem block reflect design features that are more dynamic and cumulative in terms of content and format, as depicted toward the right side of the continuum in Exhibit 2. For example, readers are constrained by specific purposes and role expectations about how to engage with provided texts. The four tasks (and related sub-tasks) are tightly structured so that one task builds on the previous, such that readers are asked to learn more about the project goals and get a general sense of the public's comments before they are asked to gain a deeper understanding of the historical significance of the proposed park ~~design: the inclusion of a 13-year-old African-American girl named Keisha who appears on illustrated signs throughout the park. Park designers have proposed including signs of Keisha in many park locations to provide details about the African-American community's history in the Hill.~~.

The test block also includes opportunities for students to engage with several interconnected digital texts (e.g., excerpts from social media, search engine results, and multimedia websites and online news articles) that represent the perspectives of different kinds of community members and cuts across issues of contemporary and historical relevance. Throughout the block, readers ~~engage with a collection of~~ XX ~~historical and contemporary multimodal texts to develop~~are asked to activate and employ their personal, cultural, and civics

134

knowledge and resources by drawing on textual evidence in multiple modes to make thoughtful interpretations and evaluations of the text. Of note, several UDEs and dynamically formatted items are designed to motivate and guide students through the series of challenging assessment tasks in a multilayered digital environment.

Specific Reading Purpose(s) and Reader Role. At the beginning of the assessment (see Exhibit 18), students learn that the city has recently unveiled the park plan to the public on its website and city residents have been invited to share their reactions on various social media. Students are also introduced to three high school aged task characters selected by the Mayor to help compile comments in preparation for a series of public working meetings (see Exhibit 19). In a school partnership with the city, the three high schoolers have invited other students to help them organize comments from different community members. This situation inspires the question/problem that guides readers' inquiry in the assessment block: How do different community members feel about the proposed park project and what interests inform their comments?

Exhibit 18. A social studies context and reader role serve to situate readers in a Grade 12 Reading to Solve A Problem block involving several interconnected digital texts

Exhibit 19. Same-aged task characters and a task-based UDE in the form of four task-specific purposes serve to guide and motivate readers in the RSP block

Your Task

You will work with three high school students who were selected by the mayor to lead the project:

Kai   Moises   Jasmine

To accomplish this goal, you will do four tasks:

1. Read multiple documents to learn about the park plan and what different community members think about the plan.

2. Learn about the history of Pittsburgh's Hill District and how that history is related to the park's design.

3. Describe some of the benefits and concerns about the park from different perspectives, or viewpoints, including your own.

4. Share your work with the student project leaders for a meeting with the Mayor.

NEXT

Task Components (Tasks, Text(s), and Items).

Tasks. To support their inquiry, students are told they will read multiple documents and respond to items situated in four purpose-driven tasks to: a) learn more about the proposed park plan and keep notes about what different community members think about the plan; b) learn about the history of Pittsburgh's Hill District and how that history is related to the park's design; c) synthesize some of the benefits and concerns about the park from different perspectives, including their own and d) share their work with the student project leaders for a meeting with the Mayor. Several task-based UDEs (e.g., graphic organizers and purpose setting statements) and motivational UDEs (three student avatars, a recent event, and an ~~understanding of~~ opportunity to express their own opinions about the ~~Hill District's history~~project) serve to guide and ~~then clarify~~motivate readers to engage with the ~~planned vision of~~block.

Texts. After learning about the four task-specific purposes in this social studies block, readers engage with a digital text set that contains important information and viewpoints related to the proposed park ~~to different~~plan. These include Twitter comments from community members ~~of the~~; a set of search engine results and pull-down menu items from a website; and text passages on websites about the project embedded with comments from Pittsburgh ~~community. While some of the items~~residents, photographs, a short video, and an artist's rendering of the park plan. With each new text, readers learn more about proposed features of the park plan that help to build their understanding of how different community members view the park's features from various perspectives and how the history of Pittsburgh's Hill District is relevant to the park's plan.

Comprehension Items. Item response types would vary from simple multiple choice to short answer or hybrid constructed response items to give ~~students~~readers different kinds of opportunities to demonstrate their understanding ~~and~~ in the block and apply that understanding to

136

solve the problem. While some items give students opportunities to demonstrate their understanding and develop thinking within a specific text, other items are designed to assess how readers ~~integrate ideas~~navigate and ~~perspectives~~make meaning across sources representing multiple ~~sources~~ and diverse perspectives.

After being asked to read text and watch a short video on a website about the park project (Exhibit 20), sample questions may, for example, include single or multiple response formats for multiple choice items that ask readers to locate and recall important details about the project from the passages and the video (Exhibits 21 and 22). Other questions might assess students' ability to integrate and interpret textual and visual information from an artist's rendering of the site improvement plan on a different website (see Exhibit 22). **Task-based UDEs** (e.g., one of three task characters) provide short prompts (shown at the top of Exhibits 20 and 23) designed to cue the reader about the steps they are completing as they read across different sources to solve the problem.

Exhibit 20. A Grade 12 RSP block illustrating the directions that readers are asked to follow as they engage with texts and items. The task character reminds the reader of the specific purpose and the first task

Exhibit 21. A Grade 12 Locate and Recall item illustrating a multiple-selection multiple choice response format

Exhibit 22. A Grade 12 Locate and Recall item illustrating a single-selection multiple choice item response format

Exhibit 23. **Two Grade 12 items that ask readers to Integrate and Interpret (item 1) and Locate and Recall (item 2) textual and visual information from an artist's rendering of the site improvement plan published on a website**

Examples of short constructed-response items earlier in the block might ask readers to integrate and interpret information about how park designers plan to modify the city's use of natural resources to address environmental concerns (Exhibit 24). Later in the block, readers might be asked to integrate and interpret information in an online newspaper article about the historical significance of the park's design (Exhibit 25) or to analyze and evaluate the requests of some community members to include park features that honor the history of their neighborhood (Exhibit 26). Also depicted in Exhibit 26 is a **task-based UDE** in the form of a task character that serves to remind students of their reading purpose in the second task.

Exhibit 24. A Grade 12 RSP short constructed-response item that asks readers to integrate and interpret information about how park designers plan to address environmental concerns

Exhibit 25. A Grade 12 short constructed-response item with a look-back button (task-based UDE) that asks readers to integrate and interpret information in an online newspaper article about the historical significance of the park's design

Exhibit 26. A Grade 12 short constructed-response item that asks readers to integrate and interpret information on a web page with a look-back button (task-based UDE). The task character reminds readers of the specific purpose of the second task

Other potential items might ask readers to locate and evaluate the relevance of search engine results pertaining to the historical significance of some of the park's features (see Exhibit 27) or locate (navigate to) and then analyze information from a website's menu to evaluate the expertise of the group responsible for publishing information about the park project (see Exhibits 28 and 29 respectively). Both of these tasks and items can be designed to collect timing and navigation process data about the choices readers make as they navigate multilayered digital environments such as search engines and websites with menus.

Exhibit 27. A Grade 12 selected response zone item designed to capture process data about which link is selected paired with a short constructed response scored item that asks readers to analyze and evaluate the relevance of their search engine choice

TASK 2. Now, can you help us with some Internet research to find background information to share?

Directions: Read the list of Google Search Results. Choose the link most likely to have information about the history of Pittsburgh's Hill District and why this history is relevant to the park's plan. Then, use the box to explain why that link is the best choice. Use evidence from the search engine results to support your thinking.

https://www.bostonmagazine.com › news › 2017/04/05
A Park Beneath the I-93 Overpass Will Open in June - Boston
Apr 5, 2017 — A new park, called Ink Underground, is scheduled to open under the I-93 overpass in June. It will stretch eight acres and feature public art.

https://archive.triblive.com › local › pittsburgh-allegheny
Pittsburgh City Council OKs 'cap' park over Crosstown ...
Jul 2, 2018 — A project to build a 'cap' over Pittsburgh's Crosstown Boulevard and ... A proposed 'cap' linking the Lower Hill District and downtown Pittsburgh would include a city park and ... office redevelopment plan for the 28-acre former Civic Arena site. ... The Shrines of Pittsburgh

https://pittsburghpa.gov › dcp › hill-district
The Greater Hill District Master Plan | pittsburghpa.gov - City of ...
The Neighborhood. As one of Pittsburgh's earliest and largest neighborhoods, the Hill District continues to play an important role in the story of African-Americans ...

https://www.post-gazette.com › 2019/06/14 › stories
Righting a wrong: New park over I-579 to reconnect ...
Jun 14, 2019 — A $32 million effort to "right a wrong" that occurred half a century ago gets its start Friday. A slew of politicians and stakeholders will gather in a parking lot near PPG Paints Arena to break ground on a three-acre park that will straddle Interstate 579/Crosstown Boulevard.

Your explanation:

NEXT

Exhibit 28. A Grade 12 item selected response zone item designed to capture process data about how readers navigate through hyperlinked web pages

TRIB LIVE | SEA | Righting a Wrong | Next City

Pittsburgh City Council on Tuesday authorized about $32 million in spending for the Interstate 579 "cap" designed to link the Lower Hill District and Downtown via a three-acre park.

PennDOT will oversee the work expected to start by July and end in late 2021, according to the Sports and Exhibition Authority of Pittsburgh and Allegheny County.

Plans call for handicapped-accessible pedestrian pathways, an amphitheater, stage, lawns, landscaping, recreation areas, art and replacement of walkways in the area.

"The 'cap' project will be transformative for the Hill District by removing a physical barrier and re-establishing connectivity to centers of employment, education and services in Downtown Pittsburgh," according to the SEA website.

I-579, known as the Crosstown Expressway, runs between the Hill and Downtown. The cap will essentially serve as a large bridge deck made of 8-inch-thick, reinforced concrete slabs supported by beams and pillars.

It will complement the Penguins' long-awaited $450 million residential, retail and office redevelopment plan for the 28-acre former Civic Arena site.

The SEA in 2016 received a $19 million federal grant for the work and is kicking in an additional $5.2 million. The remaining funding is coming from the state.

Council unanimously approved the allocation without comment.

Let's try to find out more about the Sports and Exhibition Authority (SEA) to see why they might be spending so much money to support the park project.

Directions: Select the link that will tell you more about the Sports and Exhibition Authority.

142

Exhibit 29. A Grade 12 critical online resource evaluation item that asks readers to analyze and evaluate the extent to which an organization has the appropriate qualifications to publish details about the proposed park plan on their website using a hybrid constructed response

Dynamic response items in the testing block can also be used to capture process data (e.g., how long students take to complete the item and the order of selections and answer changes) while assessing reading comprehension performance. The item in Exhibit 30, for example, asks readers to analyze and evaluate a small set of comments shared on social media in order to characterize the interests of different community members in relation to the proposed park plan. In this context, the drag-and-drop dynamic response format provides two additional functions; it serves as an alternative to writing each response as well as a **task-based UDE** to guide the language students use to classify comments into categories of accurately worded perspectives. This particular task-based UDE is also designed to introduce students to perspectives they will be asked to consider later in the testing block as part of the culminating Use and Apply task.

Exhibit 30. A Grade 12 dynamic response item that asks readers to analyze and evaluate four comments on social media. The drag-and-drop response format serves as an alternative to writing and also serves as a task-based UDE to guide students' classification of items into categories of accurately worded perspectives

As was noted in Chapter 3, NAEP should continue the trend of exploring the use of other interactive or dynamic response formats made possible with emerging digital tools. To that end, the next pair of items (Exhibits 31 and 32) serves to provide an illustrative example of how task-based UDEs might be used alternatively to compare how readers engage with comprehension items that use different types of response formats.

In both instances, readers are asked to categorize comments from community members about the park project and the intentional pairing of motivation and task-based UDEs serve to guide students and sustain their willingness to persist with multiple document inquiry tasks. Exhibit 31 applies a multiple-selection response format with a **task-based UDE (table) and motivational UDE (task character)** that serve to support readers as they engage in one particular item in the block. That is, the table is designed to first help readers focus their attention on relevant comments on the left side (rather than referring back to them in the original text) and then ~~apply their understanding to the task at hand.~~, match each comment with one or more specific benefits on the right.

~~At the beginning of the assessment block (see Exhibit XXX), 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. City residents have posted comments and questions (depicted in a series of twitter posts).~~

~~Insert something here about the motivational UDE's in the authentic Twitter posts and the overpass knowledge-based UDE and the role they play in the early part of the block.~~

The four twitter posts are designed to inspire the question that guides readers' inquiry in the block: Why does Keisha matter to the city park project? As test takers proceed, they are introduced to this question (see Exhibit XXX) and the four task-specific purposes for engaging with the texts and comprehension items in the block: 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).

Exhibit 10. Readers are Situated Within a Disciplinary Context and Broad Purpose in the Reading to Solve a Problem *Hill District* Block

Your goal is to help the public understand the history of Pittsburg's Hill District and generate an answer to this question: **Why does Keisha matter to the park city project and what role does she play in fulfilling the park planner's vision?**

To accomplish this goal, you will:

1. Explore the background history of Pittsburg's Hill District to understand what has happened to "The Hill" from the 1940's to the present.
2. Examine different perspectives about the proposed park plan to determine what park visitors might consider confusing or concerning about the plan.
3. Apply your understanding of the issues to imagine how you would respond to the public's concerns.

You will work with Mr. Ramirez, a regional historian, and three teens: Brian, Diana, and Michael

Mr. Ramirez

NEXT

Tasks

Readers are asked to engage in purpose-driven tasks 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 encounter multiple texts about the history of the Hill District, to help 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 asked 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 carry 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 and texts 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.

When choosing texts, test developers take into consideration the length and level of complexity to ensure selected texts and related tasks are suitable for 8th graders completing the entire block in 20-40 minutes (e.g., passage length, structures, vocabulary, knowledge demands, motivational features).

Items

Comprehension processes are identified throughout the block 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 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 block with variations in attention paid to unique text features and task demands as well as qualitative differences within each comprehension target category.

Universal Design Elements

As shown in Exhibit 4.19, the block design includes a range of digitally enhanced UDEs as readers comprehend texts, respond to items, and reflect on their performance. In the initial stage, an task character (a regional historian designated as a knowledge-based UDE) presents the reader with a primary purpose for reading; then, the reader (alongside task character classmates that represent motivational UDEs) 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.

We can insert a visual that illustrates what these task characters might look like with these dialogue prompts.

Task-based UDE's 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). These task-based design elements (a graphic timetable and a searchable resource compilation) also serve as motivational UDE's in that they are designed to assist with organizing and analyzing information throughout the testing block while also helping to facilitate real-world connections and sustain 8th graders willingness to persist in this block's challenging collection of tasks.

We can insert a visual that illustrates what this timeline image might look like next to a list of keywords. Julie could draft an idea?

Diverse but intuitive response formats can be selected to facilitate reader engagement and reduce the cognitive memory load involved in expressing responses to test items designed to measure comprehension performance. Students are likely to benefit from embedded task guidance provided by task character guides and/or a graphical overview of block-specific reading

147

In contrast, Exhibit 32 engages readers in a similar matching process, but for this item, a task character **(motivational UDE)** ask readers to move each comment into the appropriate cells of a table that is part of a retractable digital notepad **(task-based UDE** marked near a blue arrow to illustrate how it can be minimized and maximized on the screen as needed); readers use the notepad to store, organize, and recall important details as they read across multiple sources to solve the problem at hand. Similar to how students engage in reading across multiple documents outside of a testing environment, the digital notepad enables students at several points in the testing block to click on the notepad (which makes the table appear) to add and organize details as they continue to learn more and build a deeper understanding about how different community members feel about the park project from their varied and diverse perspectives. Exhibit 33 illustrates how the same notepad could have been paired with a different item earlier in the task when students were reading on a different website.

Of course, as was also noted in Chapter 3, when selecting the format of any particular item, developers 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. Pairing the development of any innovative task-based UDEs with careful piloting efforts will ensure that design features yield their intended outcomes for as many students as possible.

Exhibit 31. A Grade 12 multiple-selection response grid item with a task-based UDE (table) and motivational UDE (task character) that serve to support readers as they engage in one particular item in the RSP block

Exhibit 32. A Grade 12 dynamic matching response grid item with a motivational UDE (task character) and task-based UDE (retractable digital notepad) that serve to support readers at multiple points in the RSP block as they read across multiple sources to solve the problem at hand

I found a lot of different opinions and perspectives on the park in this website. If you can help organize these by topic, I'll add them to our summary report for the Mayor.

Directions. Below are comments from two community members. Determine which perspective best applies to each comment and if that person's comment would be considered a benefit or concern. Then select and drag each comment to the appropriate box in the table in your notepad.

Notepad

Longtime Hill District Resident Brenda Tate: For Brenda Tate, who has lived on the same block of Webster Avenue in the Hill for all of her 70 years, the park once again will give her the chance to traverse Wylie Avenue to the park then into Downtown and back. "There won't be separation. There will be a clear avenue to come back and forth. It's symbolic," she said.

Ms. Tate, who with her 98-year-old aunt will be attending Friday's groundbreaking, sees positives in the park's construction. "It will be a nice green space, a welcoming space, for people who want to come into the community," she said. (*supportive member of the Hill District*)

City Councilman R. Daniel Lavelle: "What we're going to begin doing [Friday] is finally righting those wrongs of 50 or 60 years ago," added Mr. Lavelle, who represents the Hill. While the park is important, Mr. Lavelle said the greater value lies in providing business and job opportunities within the arena redevelopment for Hill residents and minorities. (*city councilman who represents the Hill district*)

	Benefits	Concerns
Recreational Perspective		
Environmental Perspective		
Economic Perspective		
Historical Perspective		

NEXT

Exhibit 33. A Grade 12 dynamic matching response grid item with a task-based UDE (retractable digital notepad) that serves to support readers at another point in the RSP block as they read across multiple sources to solve the problem at hand

TRIB LIVE	SEA	Righting a Wrong	Next City

Pittsburgh City Council on Tuesday authorized about $32 million in spending for the Interstate 579 "cap" designed to link the Lower Hill District and Downtown via a three-acre park.

PennDOT will oversee the work expected to start by July and end in late 2021, according to the Sports and Exhibition Authority of Pittsburgh and Allegheny County.

Plans call for handicapped-accessible pedestrian pathways, an amphitheater, stage, lawns, landscaping, recreation areas, art and replacement of walkways in the area.

"The 'cap' project will be transformative for the Hill District by removing a physical barrier and re-establishing connectivity to centers of employment, education and services in Downtown Pittsburgh," according to the SEA website.

I-579, known as the Crosstown Expressway, runs between the Hill and Downtown. The cap will essentially serve as a large bridge deck made of 8-inch-thick, reinforced concrete slabs supported by beams and pillars.

It will complement the Penguins' long-awaited $450 million residential, retail and office redevelopment plan for the 28-acre former Civic Arena site.

The SEA in 2016 received a $19 million federal grant for the work and is kicking in an additional $5.2 million. The remaining funding is coming from the state.

Council unanimously approved the allocation without comment.

Directions: Click the notepad to open your notes. Then answer this question.

What do people who like to exercise think about the proposed park plan?

Use your notepad to briefly describe a benefit and a concern about the park plan from a recreational perspective. Use details from the text and the video to support your answer.

Notepad

	Benefits	Concerns
Recreational Perspective		
Environmental Perspective		
Economic Perspective		
Historical Perspective		

NEXT

149

Culminating Task. Toward the end of the Reading-to-Solve-A-Problem task, the three task characters remind students they are close to accomplishing their goal. In the first part of the task (Exhibit 34), students are asked to use what they learned about what different community members think about the proposed park plan (as stored in their digital notepads) and apply that understanding to provide evidence-based descriptions of their benefits and concerns from a certain perspective in order to help the task characters submit their final report to the Mayor. By suggesting "this is a big task so can you help with two of the perspectives and then I'll find the other three?", the high-school aged avatars recognize the difficulty of the task and provide support, as a **motivational UDE,** while still asking students to demonstrate their ability to use and apply what they have learned about the views of different community members in preparation for the final report. Readers are also reminded that they have access to the four websites they read and their digital notepad **(task-based UDEs)** to help them accomplish this culminating task.

For the second part of the task, students are asked to share their own evidence-based views of the park proposal plan and the task characters promise to also include their opinions in their final report. This item serves to validate the student's own voice and agency as an important contributor to the group's final summary. Exhibit 35 illustrates how this item might look using a short-constructed response format, similar to those in existing NAEP assessment blocks, and Exhibit 36 is included to depict what an item might look like longer into the future, as NAEP continues to explore alternative response formats that offer authentic opportunities for students to choose their preferred response format (e.g., written or audio recording) to express their own opinions to the problem posed by this testing block. Again, pairing the development of these innovative features with new considerations for scoring and careful piloting efforts will ensure that design features yield their intended outcomes for as many students as possible while never unintentionally disadvantaging some populations of students.

Exhibit 34. This Use and Apply item with open-constructed response format illustrates the use of a task character (motivational UDE) that reminds students of their goal, recognizes the difficulty of the task, and provides support.

Exhibit 35. This final Use and Apply item with open-constructed response format illustrates the use of a task character (motivational UDE) who reminds students they have accomplished their goal and validates the test-taker's role by inviting them to use what they learned and apply that understanding by sharing their own opinion.

Task 4. Now that you know more about the features of the park plan and the perspectives of Pittsburgh residents, you probably have your own opinion too! We'd love to include your own opinions in our final summary for the Mayor too.

Notepad

Directions. Imagine you lived in Pittsburgh and will attend the community meeting to express your views. Follow the directions to share your opinion. You can also select the notepad to view your notes or click the links on the left to view any of the sources you read.

Choose the perspective (recreational, environmental, economic, or historical) that best relates to your own interests in the CAP Project and summarize briefly what you think about the park proposal from that perspective. Support your thinking using evidence from the text.
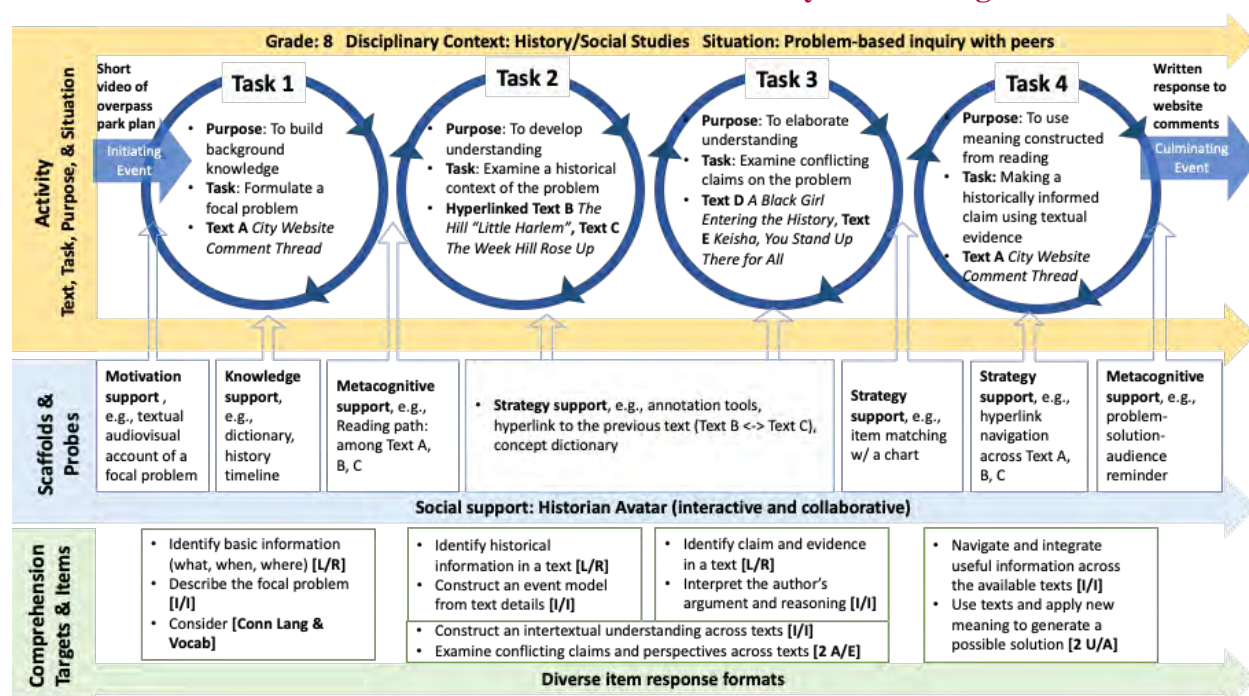
Sources

TRIB LIVE

SEA Website

Righting a Wrong Website

Next City Website

Exhibit 36. This alternative format for the final Use and Apply item with open-constructed response format illustrates the use of motivational UDEs for two purposes: a task character who invites students' own opinion paired with an opportunity to choose their preferred format (text or audio) for expressing their opinion.

Performance Evidence and Indicators

      . *Scores from the Hill District block reveals what Grade 812 students can do when Reading to Solve a Problem in a social studies context. Ultimately, NAEP produces descriptions of what 8th12th graders (or sub-groups of 8th12th graders) can do in each disciplinary reading context. Thus, from students' participation in the Hill District block (and other assessment blocks designated as Reading to Solve a Problem in social studies contexts), it is possible to characterize how well eighth gradeGrade 12 students are able to comprehend and use multiple sources while engaging in social-studies inquiries involving a collection of relatively short but nonetheless complex multilayered digital texts and a range of digitally enhanced items and access tools.*

Grade: 8   Disciplinary Context: History/Social Studies   Situation: Problem-based inquiry with peers

Activity — Text, Task, Purpose, & Situation

Short video of overpass park plan — Initiating Event

Task 1
- **Purpose:** To build background knowledge
- **Task:** Formulate a focal problem
- **Text A** *City Website Comment Thread*

Task 2
- **Purpose:** To develop understanding
- **Task:** Examine a historical context of the problem
- **Hyperlinked Text B** *The Hill "Little Harlem"*, **Text C** *The Week Hill Rose Up*

Task 3
- **Purpose:** To elaborate meaning constructed from reading
- **Task:** Examine conflicting claims on the problem
- **Text D** *A Black Girl Entering the History*, **Text E** *Keisha, You Stand Up There for All*

Task 4
- **Purpose:** To use meaning constructed from reading
- **Task:** Making a historically informed claim using textual evidence
- **Text A** *City Website Comment Thread*

Written response to website comments — Culminating Event

Scaffolds & Probes

- **Motivation support**, e.g., textual audiovisual account of a focal problem
- **Knowledge support**, e.g., dictionary, history timeline
- **Metacognitive support**, e.g., Reading path: among Text A, B, C
- **Strategy support**, e.g., annotation tools, hyperlink to the previous text (Text B <-> Text C), concept dictionary
- **Strategy support**, e.g., item matching w/ a chart
- **Strategy support**, e.g., hyperlink navigation across Text A, B, C
- **Metacognitive support**, e.g., problem-solution-audience reminder

Social support: Historian Avatar (interactive and collaborative)

Comprehension Targets & Items

- Identify basic information (what, when, where) [L/R]
- Describe the focal problem [I/I]
- Consider [Conn Lang & Vocab]

- Identify historical information in a text [L/R]
- Construct an event model from text details [I/I]
- Identify claim and evidence in a text [L/R]
- Interpret the author's argument and reasoning [I/I]
- Navigate and integrate useful information across the available texts [I/I]
- Use texts and apply new meaning to generate a possible solution [2 U/A]

- Construct an intertextual understanding across texts [I/I]
- Examine conflicting claims and perspectives across texts [2 A/E]

Diverse item response formats

E.B. White

The last example offers a sketch of what a Grade 8 Reading to Develop Understanding in a Literature Context block might look like. This example illustrates what a block might look like if it occupied a space along the left end of the continuum portrayed in Exhibit 2. Here, students have more time to develop deep understanding of the texts. Tasks are relatively simple, and so fewer digital design features are needed to support the complexity of the task. When fully developed, this block should provide a good opportunity for students to demonstrate reading to develop understanding, by answering text-based questions that promote close reading of two texts as well as drawing inferences about how the ideas in the two texts inform one another.

Block Components (Disciplinary Context, Purposes, and Reader Role). In this example, students read and answer questions about two texts representing common literary genres: (a) a *biographical sketch* about the author E. B. White, and (b) a short human-interest *essay* by him. Some of the items will query the sketch, others will query the essay, and one item will require reasoning across the texts. These texts are a part of a NAEP released block that was used in the 2011 NAEP Assessment. The texts appear here (in Exhibits 46 and 47), as they did in that assessment.

At the outset, readers are provided a specific reading purpose and informed about the role (working on their own) they will be asked to assume during the block, composed of two common literary genres—a biographical sketch and a human-interest essay (see Exhibit 37).

Exhibit 37. Introduction to EB White

Introduction

You will read two texts: (1) a biographical sketch about the author E. B. White, most famous for writing *Charlotte's Web*, and (2) an essay that White wrote for The New Yorker magazine.

You will answer questions about each text. Then, you will explain how the description of E. B. White in DiConsiglio's biographical sketch applies or does not apply to the narrator of E.B. White's essay, *Twins*.

NEXT

Task Components: Tasks, Text(s), and Items). This E. B. White block has three tasks that include, 1) Reading and answering questions about the biographical sketch, *Not Just for Kids Anymore;* 2) Reading and answering question about the essay, *Twins*, and 3) Reasoning across the two texts to explain how what was learned in *Not Just for Kids Anymore* helps to understand E.B. White, the narrator of the essay, *Twins*. See exhibits See Exhibit 38, which shows task 1.

Exhibit 38. Introduction to the grade 8 EB White literature block

Task 1: Read the biographical sketch to learn about how DiConsiglio viewed E. B. White's career. Then answer the 4 questions about this text.

Meet the author: E. B. White, the author of children's classics Charlotte's Web and Stuart Little, was also a great essayist.

Not Just for Kids Anymore

By John DiConsiglio

"I have a lot of the cat in me," said author E. B. White, "and cats are not joiners."

Perhaps that is why White, one of the country's greatest writers, is so hard to label. His essays for *The New Yorker* appealed to an urbane crowd, but he is best remembered for his children's books. He loved the bustle of New York City, but was happy raising chickens on a Maine farm. And just when critics thought they had him pegged as a benign satirist, he'd write a biting condemnation of the dangers of technology.

Select the answer that shows why DiConsiglio thought it was hard to label E. B. White.

A ◯ White was as happy in a crowd in New York City as on a farm in New England.

B ◯ White was a great writer.

C ◯ White was well-liked by many people.

D ◯ White could write more than one type of prose.

NEXT
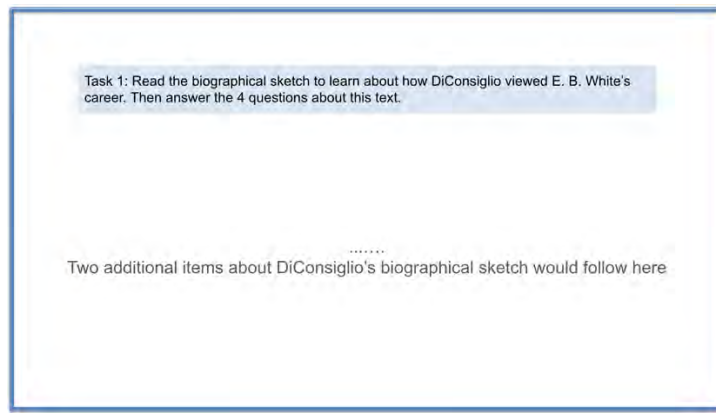
The comprehension items for Task 1 could help the reader develop understanding on

154

segments of the biographical sketch that focus on characteristics of White that might be useful in Task 3 (see Exhibit 39). Plausible segments for focus could be…

- The very first paragraph in which he compares himself to a cat.
- His adaptability (equally comfortable in NYC or Maine).
- Mood variation—benign satire to biting critique.
- The statement near the end suggesting that his essays matched his personality.
- The very last statement, suggesting that he was an eminently likeable character.

In terms of UDEs, note that there is a knowledge-based introductory UDE just before the title of the biographical sketch. Several relatively obscure terms are singled out as possible vocabulary pop-ups for a definition. No explicit motivational UDEs are provided.

Exhibit 39. Task 1 would involve additional items

Task 1: Read the biographical sketch to learn about how DiConsiglio viewed E. B. White's career. Then answer the 4 questions about this text.

....
Two additional items about DiConsiglio's biographical sketch would follow here

For Task 2, comprehension items should focus on the narrator White's statements that say something about his personality and attitudes toward the world around him (see Exhibits 40-42). Candidates for items include:

- Getting more than we bargained for and the sighting of the cow and her twins.
- White's characterization of the moose cow resentful of the onlookers
- The description of the mother and child as unaware of the special treat before their eyes
- The fawn's attempt to "hide" behind the leaf of the plant.
- One of several contrasts between the natural environment in a forest and the urban substitute of a zoo.

In terms of UDEs, similar to the biographical sketch there is a knowledge-based introductory UDE just before the title of the biographical sketch. Also several relatively obscure terms are singled out as possible vocabulary pop-ups for a definition. No explicitly motivational UDEs are provided.

Exhibit 40. Task 2 for the grade 8 EB White block illustrating an Integrate and Interpret item with a short constructed response item format

Task 2: Read E.B. White's essay in light of DiConsiglio's biographical sketch, making connections as you read. Then, answer the questions.

E. B. White was not only a great author for children, he was also the preeminent essayist of his time. This essay, written as a "Talk of the Town" piece for The New Yorker, provides a hint of his powers..

Twins

On a warm, miserable morning last week we went up to the Bronx Zoo to see the moose calf and to break in a new pair of black shoes. We encountered better luck than we had bargained for.

The cow moose and her young one were standing near the wall of the deer park below the monkey house, and in order to get a better view we strolled down to the lower end of the park, by the brook. The path there is not much traveled. As we approached the corner where the brook trickles under the wire fence, we noticed a red deer getting to her feet. Beside her, on legs that were just learning their business, was a spotted fawn, as small and perfect as a trinket seen through a reducing glass. They stood there, mother and child, under a gray beech whose trunk was engraved with dozens of hearts and initials. Stretched on the ground was another fawn, and we realized that the doe had just finished twinning. The second fawn was still wet, still unrisen. Here was a scene of rare sylvan splendor, in one of our five favorite boroughs, and we couldn't have asked for more. Even our new shoes seemed to be working out all right and weren't hurting much.

White begins the essay by suggesting that "we encountered better luck than we had bargained for." What was the better luck?

Use the blank box below to type your response. Use evidence from the text to explain your thinking.

NEXT

Exhibit 41. Task 2 continues for the grade 8 EB White block illustrating an Analyze and Evaluate item with a multiple choice item response format

Task 2 (continued).

The doe was only a couple of feet from the wire, and we sat down on a rock at the edge of the footpath to see what sort of start young fawns get in the deep fastnesses of Mittel Bronx. The mother, mildly resentful of our presence and dazed from her labor, raised one forefoot and stamped primly. Then she lowered her head, picked up the afterbirth, and began dutifully to eat it, allowing it to swing crazily from her mouth, as though it were a bunch of withered beet greens. From the monkey house came the loud, insane hooting of some captious primate, filling the whole woodland with a wild hooroar. As we watched, the sun broke weakly through, brightened the rich red of the fawns, and kindled their white spots. Occasionally, a sightseer would appear and wander aimlessly by, but of all who passed none was aware that anything extraordinary had occurred. "Looka the kangaroos!" a child cried. And he and his mother stared sullenly at the deer and then walked on.

As the birth of the twins happened for all to see, how does the narrator contrast what he and his partner saw with what others saw?

A ○ White and his partner saw something special but others saw only the ordinary. ⊖

B ○ White and his partner saw a moose but others saw a kangaroo. ⊖

C ○ White and his partner saw a monkey but others saw a the mother moose. ⊖

D ○ They all saw the deer and walked on. ⊖

NEXT

Exhibit 42. Additional items accompany task 2

Task 2 (continued).

....

Three additional items about White's essay, *Twins*, would follow here

For Task 3, which was foreshadowed by the original block-specific purpose at the outset, both texts are involved. A task-based UDE, in the form of a partially completed note-taking chart (see Exhibits 43 and 44), might be provided to assist students in organizing their response to a final Use and Apply extended constructed response item (see Exhibit 45).

Exhibit 43. An Integrate and Interpret item illustrating a matching item response format

Task 3: Comparing ideas across the two passages

The final question (item 10) will require you to show how the ideas from *Not Just for Kids Anymore* apply to the narrator of the essay, *Twins*.
To prepare for that final item, fill out the chart below by moving phrases from the idea box into the the blank spaces in the chart.

1. Idea from *Not Just for Kids Anymore*	2. How this applies to the narrator of *Twins*
Cats are not joiners.	White and his companion hung back from the others who could see the moose.
He could adapt to many settings.	
	He was critical of the mother and child, who seemed not to appreciate the incredible good fortune of witnessing the twin birth.
He was comfortable on a rural farm with animals.	

Idea Box
- When at the zoo, the narrator was able to sit back and enjoy the birth of the twins.
- He showed great respect for the animals at the zoo.
- He is capable of biting criticism.
- He graduated from Cornell University.

NEXT

Exhibit 44. Integrate and Interpret item illustrating resetting of item responses from prior item

Task 3: Completed Chart: Comparing ideas across the two passages

No question to answer on this screen. Below you will see the chart from the previous page with the phrases from the idea box when they have been dragged into the right spaces in the chart. You can refer back to this chart when you complete the next (and last) item on this block.

1. Idea from *No Longer Just for Kids*	2. How this applies to the narrator of *Twins*
Cats are not joiners.	White and his companion hung back from the others who could see the moose.
He could adapt to many settings.	When at the zoo, he was able to sit back and enjoy the birth of the twins.
He was capable of biting criticism.	He was critical of the mother and child, who seemed not to appreciate the incredible good fortune of witnessing the twin birth.
He was comfortable on a rural farm with animals.	He showed great respect for the animals at the zoo.

Idea Box
- When at the zoo, the narrator was able to sit back and enjoy the birth of the twins.
- He showed great respect for the animals at the zoo.
- He was capable of biting criticism.
- He graduated from Cornell University.

NEXT

After completing the drag and drop task with the chart (Exhibit 43), students receive feedback about how the chart might best have been completed in Exhibit 44. The KB UDE, called resetting, is provided so that students do not carry misconceptions into the final item in Exhibit 45.

Task 3: Compare ideas across the two passages

For the final task, you will use ideas from the biographical sketch to support your thoughts about how ideas from DiConsiglio's biographical sketch apply to the narrator of *Twins*.

Use the underlined completed chart on the previous page or go back to either passage to get ideas to support your answer. Type your answer into the box below.

DONE

As suggested earlier, the EB White block sketch provides an example of what blocks might look in under the auspices of the 2026 when they are developed with an RDU Broad Purpose as the driving force in design. Blocks like these have long been a part of the NAEP Reading Assessment Portfolio and will continue to be included going forward. For the convenience of the reader, the full version of the two texts used for this block appear in Exhibits 46 and 47.

Exhibit 46. The First Text for the E. B. White Task: A Biographical Sketch. Meet the author: E. B. White, the author of children's classics Charlotte's Web and Stuart Little, was also a great essayist.

Not Just for Kids Anymore

"I have a lot of the cat in me," said author E. B. White, "and cats are not joiners."

Perhaps that is why White, one of the country's greatest writers, is so hard to label. His essays fo*r The New Yorker* appealed to an urbane crowd, but he is best remembered for his children's books. He loved the bustle of New York City, but was happy raising chickens on a Maine farm. And just when critics thought they had him pegged as a benign satirist, he'd write a biting condemnation of the

159

dangers of technology.

E. B. White and Minnie, his dachshund, at *The New Yorker* offices in the late 1940s.

The son of a piano manufacturer, Elwyn Brooks White was born in Mount Vernon, New York, in 1899. His family was prosperous, and White was raised with the mix of sophistication and common sense that would mark his writing. After graduation from Cornell University, White spent a year as a newspaper reporter in New York City, then decided to drive across the country with a friend in a Model T Ford. The trip gave White a lifetime of anecdotes, and spawned a legend or two. "When they ran out of money," White's friend, James Thurber, noted, "they played for their supper—and their gasoline—on a fascinating musical instrument that White had made out of some pieces of wire and an old shoe."

When White returned to New York City in the mid-1920s, he spent a few years bouncing between advertising jobs and unemployment before trying his hand again at writing Borrowing his brother's typewriter, he began pounding out sketches and poems. On a lark, he sent some essays to a fledgling magazine called *The New Yorker.* Since its founding in 1925, the magazine had struggled to find its niche, and White's work helped put *The New Yorker* on the map. His essays were funny and sophisticated; they spoke equally to socialites and cab drivers, professors and plumbers. Through his essays, which he wrote for nearly

50 years, White helped give *The New Yorker* its voice and identity.

In 1945, already a leading literary figure, White embarked on his second career: writing children's books. He moved from New York to a farm in Maine, where he raised chickens and geese. Seeking a way to amuse his nieces and nephews, White started to write stories for them. "Children were always after me to tell them a story and I found I couldn't do it," he said. "So I had to get it down on paper."

A vivid dream about a mouselike character led to S*tuart Little*. Then, in 1952, White published *Charlotte's Web*. The book, which was inspired by White's own farm animals, is arguably the most famous children's story published in the 20th century.

By the time he died from Alzheimer's disease in 1985, White's essays had appeared in more college anthologies than those of any other writer. Many said his essays matched his personality: subtle without being simple, critical without being mean.

Indeed, one New *York Time*s critic wrote, "There are times reading an E. B. White book of essays when you think he must be the most likable man of letters alive. If you are some kind of writer yourself, you probably want to imitate him."

-By John DiConsiglio

E. B. White was not only a great author for children, he was also the preeminent essayist of his time. This essay, written as a "Talk of the Town" piece for The New Yorker, provides a hint of his powers.

Twins

by E. B. White

On a warm, miserable morning last week we went up to the Bronx Zoo to see the moose calf and to break in a new pair of black shoes. We encountered better luck than we had bargained for.

The cow moose and her young one were standing near the wall of the deer park below the monkey house, and in order to get a better view we strolled down to the lower end of the park, by the brook. The path there is not much traveled. As we approached the corner where the brook trickles under the wire fence, we noticed a red deer getting to her feet. Beside her, on legs that were just learning their business, was a spotted fawn, as small and perfect as a trinket seen through a reducing glass. They stood there, mother and child, under a gray beech whose trunk was engraved with dozens of hearts and initials. Stretched on the ground was another fawn, and we realized that the doe had just finished twinning. The second fawn was still wet, still unrisen. Here was a scene of rare sylvan splendor, in one of our five favorite boroughs, and we couldn't have asked for more. Even our new shoes seemed to be working out all right and weren't hurting much.

The doe was only a couple of feet from the wire, and we sat down on a rock at the edge of the footpath to see what sort of start young fawns get in the deep fastnesses of Mittel Bronx.

The mother, mildly resentful of our presence and dazed from her labor, raised

one forefoot and stamped primly. Then she lowered her head, picked up the afterbirth, and began dutifully to eat it, allowing it to swing crazily from her mouth, as though it were a bunch of withered beet greens. From the monkey house came the loud, insane hooting of some captious primate, filling the whole woodland with a wild hooroar. As we watched, the sun broke weakly through, brightened the rich red of the fawns, and kindled their white spots. Occasionally, a sightseer would appear and wander aimlessly by, but of all who passed none was aware that anything extraordinary had occurred. "Looka the kangaroos!" a child cried. And he and his mother stared sullenly at the deer and then walked on.

In a few moments the second twin gathered all his legs and all his ingenuity and arose, to stand for the first time sniffing the mysteries of a park for captive deer. The doe, in recognition of his achievement, quit her other work and began to dry him, running her tongue against the grain and paying particular attention to the key points. Meanwhile the first fawn tiptoed toward the shallow brook, in little stops and goes, and started across. He paused midstream to make a slight contribution, as a child does in bathing. Then, while his mother watched, he continued across, gained the other side, selected a hiding place, and lay down under a skunk-cabbage leaf next to the fence, in perfect concealment, his legs folded neatly under him. Without actually going out of sight, he had managed to disappear completely in the shifting light and shade. From somewhere a long way off a twelve-o'clock whistle sounded. We hung around awhile, but he never budged. Before we left, we crossed the brook ourself, just outside the fence, knelt, reached through the wire, and tested the truth of what we had once heard: that you can scratch a new fawn between the ears without starting him. You can indeed.

Footnote

Sample items in the framework are being provided to exemplify key concepts in the framework and do not represent items that will be used on future NAEP assessments. These sample items may not represent accurately the full set of NAEP style guide and other test specifications. Tasks presented with multiple sample items are provided to help readers of the framework envision how theoretical ideas in the framework might guide assessment design, but they do not represent fully expectations for enacting the NAEP style guide and other test specifications.

Exhibit 1. Example of a Matching Selected Response Item for a Webpage Text from PISA's Rapa Nui Block

PISA

Rapa Nui
Question 6 / 7

Refer to all three sources on the right by clicking on each of the tabs.

Drag and drop the causes, and the effect they have in common, into the correct places in the table about the theories.

The Theories

Cause	Effect	Supporters of the Theory
		Jared Diamond
		Carl Lipo and Terry Hunt

The moai were carved in the same quarry.	Polynesian rats ate tree seeds and as a result no new trees could grow.	Settlers used canoes to bring Polynesian rats to Rapa Nui.
The large trees disappeared from Rapa Nui.	Rapa Nui residents needed natural resources to move the moai.	Humans cut down trees to clear land for agriculture and other reasons.

Blog   Book Review   **Science News**

← → ⟳   www.sciencenews.com/Polynesian_rats_Rapa_Nui

SCIENCE NEWS

Did Polynesian Rats Destroy Rapa Nui's Trees?

By Michael Kimball, Science Reporter

In 2005, Jared Diamond published *Collapse*. In the book, he described the human settlement of Rapa Nui (also called Easter Island).

The book caused a huge controversy soon after its publication. Many scientists questioned Diamond's theory of what happened on Rapa Nui. They agreed that the huge trees had disappeared by the time Europeans first arrived on the island in the 18th century, but they did not agree with Jared Diamond's theory about the cause of the disappearance.

Now, two scientists, Carl Lipo and Terry Hunt, have published a new theory. They believe that the Polynesian rat ate the seeds of the trees, preventing new ones from growing. The rat, they believe, was brought over either accidentally or purposefully on the canoes that the first human settlers used to land on Rapa Nui.
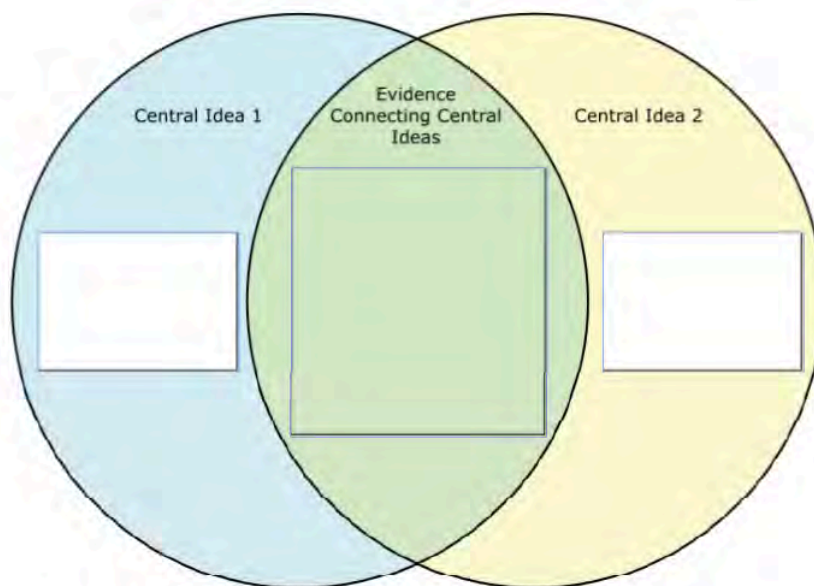
Studies have shown that a population of rats can double every 47 days. That's a lot of rats to feed. To support their theory, Lipo and Hunt point to the remains of palm nuts that show the gnaw marks made by rats. Of course, they acknowledge that humans did play a role in the destruction of the forests of Rapa Nui. But they believe that the Polynesian rat was an even greater culprit among a series of factors.

Exhibit 1, from PISA's Reading Literacy test for 15-year-olds, illustrates a matching item response format. After reading a webpage, students are asked to "drag and drop" the causes and effects offered at the bottom of the table into the appropriate places in the table.

Choose **two** central ideas that are developed in the passage from "Biopiracy in India: The Case of the Aubergine." Drag each idea into one of the sections of the Venn diagram labeled Central Idea. Then, drag the quotation that illustrates the relationship between the two central ideas to the central section of the Venn diagram.

The government of India may be too disorganized to stand up to corporations like Monsanto.	Genetically-modified foods are currently controlled by a few large corporations.	Farmers in India have not been paid for the seeds and crops they have cultivated for generations.	Monsanto has not lived up to the pledges it has made in public statements.	Developments in agricultural technology should be used to help feed the people who are now going hungry.
"In response, the national biodiversity authority has announced its plans to prosecute Monsanto for carrying out this research without seeking its permission and the consent of hundreds of thousands of farmers who have cultivated these varieties for generations." (paragraph 5)	"More than half (53 per cent) of all genetically modified and organic seeds traded worldwide are owned by three multinational companies, according to the environmental group Greenpeace." (paragraph 12) "This is all the more poignant in India, where thousands of debt-ridden farmers have in recent years resorted to taking their own lives to escape the misery of crop failure and financial ruin." (paragraph 12)	"The world's top ten agro-chemical companies own almost 75 per cent of all seeds globally." (paragraph 12)		"In developing nations where farmers often rely on subsistence agriculture to eke out meager livings, the controversial and highly lucrative industry of genetic engineering is thrown into sharper relief against a backdrop of widespread poverty." (paragraph 12)

Example 2, from a PARCC Grade 12 task, illustrates a matching format. Students are asked to "drag" the ideas into the venn diagram.

Exhibit 3. Example of a Zones Selected Response Item Format and the Use of Task Characters from ePIRLS' Mars Block

Exhibit 3, from an ePIRLS task for grade 4 students, illustrates a zones item format. The item asks students to "click on the website tab 'Rover Called Curiosity'." To do so, students must click on the tab of the webpage with the same title. This item also illustrates the use of task characters, or avatars. An animated icon of a teacher shows "Mr. Webster," and another one shows the "Student," who is the test taker.

Exhibit 4. Example of a Grid Selected Response Item from PISA's Rapa Nui Block

PISA

Blog **Book Review** **Science News**

www.academicbookreview.com/Collapse

Rapa Nui
Question 3 / 7

Refer to the Review of Collapse on the right. Click on the choices in the table to answer the question.

Listed below are statements from the Review of *Collapse*. Are these statements facts or opinions? Click on either **Fact** or **Opinion** for each statement.

Is the statement a fact or an opinion?	Fact	Opinion
In the book, the author describes several civilizations that collapsed because of the choices they made and their impact on the environment.	○	○
One of the most disturbing examples in the book is Rapa Nui.	○	○
They carved the moai, the famous statues, and used the natural resources available to them to move these huge moai to different locations around the island.	○	○
When the first Europeans landed on Easter Island in 1722, the moai were still there, but the trees were gone.	○	○
The book is written well and deserves to be read by anyone who is concerned about the environment.	○	○

Review of *Collapse*

Jared Diamond's new book, *Collapse*, is a clear warning about the consequences of damaging our environment. In the book, the author describes several civilizations that collapsed because of the choices they made and their impact on the environment. One of the most disturbing examples in the book is Rapa Nui.

According to the author, Rapa Nui was settled by Polynesians sometime after 700 CE. They developed a thriving society of, perhaps, 15 000 people. They carved the moai, the famous statues, and used the natural resources available to them to move these huge moai to different locations around the island. When the first Europeans landed on Rapa Nui in 1722, the moai were still there, but the trees were gone. The population was down to a few thousand people who were struggling to survive. Mr. Diamond writes that the people of Rapa Nui cleared the land for farming and other purposes and that they over-hunted the numerous species of sea and land birds that had lived on the island. He speculates that the dwindling natural resources led to civil wars and the collapse of Rapa Nui's society.

The lesson of this wonderful but frightening book is that in the past, humans made the choice to destroy their environment by cutting down all the trees and hunting animal species to extinction. Optimistically, the author points out, we can choose **not** to make the same mistakes today. The book is written well and deserves to be read by anyone who is concerned about the environment.

Exhibit 4, from PISA's Reading Literacy test for 15-year-olds, illustrates the use of a grid item response format to efficiently collect data about students' ability to analyze multiple fact/opinion statements.

Exhibit 5. Example of a Zones Item for an Internet Text from ePIRLS' "Elizabeth Blackwell" Block

ePIRLS Online Reading 2016

http://www.google.com?q=doctor+elizabeth+blackwell

TIME LEFT 39:27

PROGRESS

Google

Google   Doctor Elizabeth Blackwell

Elizabeth J. Blackwell — Film Archive
www.filmarchive.com/eblackwell_(film)
Elizabeth Blackwell, (born Nov 2, 2002) is a child actress who is famous for her appearance in the series of movies, "Magic Mischief"...

Doctor Elizabeth Blackwell — Her Story
www.womenshistory.org/dreblackwell
Doctor Elizabeth Blackwell, (3 February 1821 – 31 May 1910) is recognized as the first woman to become a **doctor** in the United States of America...

Elizabeth Blackwell Medal
www.elizabethblackwellmedal.org
The **Elizabeth Blackwell** Medal is an award given to a **doctor** who shows exceptional...

Doctor Blackwell visits the jungle – Blossom Books
www.blossonmbooks.com/blackwell
Doctor Blackwell visits the jungle is a children's picture book by Sarah Schubert. Published in 2010...

ePIRLS
Class Project

Mr. Webster
Today, we're going to read about the first woman doctor in the United States and England— Doctor Elizabeth Blackwell.

Mr. Webster
Let's begin by using "Google" to search the Internet.

1.
Look at the Google search results, at left.

Student

Click on the link that is most likely to have information about the life and achievements of Doctor Elizabeth Blackwell.

Exhibit 5, from ePIRLS' assessment for grade 4 students, provides an example of the use of a zones item format. Here, students are asked to "Click on the link that is most likely" to have the requested information – in this case, "information about the life and achievements of Doctor Elizabeth Blackwell." This exhibit also illustrates the use of an Internet text in the form of a search engine results page.

Exhibit 6. Example of an In-line Choice Item from ePIRLS' Mars Block That Also Collects Process Data on Where Students Click on the Web Page

Exhibit 6, from ePIRLS' assessment for grade 4 students, asks students to use the digital diagram to answer questions by selecting responses from a drop-down menu (an in-line choice item). This item also collects **process data** of where on the graphic stimulus students click. While the clicks are not scored as items, they allow test makers to collect valuable information about why students might perform the way that they do. Such information can be useful for test development and also for outside researchers.

Exhibit 7. Example of a Short Constructed Response Item from PISA's Galapagos Islands Block

Exhibit 7, from PISA's Reading Literacy test for 15-year-olds, illustrates a short constructed response. Here, students are given a small text box and asked to write about a key difference they read about in the approach taken to two different conservation programs.

Exhibit 8. Example of a Fill in the Blank Item Response Format from ePIRLS' Mars Block

Exhibit 8, from ePIRLS for grade 4 students, illustrates the use of a fill in the blank item response format for a digital website text that is a graphic. Here, students are asked to use the graphic to identify the "names of the three planets between Mars and the Sun." To give their answers, students type each name ("Mercury," "Venus," and "Earth") into three separate text fields.

Exhibit 1. Example of a Specific Reading Purpose and a Knowledge-based UDE from PISA's Rapa Nui Block
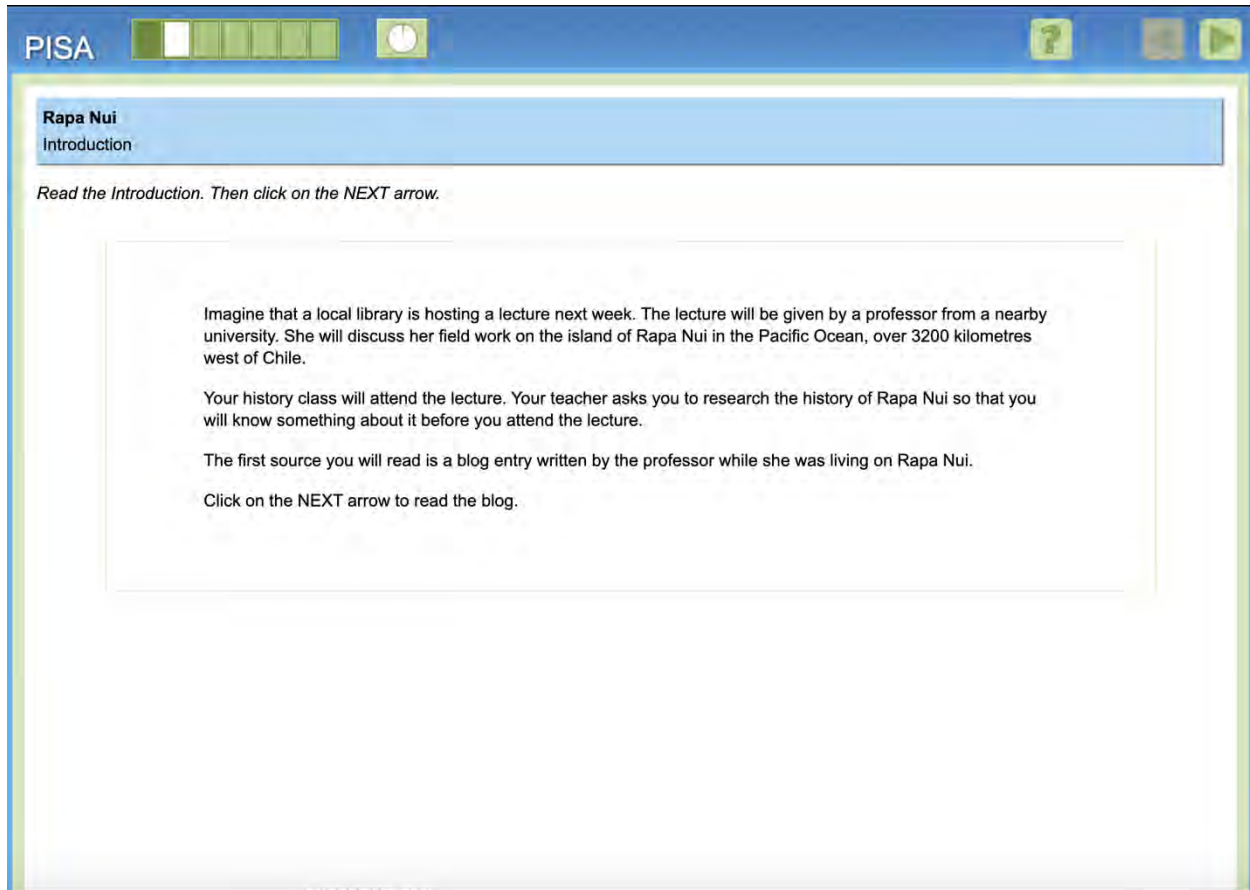
Exhibit 1, from PISA's Reading Literacy test for 15-year-olds, illustrates how readers are situated, at the beginning of the block, within a specific reading purpose: To conduct research on the history of Rapa Nui in order to prepare for a lecture at a local library. This example also illustrates a knowledge-based UDE in which students are introduced to the first source they will read – a blog entry written by a professor while living in Rapa Nui.

Exhibit 2. Example of a Task-based UDE from the Smarter Balanced Items Published by The Regents of the University of California

Student Directions for Part 2

You will now review your sources, take notes, and plan, draft, revise, and edit your article. You may use your notes and refer to the sources. Now read your assignment and the information about how your article will be scored; then begin your work.

Your Assignment:

In your school, the Science Club is encouraging students to provide articles for its new website. For your contribution to the website, you will write an explanatory article about improving memory.

Using more than one source, develop a thesis/controlling idea to explain how to improve memory. Once you have a thesis/controlling idea, select the most relevant information to support your thesis/controlling idea. Then, write a multi-paragraph explanatory article explaining your thesis/controlling idea. Clearly organize your article and elaborate on your own ideas. Develop your ideas clearly and use your own words, except when quoting directly from the sources. Be sure to reference the source title or number when quoting or paraphrasing details or facts from the sources.

Explanatory Scoring

Your explanatory article will be scored using the following:

1. **Organization/Purpose:** How well did you state your thesis/controlling idea and maintain your thesis/controlling idea with a logical progression of ideas from beginning to end? How well did you narrow your thesis/controlling idea so you can develop and elaborate on the conclusion? How well did you consistently use a variety of transitions? How effective were your introduction and your conclusion?

2. **Elaboration/Evidence:** How well did you integrate relevant and specific information from the sources? How effective were your elaborative techniques? How well did you clearly state ideas using precise language that is appropriate for your audience and purpose?

3. **Conventions:** How well did you follow the rules of grammar usage, punctuation, capitalization, and spelling?

Now begin work on your explanatory article. Manage your time carefully so that you can:

- plan your multi-paragraph article,

- write your multi-paragraph article, and

- revise and edit the final draft of your multi-paragraph article.

Word-processing tools and spell-check are available to you.

For Part 2, you are being asked to write a multi-paragraph article, so please be as thorough as possible. Type your response in the space provided. The box will expand as you type.

Remember to check your notes and your prewriting/planning as you write, and then revise and edit your article.

Exhibit 2, from the Smarter Balanced test for grade 8 students, illustrates a task-based UDE in the form of scoring criteria and steps for writing an explanatory article. Additionally, the example illustrates the use of an extended constructed response item in the form of what would be a Use and Apply comprehension target in the 2026 NAEP Reading Assessment. See Appendix E for additional examples of different response formats.

Exhibit 3. Example of a Motivational UDE, from NAEP's "Tough as Daisy" Block

I'm the only girl at
the sign-up desk.

Tough as Daisy

by David M. Simon

The sign on the YMCA door says *Wrestling Tournament Today.*

I enter the gym and take a deep breath. It smells like old sweat socks and the stuff they use to wash wrestling mats.

I love that smell. Weird, huh? Not to me.

My dad always says, "Pound for pound, no one's as tough as Daisy."

I see my family in the stands. I wave to them and smile, but I'm nervous.

Lots of boys are already on the mats, loosening up. I'm the only girl at the sign-up desk. Some of

Exhibit 3, from a NAEP grade 4 block, illustrates a motivational UDE in the form of an illustration and caption. Together, the illustration and caption reading, "I'm the only girl at the sign-up desk." serve to pique readers' interest in the text. The illustration and caption also serve as a knowledge-based UDE because they introduce the text by offering key plot information (a girl standing in line, among only boys).

Exhibit 4. Example of Two Knowledge-based UDEs from NAEP's "Five Boiled Eggs" Block

Exhibit 4, from a NAEP Grade 4 block, illustrates two knowledge-based UDEs. The first knowledge-based UDE appears in the form of an introduction to the story "Five Boiled Eggs," which introduces students to Nasreddin Hodja, a character in the story whose last name means "teacher" in Turkish. The second knowledge-based UDE appears in the form of a vocabulary pop-up box defining the Turkish word "akche."

<u>Exhibit 5. Two Examples of Knowledge-based UDEs in the Form of Passage Introductions from a Released NAEP 2019 Block on E. B. White</u>

<u>Example 1</u>

> *Meet the author:* E. B. White, the author of children's classics Charlotte's Web *and* Stuart Little, *was also a great essayist.*
>
> # Not Just for Kids Anymore
>
> "I have a lot of the cat in me," said author E. B. White, "and cats are not joiners."
>
> Perhaps that is why White, one of the country's greatest writers, is so hard to label. His essays for *The New Yorker* appealed to an urbane crowd, but he is best remembered for his

<u>Example 2</u>

> *E. B. White was not only a great author for children, he was also the preeminent essayist of his time. This essay, written as a "Talk of the Town" piece for* The New Yorker, *provides a hint of his powers.*
>
> # Twins
>
> by E. B. White
>
> On a warm, miserable morning last week we went up to the Bronx Zoo to see the moose calf and to break in a new pair of black shoes. We

<u>Exhibit 5 illustrates two different written introductions, one for each of two texts. In Example 1, a knowledge-based UDE appears in the form of an introduction to an article about the writer E. B. White. In Example 2, a knowledge-based UDE appears in the form of an introduction to an essay by E. B. White, which explains that the author of the essay is also a children's author.</u>

Exhibit 6. Example of Three Knowledge-based UDEs in the Form of Passage Introductions from the Michigan Student Test of Educational Progress

Source #1
You have found an article that describes how animals survive in different environments, the places where plants and animals live.

Source #2
You have found an article from *Appleseeds* magazine that describes how some animals build their homes.

Source #3
You have found an article that discusses plants and animals that live in the same place. The article describes how these plants and animals depend on each other to stay alive.

Exhibit 6, from Michigan's reading assessment for grade 4 students, illustrates three knowledge-based UDEs in the form of passage introductions for each of three different sources within a block. In this task, students are asked to learn from reading each source and to then write an informational article using what they have learned.

(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

AERA/APA/NCME. (2014). Standards for educational and psychological testing. Washington, DC: Author.

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

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

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

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. (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.

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.

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

Bowen, N., Bowen, G. & Ware, W. (2002). Neighborhood Social Disorganization, Families, and the Educational Behavior of Adolescents. *Journal of Adolescent Research*, 17 (5), 468-490.

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

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 793-828). John Wiley & Sons.

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.

Damasio, 1995

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.

Durán, 2006

Durlak, J., Domitrovich, C., Weissberg, R., & Gullotta, T. (Eds.). (2015). *Handbook of social and emotional learning: Research and practice*. Guilford Press.

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. O'Neill, et al. (2005). Ability self-perceptions and subjective task values in adolescents and children.

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

Elias, M. (2019). What if the doors of every schoolhouse opened to social-emotional learning tomorrow: Reflections on how to feasibly scale up high-quality SEL. *Educational Psychologist, 54*(3), 233–245.

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*, 34, 4, 293–309

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., Hamm, J., Dawes, M., Barko-Alva, K., & Cross, J. (2019). Promoting inclusive communities in diverse classrooms: Teacher attunement and social dynamics management. *Educational Psychologist, 54*(4), 286–305.

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/10.1177/0022057416196000303

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.

Goldman S. (2012). Adolescent literacy: learning and understanding content. *The Future of Children / Center For the Future of Children, the David and Lucile Packard Foundation*. 22: 89-116. PMID 23057133

Goldman, S. R. (2018). Discourse of learning and the learning of discourse. Discourse Processes,55(5–6), 434–453.

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.

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

~~Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371-395.~~

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. L. (2016). Engagement and motivational processes in reading. In Afflerbach, P. (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 41–53). Routledge.

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., & You, W. (2012). Instructional contexts for engagement and achievement in reading. In S. Christensen, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 601-635). New York: Springer Science

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.

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.

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.

Hopkins, Thompson, Linquanti, August, & Hakuta, 2013

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

Jaeger 2003

Jaeger, 1998

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

Johnstone, C., Altman, J., & Thurlow, M. (2006). A state guide to the development of universally designed assessments. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Katz, M.L., Brynelson, N., and Edlund, J.R. (2019). "Enacting Rhetorical Literacies: The Expository Reading and Writing Curriculum in Theory and Practice." seventh edition, *Theoretical Models and Processes of Literacy*, D. Alvermann, N. Unrau, and M. Sailors (Eds.). New York: Routledge (pp. 533-562).

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*. London, UK: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, 2019

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. (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., Enciso, P. & Lesau, N. NY, Routledge.

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

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.

Linn and Dunbar 1992

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.

Mahoney, J.L., Durlak, J.A., & Weissberg, R.P. (2018). An update on social and emotional learning outcome research. *Phi Delta Kappan, 100*(4), 18–23.

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

Markus and Kitayama 1991

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

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

NAEP Nation's Report Card, 2019

NAEP Reading Special Study, 2019

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, 2014

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. (2019). Reading Framework of the 2019 National Assessment of Educational Progress. Washington, DC: Author

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 2017?

National Research Council, 2018?

National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits.* Washington, DC: The National Academies Press.

National Research Council. (2010). *Language Diversity, School Learning, and Closing Achievement Gaps: A Workshop Summary.* 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., Wang, J., Sabatini, J. (2019). How Much Knowledge Is Too Little? When a Lack of Knowledge Becomes a Barrier to Comprehension Psychological. *Science*, Vol. 30(9) 1344–1351

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

Pearson, P. D., Hansen, J., & Gordon, C. (1979). The effect of background knowledge on young children'schildren's comprehension of explicit and implicit information. Journal of Reading Behavior, 11(3), 201-209210.

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.

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.

Pitzer, J & Skinner, E. (2017). Predictors of changes in students' motivational resilience over the school year: The roles of teacher support, self-appraisals, and emotional reactivity. *International Journal of Behavioral Development*, 41(1) 15–29.

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., 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.

Schaffner, E., Schiefele, U., & Ulferts, H. (2013). Reading Amount as a Mediator of the Effects of Intrinsic and Extrinsic Reading Motivation on Reading Comprehension. *Reading Research Quarterly*, 48, 369-385.

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. and Cole, M (1981). *The Psychology of Literacy*. Cambridge, Mass.: Harvard University Press.

Scribner, S., & Cole, M. (1978). Literacy without schooling: Testing for intellectual effects. *Harvard Educational Review, 48*(4), 448–461.

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

Slemmons, K., Anyanwu, K., Hames, J., Grabski, D., Mlsna, J., Simkins, E., & Cook, P. (2018). The impact of video length on learning in a middle-level flipped science setting: implications for diversity inclusion. Journal of Science Education and Technology, 27(5), 469-479.

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. (~~2010). Assessing the cultural validity of assessment practices: An introduction. Basterra, In M. R., Trumbull, E., & Solano-Flores, G. (Eds.). *Cultural validity in assessment: A guide for educators* (pp. 3-21). New York: Routledge.~~

~~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.

~~Taft, M. L., & Leslie, L. (1985). The effects of prior knowledge and oral reading accuracy on miscues and comprehension. *Journal of Reading Behavior, 17*(2), 163-179.~~

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/

~~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, University of Minnesota. Retrieved from https://nceo.umn.edu/docs/OnlinePubs/Synth44.pdf

Thompson, S. J., Thurlow, M., & Malouf, D. (2004). Creating better tests for everyone through universally designed assessments. *Journal of Applied Testing Technology, 6*, 1–15.

Tomasello, M. (2016). The ontogeny of cultural learning. *Current Opinion in Psychology*, *8*, 1–4, doi:10.1016/j.copsyc.2015.09.008

Torppa, Eklund, Sulkunen, Niemi & Ahonen, 2018

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 den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A "landscape" view of reading. Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton, & A. C. Greasser (Eds.), *Models of understanding text* (pp. 165–187). Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

Weiner 1985

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., & Wentzel, K. R. (2007). Introduction to motivation at school: Interventions that work. *Educational Psychologist, 42*(4), 191–196.

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

Zelazo, P. D. (2013). Developmental psychology: A new synthesis. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology* (Vol. 1: Body and Mind). https://doi.org/10.1093/oxfordhb/9780199958450.013.0001

Overview of the NAEP Reading Assessment and Projections

April 23, 2021

This document has been prepared in response to questions from the National Assessment Governing Board membership and staff regarding the current operational NAEP Reading Assessment in relation to the most recent draft of the 2026 updated framework. This document has three parts:

    I.   **Description of the Current Operational NAEP Reading Assessment in Relation to the Most Recent Draft of the 2026 Updated Reading Framework**
    II.   **Support Features, Relevant Research, and Development Processes in the Current Operational NAEP Reading Assessment**
    III.   **Implementing the Updated Framework and Maintaining Trend**

    I.   DESCRIPTION OF THE CURRENT OPERATIONAL NAEP READING ASSESSMENT IN RELATION TO THE MOST RECENT DRAFT OF THE 2026 UPDATED READING FRAMEWORK

Starting with the 1992 NAEP Reading Framework, a driving principle for the NAEP Reading assessment has been authenticity as a means of establishing face validity. Authenticity in the context of the NAEP reading assessment means that, to the extent possible, the assessment should reflect the reading experiences of students outside of the testing context. For example, the 1992 NAEP reading assessment was one of the first large-scale assessments to use only full-length, naturally occurring texts as its stimulus reading materials. The move to digital assessment under the current framework has allowed the NAEP reading assessment to reflect the digital reading experiences students encounter on a daily basis both inside and outside of school contexts. The draft 2026 updated framework continues to reflect the principle of authenticity.

Definition

The current NAEP Reading Framework lists the following definition: "The NAEP Reading Assessment is guided by a definition of reading that reflects scientific research, draws on multiple sources, and conceptualizes reading as a dynamic cognitive process." This definition

applies to the assessment of reading achievement on NAEP and states that reading is an active and complex process that involves:

- Understanding written text;
- Developing and interpreting meaning; and
- Using meaning as appropriate to type of text, purpose, and situation.

The draft 2026 updated framework maintains the current construct of reading comprehension while expanding the definition to include, "to explicitly acknowledge the sociocognitive processes of reading. Reading comprehension is defined as making meaning with text and four key features are highlighted—contexts, readers, texts, and activities." More specifically, the draft 2026 framework says the following:

"Reading comprehension is making meaning with text, a complex cognitive process shaped by students' social and cultural influences. To comprehend, readers:

- Engage with text in print and multimodal forms;
- Employ personal resources that include foundational reading skills, language, knowledge, and motivations;
- Extract, construct, integrate, critique, and apply meaning in activities across a range of contexts."

Testing Experience

The NAEP reading assessment transitioned from a paper-based assessment (PBA) to a digitally-based assessment (DBA) in 2017. In the most recent DBA in 2019, each student's assessment session began with a tutorial that included student interactions with the tools and interface and concluded with a 3-minute practice session. Following the tutorial and practice session, students worked through two 30-minute cognitive blocks. The second block was followed by a 15-minute survey questionnaire.

Texts

In accordance with the 2019 NAEP Reading Framework, which was first implemented in 2009, there are two broad categories of passages that make up the NAEP reading assessment: literary and informational. Literary texts include fiction, literary non-fiction, and poetry. Informational texts include exposition, argumentation or persuasive texts, and procedural texts.

The draft 2026 framework calls for three types of texts—literature, social studies, and science—and the texts in the 2019 operational pool fall easily into these three categories.

Items

After the passages are reviewed and approved by the Governing Board, items are written to assess three *cognitive* targets under the current framework. The current framework specifies the three cognitive targets as: Locate/Recall, Integrate/Interpret, and Critique/Evaluate.

The draft 2026 updated framework proposes four ***comprehension*** targets: Locate/Recall, Integrate/Interpret, Analyze/Evaluate, and Use and Apply. The addition of Use and Apply addresses the need to assess students' ability to apply the understandings they have gained from interacting with the stimulus materials for a given purpose (e.g., preparing a page of a website or writing a message to the school board).

The current NAEP Reading Framework calls for the following **item types**:

- **Selected response** – This item type encompasses traditional, single-answer, multiple-choice items as well as more complex items that require multiple selections to be answered correctly. NAEP's shift to digitally-based assessment allowed for the introduction of technology-enhanced items, which include matching (drag and drop), grid, and select-in-passage items. Most selected response items are scored dichotomously (correct or incorrect), but more complex selected response items may be scored for partial credit.
- **Constructed response, short and extended** – This item type requires students to generate a written response. Short constructed response items can be answered with a few words or sentences and extended constructed response items may elicit a short paragraph. These items are scored by humans, using a scoring rubric. Short constructed response items are scored with 2- or 3-point rubrics. Extended constructed response items use a 4-point rubric.

Percentages of each item type are specified in the framework for each grade. Typically, NAEP reading blocks include one extended constructed response item, three to five short constructed response items, and three to seven selected response items. The typical NAEP reading block includes a total of 9–11 items.

The draft 2026 updated framework recommends continuing with these item types and provides percentage ranges for selected response, short constructed response, and extended constructed response items. The draft framework also encourages the continued use and exploration of technology enhanced item types.

Reporting

Results of the NAEP reading assessment are reported on a 0–500-point scale. Three scores are reported at each grade level: a composite, or overall reading score, and two sub-scale scores, one for literary texts and one for informational texts. The draft 2026 updated framework maintains the 0–500-point scale and recommends reporting at each grade level: a composite score and three sub-scale scores—reading to engage in literature, reading to engage in science, and reading to engage in social studies contexts.

NAEP Contextual Questionnaire Items

Following the completion of two cognitive blocks, students respond to a 15-minute survey questionnaire. There are two sections to the Contextual Questionnaire: Core and Reading-specific. Core survey items collect information on students' demographic characteristics, opportunities to learn in and out of the classroom, and educational experiences.

Reading-specific survey items focus on reading-related activities and experiences in and out of school. These items are designed to inform interpretations of the results.

In addition to the student questionnaires, teachers and administrators in schools that participate in NAEP also complete their own NAEP Questionnaires.

The draft 2026 updated framework maintains the current approach to the survey questionnaires along with recommendations for changes to the specific items in the reading surveys.

Assembling the NAEP Assessment Via Assessment Blocks

Each NAEP reading assessment is administered in two 30-minute assessment blocks, followed by a 15-minute block of contextual items. Although each student sees only two blocks, there are multiple blocks in each operational assessment as shown in the chart below. Matrix sampling of students and blocks enables NAEP to cover a broad range of content, while also minimizing the burden for students and schools.

Table 1a summarizes the number of NAEP reading assessment blocks administered in the 2019 operational assessment for grades 4, 8, and 12. Typically, each block contains 9–11 items.

Table 1a. 2019 Operational NAEP Reading Assessment Blocks and Item Pool

	Grade 4	Grade 8	Grade 12
Total Number of Blocks	12	15	15
Total Number of Items	118	149	132

Types of Assessment Blocks

Currently, two types of blocks make up the NAEP operational reading assessment: discrete blocks and scenario-based task blocks.

Discrete item (DI) blocks provide general instructions for students to read the passage and provide answers to each assessment item relating to the passages that are presented. All texts and all items are always available for student access and use. The current operational pool of DI blocks is comprised of both transadapted and newly developed blocks as described below.

- **Transadapted blocks** – These blocks are digital renditions of the assessment blocks used in the paper and pencil era of NAEP. These DI blocks make up about two-thirds of the current operational assessment.

- **Newly developed blocks** – These blocks were developed specifically for a digital platform. To take full advantage of the digital format, some of these blocks use print texts and some use texts that are "digitally native" and multi-modal. Some passages

contain embedded hyperlinks and videos. (Note that videos are not used as introductions to texts.) Items addressing video content do so in relation to passage content.

Scenario-based Task (SBT) blocks use both print and digitally native, multi-modal texts. In contrast with DI blocks, students can only access texts and questions sequentially, as the SBTs control the order in which students read texts and items and respond to questions. In this way, students are presented with sources and stimulus materials as needed to respond to items. Videos appear both as part of the texts that students read and as additional content but are not used as introductions to texts. Items addressing video information always do so in relation to the written text.

Table 1b summarizes the number of NAEP reading assessment discrete and scenario-based blocks in the current operational assessment for grades 4, 8, and 12.

Table 1b. 2019 Operational NAEP Reading Assessment Discrete and Scenario-Based Blocks

Block Type	Grade 4	Grade 8	Grade 12
Scenario-based Task Blocks	2	2	2
Discrete Blocks (Transadapted)	7	10	11
Discrete Blocks (Newly developed for DBA)	3	3	2
Total	**12**	**15**	**15**

II.   SUPPORT FEATURES, RELEVANT RESEARCH, AND DEVELOPMENT PROCESSES IN THE CURRENT NAEP OPERATIONAL READING ASSESSMENT

Consistent with the principle of authenticity, the current operational NAEP reading assessment uses *support features*, referred to as Universal Design Elements (UDEs) in the draft framework, that are intended to replicate the types of supports provided during reading instruction and practice in school and at home. One central principle is worth emphasizing: *all support features used in a particular block are available to all students who take that block.*

The types of support features available on the 2019 NAEP reading operational assessment include:

- Look-back buttons
- Pop-up notes
- Passage introductions
- Eliminate answer choice
- Highlighting and notetaking
- Text-to-speech on directions
- Zoom & selection of color themes
- Multi-part response frames

- Purpose statements*
- Avatars
- Graphic organizers
- Item foreshadowing
- Directions and transitions
- Item resetting

* Purpose statements are not considered UDEs in the draft 2026 updated framework.

Not all features are available in every block, but all of the current operational NAEP reading blocks include some support features. Some of these features are available for all reading blocks, and across other NAEP subjects, at the system level (e.g., highlighting, text-to-speech on directions, zoom, and color themes). Others are content-specific, including item look-back buttons, pop-up notes, passage introductions, and multi-part response frames (complex items with multiple components split into multiple response fields). Others appear only in SBTs, or a subset of SBTs, depending on the goals of the tasks, including block-specific purpose statements, avatars, graphic organizers and sequential directions and transitions.

The following subsection provides additional information about the use of pre-reading features, pop-up notes, and avatars and pop-up notes.

Pre-Reading Features

The current operational assessment includes two types of pre-reading features: ***block-specific purpose setting statements*** and ***introductions*** to specific texts, which have been developed in consultation with the Reading Standing Committee[1] and approved by the Assessment Development Committee (ADC) on behalf of the Governing Board. The current NAEP Reading Framework does not provide guidance on pre-reading features.

Purpose Setting

DI blocks include general directions to "read and answer the questions," but do not include block-specific purpose statements.

SBTs include both general directions and block-specific purpose statements. Block-specific purpose statements introduce a purpose for reading and describe the task students are to complete (e.g., gather information for a webpage or to compose an email message). Block-specific purpose statements focus on the tasks students will perform rather than on introducing specific texts. Block-specific purpose setting statements appear in six of the blocks (2 per grade) in the 2019 operational reading pool (17% of the pool).

Introductions

A small number of DI blocks include some information about the text students are about to read prior to reading. This prereading feature has appeared in NAEP Reading since before the

---

[1] The Reading Standing Committee is a diverse group of experts and state assessment staff in reading from across the nation. They advise as part of the assessment item development process, ensuring that NAEP assessment items align to the NAEP framework. There is a Standing Committee for each NAEP subject area assessment.

digitally-based assessment began in 2017. All introductions are written text; none are multi-media (video or audio) as was proposed in the draft 2026 framework.[2]

Passage-specific introductions appear in eight of the blocks across all three grades in the 2019 operational reading pool (23% of the pool). Five of these introductions were added by the test developers and three were part of the original source. In five of these instances, the introduction provides some information about the author. In two of these instances, the introduction provides context for passages that are excerpts.

Generally, there are no consensus assessment industry guidelines or standards for when/how to provide introductions, though there is an extensive research base on the role of prior knowledge in reading comprehension that provides some guidance. For example, seminal research on schema theory by John Bransford and his colleagues found that readers were only able to adequately demonstrate their reading comprehension skills with passages written in general terms when titles were provided that activated their schema/prior knowledge about the topics of the passages. This work, along with content analyses of instructional materials and cognitive labs with students, enabled NCES to implement passage introductions in the operational NAEP reading assessment.

In addition, introductions were deemed important by the Reading Standing Committee as a means of orienting the reader and as a response to the need for content and face validity evidence. In timed, on-demand assessments such as NAEP, brief framing can help to mitigate construct-irrelevant variance, and such introductions and framings are common in sources students encounter in their daily lives. Periodically, the NAEP program invites all states and participating Trial Urban District Assessment (TUDA) districts to review the entire pool of NAEP items. The most recent state/TUDA item review in 2015 included texts with introductions, and no concerns were raised regarding these features. Finally, text introductions appear in some state reading assessments, such as PARCC and Smarter Balanced.

The following examples of passage introductions from previous NAEP reading assessments come from released and publicly available blocks (i.e., not the current operational pool). The first is an introduction to a Turkish folktale called "Five Boiled Eggs." The second introduces an article about the writer, E. B. White, and the third introduces an essay by E. B. White, by explaining that the author of the essay they are about to read is also a children's author. The E.B. White passages appeared in the paper assessment and were released in 2011. The "Five Boiled Eggs" passage appeared in the paper assessment and was transadapted for the digital assessment in 2017 and released after that administration.

---

[2] Responding to the Governing Board's March 2021 Board meeting deliberations, the April 2021 draft of the 2026 framework update does not include multimedia introductions.

Example 1.

Introduction: *Nasreddin Hodja, a character in this story, is familiar in many Turkish legends. "Hodja" means teacher.*

Long ago, a poor country boy left home to seek his fortune. Day and night he traveled, stopping to eat at inns along the way. Though he ate sparingly, his money quickly dwindled until, one day, no silver *akches* ✴ remained.

Still, the boy kept walking. Soon, however, his empty belly began to ache. Staggering up to the next inn he saw, he

Example 2.
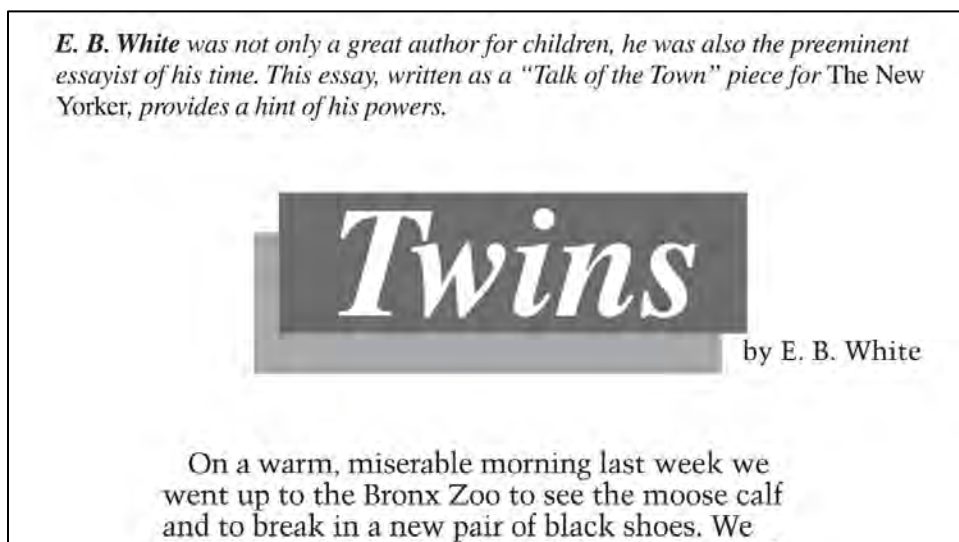
Meet the author: *E. B. White, the author of children's classics* Charlotte's Web *and* Stuart Little*, was also a great essayist.*

Not Just for Kids Anymore

"I have a lot of the cat in me," said author E. B. White, "and cats are not joiners."

Perhaps that is why White, one of the country's greatest writers, is so hard to label. His essays for *The New Yorker* appealed to an urbane crowd, but he is best remembered for his

Example 3.

E. B. White *was not only a great author for children, he was also the preeminent essayist of his time. This essay, written as a "Talk of the Town" piece for* The New Yorker, *provides a hint of his powers.*

Twins

by E. B. White

On a warm, miserable morning last week we went up to the Bronx Zoo to see the moose calf and to break in a new pair of black shoes. We
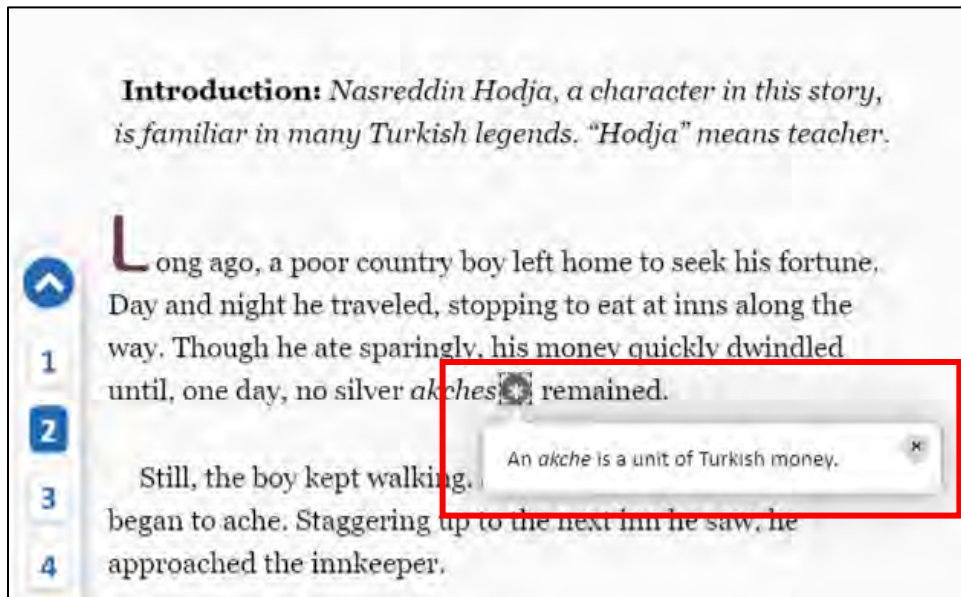
Avatars

Avatars are task characters used to create a social context and facilitate purpose-setting and transitions in SBT blocks (no discrete blocks use them) but are in and of themselves neither purpose statements nor introductions. Two of the total pool of six SBT blocks, across grades, in the 2019 operational reading pool use avatars (6% of the blocks in the total pool).

Pop-up Notes

Pop-ups are indicated by buttons in the text that signal to students that they can read more about a word or phrase. These kinds of notes appeared on the paper-based assessment (PBA) as footnotes. Pop-up notes occur in three blocks in the 2019 operational pool (9% of the blocks in the total pool). Two of the three pop-up instances provide definitions of words/terms that may be unfamiliar to the reader and are important to overall understanding. The third instance presents information that was provided in the original text. There are no assessment items directly related to the information in the pop-up notes.

The following example shows a pop-up note from the passage "Five Boiled Eggs."

Example 1.

A substantial proportion, 63%, of the entire pool of reading blocks in the 2019 operational assessment *does not contain the pre-reading features* described above. These blocks could be characterized as providing opportunities for "cold reads" and will continue to be part of the operational assessment in 2026.

Relevant NAEP Research

As noted above, two types of blocks make up the NAEP operational reading assessment: Discrete Item (DI) blocks and Scenario-based Task (SBT) blocks. At each grade level in 2019 (as noted in Table 1b above), two of 12 grade 4 blocks are SBTs, two of 15 grade 8 blocks are SBTs, and two of 15 grade 12 blocks are SBTs. The remainder are DI blocks. A special study was conducted in 2018 to examine the SBT format, relative to the current framework. For this study, researchers created discrete versions of reading SBT blocks using the same texts and items for both versions. This special study compared student performance on the same set of items and passages in a DI block versus an SBT block.

Although this study was conducted before the framework update project began, it is relevant to conversations about the framework update because SBTs involve collections of support features, which are referred to in the framework update as Universal Design Elements (UDEs). Both SBT blocks and DI blocks include UDEs.

Three of the 15 UDEs in the draft 2026 updated framework only appear in SBT blocks (i.e., avatars, sequential directions and transitions, and item resetting). The remaining 12 UDEs,

including text introductions and pop-up notes, can appear in either SBT or DI blocks.[3] This study provides no information about specific UDEs. Instead, the study examines collections of UDEs in an SBT format.

This was a randomized control trial study with a total of 3,000 students, counterbalanced for version, genre (literary and informational), and block position at each grade. Both the SBT and DI versions of blocks were delivered on tablets. Consistent with the students' experience with DI and SBT blocks in the operational assessment, students were able to move among texts and items at will in the discrete version, whereas movement between texts and items was sequential in the SBT versions.

Key findings (The differences summarized below are statistically significant.):

- Students taking the SBT versions of blocks outperformed students taking the DI versions of block in four of the six blocks.
- The advantage for the SBT versions was consistent across all NAEP subgroups (gender, race, SES, disability, ELL). In other words, there is no differential effect for any subgroup.
- The advantage of the versions with support features was consistent for low- and high-performing students in four of the six blocks.
- For the four blocks for which performance on the SBT version was significantly higher, the differences in percent correct ranged between 2% and 8%, with an average of 5%.
- The SBT-DI special study provides some indication that SBT versions of items tend to be more engaging/motivating to students than DI versions of items. This tendency could contribute to students' higher performance on SBT versions of items, compared with DI versions of items.
- Generally, reading SBT blocks tend to be equally or more difficult than DI blocks, but when comparing SBT and DI versions of the same set of items, SBTs tend to be less difficult than their DI versions.[4]
- Speededness was more of an issue in SBT versions. Revisions were made to reduce speededness before these blocks became part of the operational assessment.

---

[3] Of the 15 UDEs listed in the February 26, 2021 draft of the reading framework update, 13 already appear in the reading assessment. The 2 additions would have been: student exemplars as mentor texts (a task-based UDE) and multimedia passage introductions (a knowledge-based UDE). However, multimedia passage introductions were removed from the latest draft of the framework update. Text introductions already appear on the assessment – see earlier sections of this document on (1) pre-reading features and (2) existence of "cold reads".

[4] Because NAEP uses an Item Response Theory (IRT) model to generate scores, adding more difficult items to the NAEP Reading Item Pool will improve measurement at the high end of the score scale, i.e., detect smaller differences in student achievement for higher performers. Conversely, adding less difficult items will improve measurement on the low end of the score scale, i.e., detect smaller differences in student achievement for lower performers. The IRT methodology for scoring ensures that adding harder items to the item pool will not artificially lower scores and that adding easier items to the item pool will not artificially inflate scores.

Other Standard Research and Reviews in NAEP Item Development

NCES implements a routine research and development cycle to develop every assessment block carefully before it is introduced to the operational NAEP assessment. Each new block undergoes systematic scrutiny, typically including these steps:

1. **Text Selection.** Texts and text sets are identified by the ETS reading item development team at a rate of four for every one text or text set expected to become part of the operational assessment. Proposed texts are reviewed by the ETS bias and sensitivity review team and the ETS editorial staff and are ultimately reviewed and approved by the NCES item development staff and contractors, and the Governing Board Assessment Development Committee (ADC).

2. **Initial Item Reviews.** After passages are approved, items are developed by the ETS reading item development team. Once draft items are completed, ETS reviewers conduct editorial, cold read, bias and sensitivity, and language accessibility reviews. They are then reviewed by NCES item development staff and contractors and the reading standing committee.

3. **Pretesting.** Following initial item review, items and support features are pretested, using:
   a. Cognitive interviews with individual students to determine how they respond to proposed new texts and comprehension test items. The purpose is to determine whether the tasks actually engage students in the intended comprehension processes.
   b. Tryouts under "live" testing conditions with 50–200 students from the target population to determine whether a wide range of students can complete the blocks within the allocated time and whether all of the parts of the block are working as intended.
   c. Usability studies, which test new item or passage interactions with small groups of students.

4. **Revised Item Reviews**. After items are pretested and revised by item developers, ETS reviewers conduct editorial, bias and sensitivity, and language accessibility reviews. They are then reviewed by NCES item development staff and contractors and the standing committee. Item revisions are adjudicated with NCES item development staff, and items are submitted to the Governing Board ADC for final review and clearance for piloting. Before piloting, state/TUDA reviews may occur.

5. **Piloting.** Proposed new blocks are folded into the administration of operational blocks of a live assessment. By comparing student and item performance across the new and the old blocks, NAEP developers can determine whether the new blocks effectively scale together with the old, measuring the same underlying comprehension construct.

6. **Post-pilot Reviews**. Following the collection of pilot data (n=2500–3000 students per form), the following groups review pilot data, item level analyses, texts, and items:

- ETS reading item development team
- ETS data analysis and reporting team
- ETS Differential Item Functioning (DIF) panel
- ETS bias and sensitivity review team
- NCES item development staff and contractors
- NCES data analysis and reporting staff and contractors
- ETS editorial staff
- ETS Reading Standing Committee
- Governing Board Assessment Development Committee (ADC)

III.   IMPLEMENTING THE UPDATED FRAMEWORK AND MAINTAINING TREND

This section provides information about the implementation of the updated framework and is based on the contents of the latest draft of that document.

Following Board adoption of an updated framework, it will take time to develop the assessment. As new content is piloted and approved, old content, in particular blocks transadapted from the paper-based assessment, can be phased out. Most importantly, this gradual item development for the updated framework allows for trend to be maintained.

The 2022 and 2024 assessments will be the last operational assessments that are fully aligned to the current framework. The 2026 assessment is projected to be the first operational assessment under the updated framework. The 2026 assessment would include both trend blocks from the 2022 and 2024 operational assessments and newly developed blocks piloted in 2024, being used for the first time in an operational assessment.

In the Governing Board's previous discussions of the updated framework, concerns were expressed that there would be insufficient carryover of content to maintain trend in 2026. However, the information below indicates that maintaining trend in 2026 is possible with careful planning. The projected contents of the next three operational assessments are as follows:

- 2022:
    - Grades 4 and 8 – trend content only (re-administration of 2019)
    - Grade 12 – no assessment
- 2024:
    - Grades 4 and 8 – trend content (all blocks carried over from 2022) plus new operational content (drawn from blocks piloted in 2017 and 2019)
    - Grade 12 – trend content only (re-administration of 2019)

- 2026:
  - Grades 4 and 8 – trend content (all blocks carried over from 2024) and new operational content (drawn from blocks piloted in 2024)
  - Grade 12 – no assessment

Projected Numbers of Blocks Available for the 2026 Operational Reading Assessment

The tables below include information about the numbers of blocks in each of the following two categories that will make up the 2026 operational assessment.

1. **Trend blocks**, which consist of discrete blocks from the 2022 and 2024 operational assessments, which do not include block-specific purpose statements, and SBT blocks from the 2024 operational assessments, which do include block-specific purpose statements.
2. **New operational blocks** developed to address new aspects of the updated framework, including block-specific purpose statements and the updated comprehension targets. These blocks are being used for the first time in the 2026 operational assessment and will not become trend blocks until they are administered operationally for the second time.

The proposed approach to a gradual implementation of the updated framework has been revised since the original Overview document was submitted to Governing Board staff just prior to the March 2021 Board meeting. The March 2021 version of this document suggested adding block-specific purpose statements to three existing discrete blocks at each grade and re-piloting them in 2024. However, the most recent plan retains the existing discrete blocks, as is, and redirects the funds that would have been used for modification and re-piloting of existing blocks to the development of new blocks under the aegis of the updated framework. The current plan provides for a carryover of blocks from the 2024 to the 2026 assessment of 80% at grade 4 and 83% at grade 8. **Although ultimately an empirical question, these percentages of carryover should allow for the maintenance of trend.**[5] (See Table 3a below.)

The current plan for the 2026 development proposes new pilot development of six blocks at grades 4 and 8[6] to yield four new operational blocks. It also assumes that blocks piloted in 2017, 2019, and 2024 will be approved for operational use and that there are no public releases prior to the 2026 assessment.

Tables 3a, 3b, and 3c provide information about the composition of the 2026 operational assessment based on the current plan.

---

[5] The current NAEP reading framework – adopted in 2004 and first implemented in 2009 – included no carryover from the previous framework (0 percent) and trend was maintained. To learn more about how trend was maintained for the 2009 NAEP Reading Assessment, see the Reading Trend Study description at https://nces.ed.gov/nationsreportcard/reading/trend_study.asp.

[6] Grade 12 will not be administered in 2026 and new grade 12 development is out of scope.

Table 3a. Projected Numbers of Blocks by Status available for the 2026 NAEP Operational Reading Assessment at Grades 4 and 8

Blocks	Grade 4	Grade 8
TREND	16 (80%)	20 (83%)
NEW OPERATIONAL	4	4
Total Blocks	**20**	**24**

As a result of needing to both maintain trend and introduce new content aligned with the updated framework, the 2026 operational assessment is projected to include more blocks at each grade than the 2022 operational assessment. The grade 4 assessment would contain 11 blocks in 2022[7] and as many as 20 in 2026, and the grade 8 assessment would contain 14 blocks in 2022 and as many as 24 blocks in 2026.[8] A larger item pool is also required to support reporting goals for the updated framework, including reporting for three subscales instead of the two subscales reported under the current framework.

All of the passages and items in the blocks that would be carried over from 2024 to 2026 are consistent with the updated framework. The block-specific purposes required by the updated framework will be present in 40% of the blocks at grade 4 and 33% of the blocks at grade 8.

Tables 3b and 3c describe the contents of the projected 2026 operational assessment at each grade broken down by subscale.

---

[7] Tables 1a and 1b showed that the 2019 assessment included 12 blocks at grade 4 and 15 blocks at grade 8. However, one cross-grade 4/8 block has been dropped for sensitivity reasons, resulting in 11 blocks at grade 4 and 14 blocks at grade 8 for the 2022 assessment. Blocks sometimes need to be dropped for sensitivity reasons if they address topics that might be disturbing because of recent or ongoing current events, e.g., a hurricane, a pandemic, etc.

[8] The actual number of 2026 blocks is contingent on the contents of possible public releases in 2022 and 2024.

Table 3b. Projected Number of Blocks available for the 2026 NAEP Operational Reading Assessment by Status and Subscale at Grade 4

Blocks	Reading in Literature	Reading in Social Studies	Reading in Science	Total Blocks
TREND	7	4	5	16 (80%)
NEW OPERATIONAL	*New development would include at least one block in each of the reading in social studies and science contexts.*			4
Total Blocks				**20**

Table 3c. Projected Number of Blocks available for the 2026 NAEP Operational Reading Assessment by Status and Subscale at Grade 8

Blocks	Reading in Literature	Reading in Social Studies	Reading in Science	Total Blocks
TREND	8	6	6	20 (83%)
NEW OPERATIONAL	*New development would include at least one reading in literature block.*			4
Total Blocks				**24**

Appendices 1 and 2 on the following pages depict the movement of blocks across the 2022, 2024, and 2026 assessments at grades 4 and 8, as well as the addition of newly developed blocks.

Appendix 1. Proposed Composition of the 2022, 2024, and 2026 Assessments at Grade 4 by Context and Status

2022 Assessment	2024 Assessment	2026 Assessment
Rdg in Lit Block 1	Rdg in Lit Block 1	Rdg in Lit Block 1
Rdg in Lit Block 2	Rdg in Lit Block 2	Rdg in Lit Block 2
Rdg in Lit Block 3	Rdg in Lit Block 3	Rdg in Lit Block 3
Rdg in Lit Block 4	Rdg in Lit Block 4	Rdg in Lit Block 4
Rdg in Lit Block 5	Rdg in Lit Block 5	Rdg in Lit Block 5
	Rdg in Lit Block 6	Rdg in Lit Block 6
	Rdg in Lit Block 7	Rdg in Lit Block 7
Rdg in Science Block 1	Rdg in Science Block 1	Rdg in Science Block 1
Rdg in Science Block 2	Rdg in Science Block 2	Rdg in Science Block 2
Rdg in Science Block 3	Rdg in Science Block 3	Rdg in Science Block 3
Rdg in Science Block 4	Rdg in Science Block 4	Rdg in Science Block 4
	Rdg in Science Block 5	Rdg in Science Block 5
Rdg in SocSt Block 1	Rdg in SocSt Block 1	Rdg in SocSt Block 1
Rdg in SocSt Block 2	Rdg in SocSt Block 2	Rdg in SocSt Block 2
	Rdg in SocSt Block 3	Rdg in SocSt Block 3
	Rdg in SocSt Block 4	Rdg in SocSt Block 4
	Pilot Block A	New Op Block
	Pilot Block B	New Op Block
	Pilot Block C	New Op Block
	Pilot Block D	New Op Block
	Pilot Block E	
	Pilot Block F	

KEY
Trend Block
New Operational Block
Pilot Block

Appendix 2. Proposed Composition of the 2022, 2024, and 2026 Assessments at Grade 8 by Context and Status

2022 Assessment	2024 Assessment	2026 Assessment
Rdg in Lit Block 1	Rdg in Lit Block 1	Rdg in Lit Block 1
Rdg in Lit Block 2	Rdg in Lit Block 2	Rdg in Lit Block 2
Rdg in Lit Block 3	Rdg in Lit Block 3	Rdg in Lit Block 3
Rdg in Lit Block 4	Rdg in Lit Block 4	Rdg in Lit Block 4
Rdg in Lit Block 5	Rdg in Lit Block 5	Rdg in Lit Block 5
Rdg in Lit Block 6	Rdg in Lit Block 6	Rdg in Lit Block 6
	Rdg in Lit Block 7	Rdg in Lit Block 7
	Rdg in Lit Block 8	Rdg in Lit Block 8
Rdg in Science Block 1	Rdg in Science Block 1	Rdg in Science Block 1
Rdg in Science Block 2	Rdg in Science Block 2	Rdg in Science Block 2
Rdg in Science Block 3	Rdg in Science Block 3	Rdg in Science Block 3
Rdg in Science Block 4	Rdg in Science Block 4	Rdg in Science Block 4
Rdg in Science Block 5	Rdg in Science Block 5	Rdg in Science Block 5
Rdg in SocSt Block 1	Rdg in Science Block 6	Rdg in Science Block 6
Rdg in SocSt Block 2	Rdg in SocSt Block 1	Rdg in SocSt Block 1
Rdg in SocSt Block 3	Rdg in SocSt Block 2	Rdg in SocSt Block 2
	Rdg in SocSt Block 3	Rdg in SocSt Block 3
	Rdg in SocSt Block 4	Rdg in SocSt Block 4
	Rdg in SocSt Block 5	Rdg in SocSt Block 5
	Rdg in SocSt Block 6	Rdg in SocSt Block 6
	Pilot Block A	New Op Block
	Pilot Block B	New Op Block
	Pilot Block C	New Op Block
	Pilot Block D	New Op Block
	Pilot Block E	
	Pilot Block F	

KEY
Trend Block
New Operational Block
Pilot Block

NCES Response to the Committee on Standards, Design and Methodology (COSDAM) Reading Framework Questions

April 22, 2021

This document has been prepared in response to questions from the National Assessment Governing Board's COSDAM regarding the current operational NAEP Reading Assessment in relation to the most recent draft of the 2026 updated framework. Three groups of questions are addressed in this document: 1) questions regarding Universal Design Elements (UDEs); 2) questions about the construct(s) being measured and the feasibility of maintaining trend; and 3) questions about implementation plans, projections, and budget considerations.

Universal Design Elements (UDEs) Questions

> What research evidence was used to implement the "support features" on the assessment, in particular the passage introductions?
>
> Is there existing evidence that knowledge-based UDEs are differentially effective based on students' prior knowledge?
>
> Is there any existing evidence regarding the "effect size" of UDEs on performance?

UDEs, such as introductions, have been part of the NAEP Reading Assessment since before NAEP became a digital assessment in 2017. In general, there are no assessment industry guidelines or standards for when/how to provide introductions, though there is an extensive research base on the role of prior topic knowledge in reading comprehension that provides some guidance. For example, seminal research on schema theory by John Bransford and his colleagues found that readers were only able to adequately demonstrate their reading comprehension skills with passages written in general terms when titles were provided that served to activate their schema/prior knowledge about the topics of the passages. This work, along with content analyses of instructional materials and cognitive interviews with students, provided justification for NCES to implement passage introductions in the operational NAEP Reading Assessment. It is also the case that text introductions appear in some state reading assessments, such as the Partnership for Assessment of Readiness for College and Careers (PARCC) and Smarter Balanced Assessment Consortium.

Examining the differential effectiveness of introductions and pop-up notes based on students' prior topic knowledge would require a study that includes measures of students' background knowledge. The NAEP program has not conducted any such study. Similarly, the NAEP program does not have evidence regarding the "effect size" of UDEs on performance because NAEP is not primarily a research program. NAEP relies on data from a variety of quantitative and qualitative sources to inform its development including cognitive interviews, small-scale tryouts, content reviews, and, occasionally, special studies. All new NAEP reading blocks are evaluated in a nationally representative pilot, followed by rigorous, block- and item-level analyses, and submitted for further review by the Governing Board's Assessment Development Committee (ADC). Weak or problematic blocks are not moved on for inclusion in the operational assessment.

Pretesting through cognitive interviews and small-scale tryouts was used to explore new UDEs introduced with Scenario-based Tasks (SBTs). The evidence from pretesting indicated that the majority of students reported that SBT UDEs were helpful and not distracting.

> How much time do existing UDEs add to testing? Would additional UDEs exacerbate this further?
>
> Is there any evidence about whether this additional time might hinder performance or be distracting?

The incorporation of UDEs in reading blocks is an integral part of the development of 30-minute blocks, as opposed to an "add on." Any potential time or cognitive burden they may pose is evaluated for each block as part of the development and pretesting processes via cognitive interviews and small-scale tryouts. Independent of the impact of UDEs, speededness is evaluated and addressed for all blocks as part of the development, pretesting, and piloting processes. The majority of the UDEs recommended in the draft framework are already included in the reading blocks in the NAEP operational assessment.

> Is NCES concerned about the framework's characterization of "support features" as UDEs given how the NAEP program already characterizes "Universal Design Elements?"

NCES describes "Universal Design Elements" as a form of accommodation that is available to all students. Some of the UDEs described by NCES, such as highlighting and zoom, are considered "task-based UDEs" in the updated framework. NCES is not concerned that the updated framework adds UDEs that are not specified in its current description of UDEs.

> What should be the main takeaways from the Scenario-based Task—Discrete Block (SBT-DI) study, relating to this framework update?

The main takeaway from the SBT-DI study is that regardless of ability (low vs. high), students performed better on the SBTs[1]. The main difference between the SBT and DI versions was the use of purpose-driven introductions and a broad range of UDEs in the SBTs. The support features in SBTs examined in the study are consistent with the updated framework.

Construct/Trend Questions

> In NCES's view, does the current framework and the framework update both allow for "cold reads?"

The assessment has not been operationalized in terms of "cold reads" as the term is not defined or discussed in either the current or draft updated framework.

> What evidence/ongoing studies/best guesses are there related to the likelihood of maintaining trend? Is this only a function of how many new blocks are needed?
>
> In NCES's view, do the newly proposed UDEs (mentor texts and multi-media introductions) represent a change in the construct of reading that can threaten trend (on a conceptual level)?
>
> In NCES's view, do other strictly digital UDEs from the framework represent a change in the construct of reading that should have already threatened trend (on a conceptual level)?

The likelihood of maintaining trend is a function of both how many new blocks are needed and whether these blocks differ qualitatively from the existing blocks in terms of what they measure. That said, there is a high likelihood of maintaining trend under the updated framework. Evidence of this comes from several sources. First, the construct of reading comprehension in the updated framework has changed very little from the construct in the current framework. This means that the passages and items developed under the updated framework will not differ significantly from those that were developed under the current framework. Second, the current plan of gradual implementation results in carryover of approximately 80% of blocks from the 2024 to the 2026 operational assessments. Finally, it should be noted that trend was maintained in the implementation of the current framework when there was no carryover from the previous assessment.

The possibility that the new UDEs recommended in the draft framework (mentor texts and multimodal introductions) could threaten the construct is also an empirical question that will be investigated through means such as pretesting and, possibly, special studies and evaluated

---

[1] The magnitude of the improvement, in terms of percent correct, ranged between 2% and 8% with an average of 5%. Although the SBT versions were less difficult than the DI versions of the same texts and items, evidence from the operational assessment indicates that SBTs are of equal or greater difficulty than the DI blocks in the operational pool.

through the various reviews described in the most recent memorandum—*Overview of the NAEP Reading Assessment and Projections*—to the ADC.

The "strictly digital" UDEs introduced with SBTs in 2019 are not a threat to trend either conceptually or empirically. Conceptually, they are consistent with the construct of reading comprehension in the current framework as a means of measuring students' ability to "use meaning as appropriate to type of text, purpose, and situation" (part of the definition from the current NAEP Reading Framework [2009]). Empirically, SBT blocks containing these UDEs scale with Discrete Blocks that do not include these UDEs.

Implementation and Budget Questions

> To what extent can the current reading item pool be used to implement the framework update? Specifically, how much re-field testing is needed and how much new item development is needed?
>
> What is the cost of implementing the framework update?

NCES's proposed approach to a gradual implementation of the updated framework has been revised since the March 2021 Board meeting (see also in this packet of materials—*Overview of the NAEP Reading Assessment and Projections*). The current plan increases the percentage of trend blocks carried over from the 2024 to the 2026 assessments, to 80% at grade 4 and 83% at grade 8. Although ultimately an empirical question, these percentages of carryover should allow for the maintenance of trend.

All of the passages and items in the blocks that would be carried over from 2024 to 2026 are consistent with the updated framework. The block-specific purposes required by the updated framework will be present in 40% of the blocks at grade 4 and 33% of the blocks at grade 8. The remainder of trend blocks include general purpose statements.

NCES expects to develop and pilot six new blocks at each of grades 4 and 8 to yield four new operational blocks. It also assumes that blocks piloted in 2017, 2019, and 2024 are approved for operational use, and there are no public releases prior to the 2026 assessment.

The cost of implementing the updated framework can only be determined when an updated framework has been approved.

NAEP READING FRAMEWORK UPDATE

TECHNICAL ADVISORY COMMITTEE GUIDANCE FOR THE DEVELOPMENT PANEL

OVERVIEW

The NAEP Reading Framework Technical Advisory Committee (TAC) is a group of eight experts in psychometrics and large-scale assessment. The TAC's role in the NAEP Reading Framework update process is to support the Development Panel (DP), addressing measurement and assessment questions as they surface. Two members of the TAC attend each Development Panel meeting. After Development Panel meetings, the full TAC convenes virtually to address specific questions from the previous Panel meeting, and to provide guidance for the subsequent Panel meeting. The TAC met for the seventh time on March 26, 2021. The objective of the meeting was to discuss the full Framework document and to offer guidance to support fine-tuning the Framework in advance of submission to the Governing Board. The TAC discussion focused on primarily universal design elements (UDEs) and topical knowledge. TAC members offered more general thoughts on the 2026 Framework as well. The TAC's March 2021 feedback and recommendations are summarized below.

VALIDITY RESEARCH ON UNIVERSAL DESIGN ELEMENTS

The TAC discussion began with a focus on evidentiary standards for UDEs. In and of themselves, UDEs are neither valid nor invalid. Rather, assessment developers examine the extent to which these features minimize construct-irrelevant variance (i.e., when factors unrelated to the intended subject of the test influence performance on the test). Similarly, UDEs should not inadvertently *create* bias by providing an advantage to particular student groups.

In the assessment accommodations literature, statistical examinations for an accommodation's impact is often carried out via multiple regression (e.g., where test scores or item responses are regressed on [1] presence of a disability such as visual impairment, [2] use of an accommodation such as Braille, and [3] the interaction of [1] and [2].) From a validity standpoint, a positive interaction effect is good: it indicates that on average, the accommodation increases scores, but only for the students who are supposed to receive it.

The TAC agreed, however, that UDEs in the 2026 Reading Framework are not accommodations; None are intended to help one group of students over another. Therefore, in a multiple regression analysis focused on the 2026 Reading Framework's UDEs, looking for main effects – not interaction effects – would be the first order of business.[1]

---

[1] Note that multiple regression is one among many tools test developers use to determine whether certain features of an assessment are doing the job they were intended to do for the populations they are intended to support.

KNOWLEDGE-BASED UNIVERSAL DESIGN ELEMENTS

The TAC spent roughly half of the meeting discussing knowledge-based UDEs (e.g., a short introduction to a potentially unfamiliar topic, available to all students), including how they affect validity and fairness, how often they are used in large-scale assessments, and how their common pitfalls can be avoided.

The reading comprehension construct in the 2026 Reading Framework does not assume prior disciplinary knowledge ("items should not ask readers to draw upon text-independent domain knowledge") and the only two knowledge-based UDEs under consideration are glossaries and short introductions, both of which are standard features of large-scale summative assessment. The Framework provides reassurance that appropriate safeguards for the recommended UDEs are in place, and that the more ambitious, potentially problematic UDEs are not slated for the 2026 assessment and will instead be the subject of further validation research.

The TAC believes it will be useful for the DP to further reinforce these points in the narrative. That is, in each instance that knowledge-based UDEs are recommended in the Framework, the DP should clarify that they are based on substantial precedent, represent best practice, and, in fact, are uncontroversial. Examples will help, whether in the Framework, in the Assessment and Item Specifications, or in both documents. Examples are especially useful for the disciplinary contexts, such as reading in science, where background knowledge arguably presents the clearest potential threat to unbiased measurement of reading comprehension.

TOPICAL KNOWLEDGE AND GENERAL IMPRESSIONS OF THE REVISED FRAMEWORK

A related issue to UDEs is the relative importance of topical knowledge in reading comprehension and, by extension, the appropriate emphasis on topical knowledge in a reading comprehension assessment. The TAC reflected on the DP's treatment of topical knowledge, and the discussion served as a springboard to more general reactions to the revised draft Framework. Both discussions are summarized briefly here.

TOPICAL KNOWLEDGE

The Framework conceptualizes topical knowledge as separate from reading comprehension. So, for the same reasons knowledge-based UDEs are encouraged, items that draw upon topical knowledge are discouraged. Specifically,

> ...items should not assess knowledge sources irrelevant to the items and associated comprehension targets in a given block. For example, items should not ask readers to draw upon text-independent domain knowledge, topic knowledge, knowledge of

---

Other techniques, such as cognitive interviews and classroom tryouts (carried out for NAEP routinely during item development) generate different types of evidence, equally important to the overall validity argument.

> *technical vocabulary or idiomatic expressions, or conceptual or domain knowledge in particular subject areas.*

The DP wants to deliver a Framework that gives NAEP the best possible chance of measuring reading comprehension as it is defined above – untethered to topical knowledge. To that end, the DP sought to confirm with the TAC that (1) the Framework's stance on topical knowledge is in keeping with modern assessment practice, and that (2) the associated rationales provided in the Framework are well aligned with modern validity theory.

This topic generated relatively little discussion. The TAC was unanimous in its support both of the DP's decision to exclude topical knowledge from the NAEP Reading construct and of the convincing rationales presented in the Framework. The DP thought it was important to be clearer about topical knowledge in the NAEP Reading Framework update. As a result, the revised framework document addresses the issue head-on. This is rare in large-scale testing; only a few states even address the issue of topical knowledge in their definition of reading comprehension, and none argue that topical knowledge should be measured as a component of reading comprehension.

GENERAL IMPRESSIONS OF THE REVISED FRAMEWORK.

The topical knowledge discussion provided the TAC an opportunity to voice its support for the revised 2026 Reading Framework. Although very few states currently address potentially controversial issues such as topical knowledge, bias, and responsible reporting, the TAC agreed that there are strong arguments that NAEP has a responsibility to be clear on these issues.

LIST OF TAC MEMBERS

Derek C. Briggs, University of Colorado, Boulder

Howard Everson, SRI International

Joan Herman, National Center for Research on Evaluation, Standards, and Student Testing (CRESST)

Kristen L. Huff, Curriculum Associates

Michael Kolen, University of Iowa

Scott Marion, The National Center for the Improvement of Educational Assessment

Jennifer Randall, Center for Educational Assessment, University of Massachusetts, Amherst

Guillermo Solano-Flores, Stanford University

Information from State Assessment Programs
April 28, 2021

21 of the 50 states responded to a recent poll conducted by the Council of Chief State School Officers (CCSSO).[1] Additionally, Board staff shared the same questions with representatives from Smarter Balanced and New Meridian (which manages the assessments formerly called PARCC).

- All 21 (plus New Meridian and Smarter Balanced) have definitions of reading comprehension that <u>do not explicitly</u> include topic or background knowledge as part of what is intended to be measured on their summative assessments.

- 19 of the 21 states (plus New Meridian and Smarter Balanced) attempt to mitigate the impact of background knowledge through selecting a wide range of passage topics.

- 17 of 20 states (plus New Meridian and Smarter Balanced) attempt to mitigate the impact of background knowledge by providing general information about the passage in the assessment, e.g., briefly introducing the topic, time period, author, or context of the passage. (1 state did not respond.)

- New Meridian, Smarter Balanced, and all 21 responding states attempt to mitigate the impact of background knowledge through defining terms in the passages that may be unfamiliar to the reader.

- New Meridian, Smarter Balanced, and 14 of 15 states attempt to mitigate the impact of background knowledge through another means. (6 states did not respond)

---

[1] The 21 states that responded include: Alabama, Arizona, Arkansas, Delaware, Georgia, Iowa, Kentucky, Maryland, Mississippi, New York, North Carolina, Ohio, Oklahoma, Oregon, South Carolina, South Dakota, Virginia, Washington, West Virginia, Wisconsin, and Wyoming