READING FRAMEWORK

for the 2026 NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

NATIONAL ASSESSMENT GOVERNING BOARD
U.S. Department of Education

WHAT IS NAEP?

The National Assessment of Educational Progress (NAEP) is a continuing and nationally representative measure of trends in academic achievement of U.S. elementary and secondary students in various subjects. For nearly four decades, NAEP assessments have been conducted periodically in reading, mathematics, science, writing, U.S. history, civics, geography, and other subjects. By collecting and reporting information on student performance at the national, state, and local levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education.

THE 2020–2021 NATIONAL ASSESSMENT GOVERNING BOARD

The National Assessment Governing Board was created by Congress to formulate policy for NAEP. Among the Governing Board's responsibilities are developing objectives and test specifications and designing the assessment methodology for NAEP.

MEMBERS

Honorable Haley Barbour, Chair
Founding Partner
BGR Group
Yazoo City, Mississippi

Honorable Alice Peisch, Vice Chair
Massachusetts House of Representatives
Wellesley, Massachusetts

Dana K. Boyd
Principal
Parkland Elementary School
El Paso, Texas

Alberto M. Carvalho
Superintendent
Miami–Dade County Public Schools
Miami, Florida

Gregory J. Cizek
Guy B. Phillips Distinguished Professor of Educational Measurement and Evaluation
University of North Carolina
Chapel Hill, North Carolina

Tyler W. Cramer
CEO and Executive Manager
Remarc Associates LLC
San Diego, California

Christine Cunningham
Professor of Education and Engineering
College of Education
The Pennsylvania State University
University Park, Pennsylvania

Frank Edelblut
Commissioner
New Hampshire Department of Education
Concord, New Hampshire

Paul Gasparini
Secondary School Principal
Jamesville–DeWitt High School
DeWitt, New York

Honorable James E. Geringer
Former Governor of Wyoming
Cheyenne, Wyoming

Eric Hanushek
Hanna Senior Fellow
Hoover Institution
Stanford, California

Patrick L. Kelly
Director of Governmental Affairs
Palmetto State Teachers Association
Columbia, South Carolina

Suzanne Lane
Professor of Research Methodology
University of Pittsburgh
Pittsburgh, Pennsylvania

Tonya Matthews
Chief Executive Officer
International African American Museum
Charleston, South Carolina

Reginald McGregor
Manager, Engineering Employee Development & STEM Outreach
Rolls Royce Corporation
Indianapolis, Indiana

Mark Miller
Eighth–Grade Mathematics Teacher and Department Chair
Cheyenne Mountain Junior High
Colorado Springs, Colorado

Honorable Beverly Perdue
Former Governor of North Carolina
New Bern, North Carolina

Julia Rafal–Baer
Managing Partner & Co–Founder
ILO Group
Cranston, Rhode Island

Ron Reynolds
Executive Director
California Association of Private School Organizations
Van Nuys, California

Nardi Routten
Fourth–Grade Teacher
Creekside Elementary School
New Bern, North Carolina

Martin R. West
Massachusetts Board of Elementary and Secondary Education
Professor of Education
Harvard Graduate School of Education
Cambridge, Massachusetts

Honorable Mark White
Tennessee House of Representatives
Nashville, Tennessee

Grover J. "Russ" Whitehurst
Professor Emeritus
Stony Brook University
Fort Myers, Florida

Carey M. Wright
State Superintendent
Mississippi Department of Education
Jackson, Mississippi

Ex–officio Member
Mark Schneider
Director
Institute of Education Sciences
Washington, D.C.

Reading Framework for the 2026 National Assessment of Educational Progress

National Assessment Governing Board
U.S. Department of Education

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

This Framework is dedicated to Carol Connor and Robert Rothman, members of the Development Panel, and Matthew Gaertner, project Measurement Specialist, who passed away during the course of Framework Development. Their ideas, their words, and their commitments to a more robust and equitable educational system are embodied on every page of this Framework.

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 Office of the Chancellor, 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, Educational Studies
School of Education
University of Michigan
Ann Arbor, MI

Byeong-Young Cho*
Associate Professor, Literacy
Hanyang University
Seoul, Republic of Korea

Julie Coiro*
Professor, Education
University of Rhode Island
Kingston, RI

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
College Park, MD

Bonnie Hain*
Senior Director, Market Solutions and Services
ACT
Iowa City, IA

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

Kathleen Hinchman*
Professor Emeritus, Education
Syracuse University
Syracuse, NY

Christy Howard
Associate Professor, Literacy Studies
East Carolina University
Raleigh, NC

Panayiota Kendeou
Distinguished McKnight University Professor and Guy Bond Chair in Reading
Department of Educational Psychology
University of Minnesota
Minneapolis, MN

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

Carol Lee*
Edwina S. Tarry Professor Emerita, School of Education and Social Policy
Northwestern University
Evanston, IL

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

Mariana 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, Program Connections and Content Strategy
College Board
Coralville, IA

P. David Pearson, Panel Chair*
Evelyn Lois Corey Emeritus Professor of Instructional Science
Graduate School of 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
New Milford, 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
Superintendent
Harrisburg, PA

Paola Uccelli*
Professor, Education
Harvard University
Cambridge, MA

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

Victoria Young
Independent Reading and Writing Consultant
Austin, TX

Technical Advisory Committee

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

Howard Everson
Senior Principal Research Scientist
SRI International
Menlo Park, CA
Professor of Education (Visiting)
Graduate Center, City University of New York

Joan Herman
Director Emerita, National Center for Research on Evaluation, Standards, and Student Testing (CRESST)
Los Angeles, CA

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

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

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

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

Guillermo Solano-Flores
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 who 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 NAEP Authorization Act of 2002 (NAEP, 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.

The law specifies that NAEP's purpose is "to provide, in a timely manner, a fair and accurate measurement of student academic achievement and reporting of trends in such achievement in reading, mathematics, and other subjects[s] …" (section 303(b)(1), NAEP Authorization Act of 2002, P.L. 107–279). The NAEP Reading Assessment data will measure national, regional, and subgroup trends in reading achievement but will not target the performance of individual students or schools.

By law, NAEP assessments shall not evaluate personal beliefs or publicly disclose personally identifiable information, and NAEP assessment items shall be secular, neutral, and non-ideological and free from racial, cultural, gender, or regional bias.

Current NAEP Reading Assessment in a Digital Environment

The Governing Board, the policymaking body for NAEP, has stated that the NAEP Reading Assessment will measure reading comprehension by asking students to read passages written in English and to answer questions about what they have read. The Framework "shall focus on important, measurable indicators of student achievement … without endorsing or advocating a particular instructional approach" (National Assessment Governing Board, 2018a). Although broad implications for instruction may be inferred from the Assessment, NAEP does not specify how reading should be taught; nor does it prescribe a particular curricular approach to teaching reading.

Furthermore, the Governing Board recognizes that there is great value in ensuring continuity in the NAEP Reading Framework in order to report student achievement trends over time, which is an important function of the NAEP program.

The NAEP Reading Assessment has been administered on a digital platform since 2017. 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). Schools and students participating in NAEP assessments are supported in various ways so they can successfully engage with the digitally based assessment. For both discrete and scenario-based tasks (SBTs) assessment blocks, tools available to all students include annotation via an on-screen pencil or highlighter, selection of color themes, and zoom-in functionality. 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, multipart response frames, and more. Not all features are available in every block, but all blocks include some 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. These multimodal features serve functions that are present in authentic text (e.g., in school settings, graphics occur frequently in science passages, and videos are used to prime students' interest in a topic). The multimedia features are not designed to provide information that would increase the comprehension scores of students who would otherwise struggle to understand the text itself. 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. 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. 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 Governing Board (National Assessment Governing Board, 2019a) 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; and the Panels are also tasked with considering opportunities to extend the depth of measurement and reporting given the affordances of digital based assessment. The update process shall result in three documents: a recommended framework, assessment and item specifications, and recommendations for contextual variables that relate to student achievement in reading (p.3)

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 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 et al., 2019, p. 45). Accordingly, the Visioning Panel set forth recommendations 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.

The Governing Board has a continuing commitment to equity in our educational systems. It advances this goal by designing assessments that are inclusive and accessible for the full diversity of students who are administered the NAEP Assessments. The assessments will align with the recent standards in large-scale assessment by continuing to strive to minimize test bias to the maximum extent possible (American Educational Research Association, American Psychological Association, and National Council of Measurement in Education, 2014; International Test Commission, 2019; IRA/NCTE Joint Task Force on Assessment, 2010). Finally, the Assessment will gather data that afford opportunities to examine malleable contextual variables that may lead to greater understanding of differential student achievement.

As a result, the Visioning Panel worked 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 Visioning 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 recommendations, 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 framework that the Development Panel ultimately recommended to the Governing Board went through several iterations by the Development Panel to address feedback from various external parties and from members and committees of the Governing Board. It underwent further revisions by the Governing Board as a final step in the consensus-building process that is mandated by the NAEP law.

The Updated NAEP Reading Framework

This updated framework for the 2026 NAEP Reading Assessment acknowledges that reading is a complex process shaped by many factors. Learning—and reading—are, 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). Research evidence has highlighted that like all human learning, reading

comprehension is a meaning-making activity that involves socially and culturally specific characteristics and practices (Bronfenbrenner & Morris, 2006; Lee, 2016b, 2020; NASEM, 2018; Pacheco 2015, 2018; Skerrett, 2020; Zelazo, 2013).

Drawing from previous frameworks and newer understandings, this updated NAEP Reading Framework attends to four key features of reading comprehension—contexts, readers, texts, and activities. At the heart of the 2026 NAEP Reading Framework is the definition of reading comprehension:

> **Reading comprehension is making meaning with text,** a complex process shaped by many factors, including readers' abilities to
>
> - 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 social and cultural contexts.

This definition applies to the assessment of reading achievement on NAEP and is not intended to be an inclusive definition of reading or reading instruction.

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

- what readers know about a topic
- what readers know about texts and how they work
- internal processes, or foundational skills, needed to render texts 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
- socially and culturally situated knowledge and practices from home, community, and school 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 (NAGB, 2017a; Rogers et al., 2016). These advances have also allowed the assessment design to gather more information on experiences and factors that influence the cognitive processes central to 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 (UDEs) 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, and

- increase the quantity and quality of information that is available to users of NAEP data on student reading achievement in the United States.

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. 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 (NCES, 2018; 2020). Framework updates also reflect trends in international reading comprehension assessments, such as PISA and 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. 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 informational UDEs are included as appropriate to support 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 provide more nuanced 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 (a) socioeconomic status within race/ethnicity, whenever feasible,[1] and (b) former English learners (ELs), in addition to current ELs and non-ELs, in order to describe student performance in more detailed ways.

The Framework also proposes to measure an expanded set of contextual variables via questionnaires and the increased use of digital process data to provide more information on student performance. The contextual variables are clustered by two sets of reader characteristics—(a) cognition and metacognition and (b) engagement and motivation—and by two sets of environmental characteristics: (a) reports of school and community resources and (b) reports of teacher, instructional, and classroom supports. Ultimately, the Framework envisions a reporting system that has enhanced capacity to assist researchers, educators, and policymakers in accessing and interpreting 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 (NAGB, 2009; 2019b). Building from this previous Framework and on digital innovations, updates include consideration of three additional, research-based concepts: (a) how social and cultural experiences shape learning and development, (b) how reading varies across disciplines, and (c) how the increasing use of digital and multimodal texts impacts reading performance.

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

1 The NAEP legislation requires the reporting of "information on special groups, including, whenever feasible, information collected, cross tabulated, compared, and reported by race, ethnicity, socioeconomic status, gender, disability, and limited English proficiency" [Sec. 303(b)(2)(G) of P.L. 107-110, as amended by P.L. 107-279].

	Current Framework and Assessment	2026 Framework Update
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 – nonmoving print, graphics, or images on screen ● Dynamic – navigation across modes (print, video, other) or nonlinear locations (hypertext link)	Digital texts ● Static – nonmoving 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 (UDEs)	Digitally based assessment as of 2017 includes tools and support features: ● Highlighting and note-taking ● 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 ● Multipart response frames ● Purpose statements ● Task characters (avatars that act as partners in simulated settings)	Types of UDEs and possible examples: ● Task-based UDEs – Highlighting and note-taking – Text-to-speech on Directions and Help screens – Zoom-in and selection of color schemes – Sequential directions and transitions for reading a collection of texts – Look-back buttons to return to relevant section of text – Graphic organizers – Item foreshadowing – Multipart response frames – Samples of student writing as examples

	Current Framework and Assessment	2026 Framework Update
	● Pop-up notes for definitions of vocabulary ● Resetting by providing correct response to answered questions ● Topic or passage introductions	– Resetting by providing correct response to answered questions ● 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 ● Informational UDEs – Text providing brief topic previews – Pop-up notes for definitions of obscure words or phrases that are not part of the Comprehension Target being tested
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 the following: ● disciplinary contexts ● socioeconomic status within race/ethnicity, whenever feasible ● former 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 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 United States. 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, presents the definition's origins in policy and scholarship on reading comprehension, and 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 process shaped by many factors, including readers' abilities to
>
> - 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 social and cultural contexts.

Key Terminology in the Definition

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

Contexts. Reading comprehension is shaped by how individuals interact with one another in classrooms, families, communities, and many other social and cultural experiences. 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. In addition to the common thread of cognition involved in reading across contexts, the process of comprehension is influenced by context (Scribner & Cole, 1981; Skerrett, 2020).

Readers. Each reader brings a unique and diverse repertoire of cognitive (including metacognitive), cultural, 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, 1st 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 United States may face a challenge when reading an unfamiliar text about the game of cricket in India but could use 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.

Texts. Texts are generated by authors to communicate to readers. 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 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 are more challenging for others who have no frame of reference or interest in the topic.

Activities. Activities include all the actions readers take 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 et al., 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).

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 knowledge 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. Cognitive models of reading describe the essential role of topic knowledge in text comprehension (Graesser et al., 1994; Kendeou et al., 2014; Kintsch, 1998; McCarthy & McNamara, 2021; Pearson & Cervetti, 2015; van den Broek et al., 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., Alexander et al., 1994; Anderson, 2019; Cervetti & Wright, 2020; Kendeou & O'Brien, 2016; Pearson et al., 1979; Taft & Leslie, 1985). 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.

The essential role of knowledge in reading comprehension is not controversial; however, there is far less consensus on how to build students' knowledge to support improved reading comprehension. As an assessment, NAEP provides information about what students have learned, not what they should be learning or how they should be learning it.

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 (C-I) models proposed by theorists such as Kintsch (1998), Perfetti (1999), and van den Broek (van den Broek et al., 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" (National Assessment Governing Board, 1992, 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 (RAND Reading Study Group, 2002), 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" (RAND Reading Study Group, 2002, pp. 11–12). The term sociocultural, within RAND and as one of the many factors that shapes reading comprehension, refers to the social and cultural features and practices of contexts, such as schools, homes, and communities, in which students learn to read and engage in reading (Lee, 2020; Pacheco, 2015, 2018; Skerrett, 2020). The 2009–2019 Framework also introduced the centrality of "using meaning as appropriate to type of text, purpose, and situation" (National Assessment Governing Board, 2019b, p. 3). The 2026 NAEP Reading Assessment will continue NAEP's longstanding focus on reading comprehension rather than on foundational skills or writing.

Updating the NAEP Reading Framework

The 2026 NAEP Reading Framework is updated to reflect three research-based developments that help ensure that the NAEP Reading Assessment remains a useful measure of reading comprehension. The first is how students' social and cultural experiences shape learning and development, including the learning and development of reading comprehension. The second is how reading varies across disciplines. The third regards the 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 et al., 2016; Heath, 1983; Lee, 2017; Scribner & Cole, 1981; Smagorinsky, 2001; Street, 1984), including disciplinary communities (Goldman et al., 2016; Moje, 2008). A 2018 report of the National Academies of Sciences, Engineering, and Medicine [NASEM] 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 learner's cultural, social, cognitive, and biological contexts" (NASEM, 2018, p. 33).

This NASEM finding is also reflected in other large-scale assessments. PIRLS, the international assessment of reading for 4th-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 & Martin, 2019, 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 (Goldman & Pellegrino, 2015; 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.

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. Cronbach 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 (Mislevy, 2019). One of these complex constructs is reading comprehension.

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, 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 et al., 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 for how well U.S. students comprehend what they read in texts and situations that more closely approximate reading practices in today's schools and society as a whole. By building on past frameworks and research traditions while embracing more recent developments in assessment, NAEP will continue to both lead and 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. To that end, five additional updated components are 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 and the natural variation that readers encounter in reading in home, school, community, and workplace settings. In this way, NAEP incorporates measurement of a wide range of factors that may influence reading comprehension.

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 (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. An additional target, ***Use and Apply,*** has been included to reflect the importance of applying comprehension to new situations.

Each Comprehension Target involves inferences that readers tend to find more or less challenging in general. 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 each Comprehension Target.

Locate and Recall. The first Comprehension 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. Processing in these ways helps readers build a literal understanding of what the text "says."

Items assessing the Locate and Recall 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 Comprehension 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. 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; Goldman 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 et al., 2018; Bråten et al., 2020; Meola, 2004; Ostenson, 2014; Wineburg, 1991; Wineburg & 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 Comprehension 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 et al., 2019; Pearson et al., 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 the reading purposes, the norms established by genre and disciplinary conventions, and the expectations about what is deemed appropriate and compelling to members of the target audience (Gee, 2001; Goldman et al., 2016; Moje, 2015).

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 generating 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, and motivations. 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 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 the readers' engagement with text and that influence how readers respond to and learn from the experience of reading. 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 literature texts, such as poetry, fictional and nonfictional narratives, and criticism, provides readers opportunities for enjoyment and for reflection and analysis around these themes, including how they shed light on readers'

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. Individuals read a variety of literature texts to appreciate elements of craft and to reflect on point of view, varied perspectives and experiences, and human dilemmas relevant to solving personal, social, and ethical problems. 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 and those of other authors 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 nonschool 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 understanding of 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 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 (National Council for the 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. Individuals read a variety of social studies texts to understand historical and contemporary issues and to solve community, national, and world problems. 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 in the creation of 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 4th, 8th, and 12th graders are likely to encounter in school and nonschool settings, as is described in more detail in 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. This means that there is overlap across disciplines regarding the kinds of texts used within disciplines.

Literature Texts. NAEP will draw on literature 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. Literature 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 nonliteral interpretations (Appleman, 2017; Lee et al., 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 that has the purpose of conveying 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 et al., 2015).

Social Studies Texts. NAEP will also sample from the varied forms of texts common to the social studies. Selection 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 classroom textbooks and interpretive books and articles about events, time periods, or people. 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 the texts were generated.

Digital Platform. As initiated in 2017, the 2026 NAEP Reading Assessment will continue to 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 literature, science, and social studies texts that reflect the digital environment, an environment that is different from the world of print on paper. 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.

As initiated by NAEP in 2017, 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 Smarter Balanced, 2020, and PARCC/New Meridian, 2019). 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 in measuring reading comprehension (Coiro, 2020; Fitzgerald et al., 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 and 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 question, "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 nonschool 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 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). 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 et al., 2002; Thompson et al., 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 (Johnstone, 2003; Johnstone, Altman, & Thurlow, 2006). 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.

The 2026 NAEP Reading Assessment uses three expanded categories of UDEs: task based, motivational, and informational.

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; de Jong, 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 that offer

models of approaches to tasks before students complete similar tasks independently (e.g., Sparks & Deane, 2015). Task-based UDEs may also include the kind of resetting feature described earlier, which has been part of NAEP since 2019.

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 (Dalton & 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 subtasks 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).

Informational UDEs. Informational UDEs are designed to maximize students' ability to engage with the content that is being tested by providing relevant context. Informational UDEs do not reduce the difficulty level of assessment items; rather, they provide orientations to topics, concepts, or obscure vocabulary that students may need to make meaning from text as they read (McNamara & Kintsch, 1996; Kintsch, 1998; van den Broek & Helder, 2017). Informational UDEs consist of brief passage introductions (e.g., a short description of the author or text) to provide context about what the student is reading and vocabulary pop-ups to offer on-demand definitions of obscure words that are not part of the content being assessed. Unless video, image, or other kinds of introductions are already part of an authentic source text, topic previews may take the form of written texts only.

UDEs and the NAEP Definition of Reading Comprehension. UDEs 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 look-back 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. Additional information about UDEs is provided in Chapter 3.

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, socioeconomic status,[2] and region of the country. 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

2 The Governing Board has traditionally complied with its legislative mandate to report on achievement by socioeconomic status by disaggregating results by free- and reduced-price lunch eligibility (in all grades) and parent education (in Grades 8 and 12). The Governing Board and the NCES are currently considering refinements of this approach that may affect the operationalization of socioeconomic status under the 2026 Framework.

according to students' experiences (e.g., 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 opportunities to gather information about students and their reading. Besides their demographic characteristics and language experiences, questionnaire items can also provide information about students' reading activities in school and community settings and their perceptions of the encouragement and instructional support they receive from peers, teachers, or community agency leaders. 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.

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

The digital format, which has been implemented since 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 briefly) they linger on different segments of the texts, the items, the UDEs, or the directions. These indicators can be used to help interpret performance differences 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, socioeconomic 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 performance.

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 and the NAEP Reading Assessment

	Features of the NAEP Definition of Reading Comprehension			
Assessment Components	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"; 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.

	Features of the NAEP Definition of Reading Comprehension			
Assessment Components	Contexts	Readers	Texts	Activities
Texts	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.
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 obscure words that are not the focus of the assessment items, and instructions on how to complete assessment tasks.	Increase broad access to texts by such means as providing definitions of obscure words not measured on the assessment and offering look-back buttons.	Provide information that clarifies the nature and order of tasks and expected responses.
Contextual Variables *Questionnaire Items*	Gather information about the contexts of readers' lives and experiences in and out of school.	Gather information about demographics, motivation, and in- and out-of-school reading practices.	Gather information about the amount and kinds of texts that readers encounter in and out of school settings.	Gather information about reading activities that readers commonly engage in at school and outside of school.

	Features of the NAEP Definition of Reading Comprehension			
Assessment Components	Contexts	Readers	Texts	Activities
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 (Pellegrino et al., 2001; AERA/APA/NCME, 2014) and were used in previous NAEP Reading Assessments (National Assessment Governing Board, 2019b). 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 leveraging digital assessment features, including item response formats, UDEs, and process data, in line with principles of validity, fairness, and inclusivity (AERA/APA/NCME, 2014).

Situating Readers Within Assessment Blocks

A block is the largest organizational unit for the 2026 NAEP Reading Assessment. In a typical NAEP Reading Assessment 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 C.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 United States are engaged in the respective contexts at Grades 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 percentages of Science and Social Studies increase.				
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. RSP activities entail developing understanding or comprehending text but are 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 Block-Specific Purposes and a Reader Role

Both RDU and RSP blocks also have their own 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 block-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.

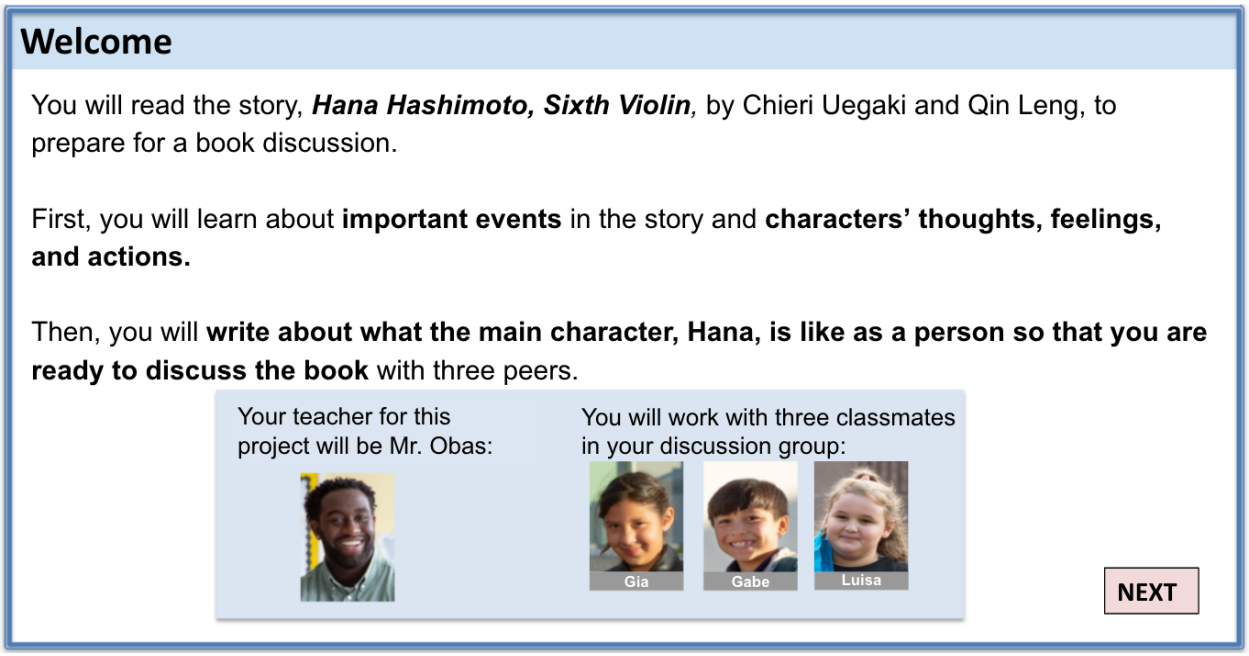
Reader roles are designed to reflect how readers typically engage with texts and each other in different contexts (e.g., 4th-grade classmates and a teacher in a literature circle discussion at school or a group of friends at home reacting to news about a local event in their town). Some blocks may ask 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*[3] by Chieri Uegaki and Qin Leng (2014) with three other 4th-grade student task characters (simulated avatar classmates). In addition to reading directions about the discussion goal, students are told they will read the story and respond to items situated in two purpose-driven tasks.

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.

3 Material from *Hana Hashimoto, Sixth Violin* written by Chieri Uegaki and illustrated by Qin Leng is used by permission of Kids Can Press Ltd., Toronto, Canada. Text © 2014 Chieri Uegaki. Illustrations © 2014 Qin Leng.

Exhibit 3.3. Block-Specific Purposes Presented at the Beginning of a Grade 4 Reading to Develop Understanding Block Using the Text *Hana Hashimoto, Sixth Violin* (a children's story), by Chieri Uegaki and Qin Leng



The photograph of Mr. Obas is sourced from https://images.all4ed.org/male-sixth-grade-math-teacher-with-protractor (photographer Allison Shelley for EDUimages). The photograph of Gia is sourced from https://images.all4ed.org/elementary-boy-with-backpack-and-girl-with-notebook/ (photographer Allison Shelley for EDUimages). The photograph of Gabe is sourced from https://images.all4ed.org/third-grade-boy-with-backpack-outside/. The photograph of Luisa is sourced from https://images.all4ed.org/fifth-grade-girl-mask-break (photographer Allison Shelley for EDUimages).

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 Assessment block has two or three 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 social media 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 the surface of Mars, and a dynamic image of two planets spinning around the sun. The item is intended to assess 4th graders' understanding of how to use embedded hyperlinks to locate and recall important information about the passage (see Mullis et al., 2017 for further detail).

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 both in and out of school (e.g., see Creer, 2018; Dobler & Azwell, 2007). The texts can represent one or more genres, modalities, or disciplines. See Exhibit A.1 in Appendix A for additional considerations for sampling text formats and modes. See Exhibit A.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 will continue to be selected using rigorous criteria that include the following:

- ***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 cohesion***. Are the texts well written and organized in ways that promote comprehension and learning? Do nonfiction 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 in the following sections with a number of principles and design considerations.

Authenticity. Most texts included in the NAEP Reading Assessment will be presented in their entirety, as students would typically encounter them. However, some texts may be excerpted from, for example, a novel, a play, 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. 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 an 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 A.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 A.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 et al., 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 text in NAEP Assessments built from the 2026 NAEP Reading Framework may continue to include video elements, consistent with previous NAEP Reading Assessments administered since 2017, 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 A.5 in Appendix A shows important 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.[4]

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, cohesion, and complexity and "considerateness" (including both qualitative and quantitative measures; e.g., conventional readability criteria, reader–text connections, language structures and vocabulary considerations; Armbruster & Anderson, 1985) and should 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 text structures and text features that are found in common across disciplinary contexts (see Exhibit A.6 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. Furthermore, texts often appear and are used in sets. Thus, it is important to determine the grade-appropriate

4 This document will be presented for Governing Board action later in 2021.

number of texts in a block 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.

Potential differences between traditional and digital texts are 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 multilayered digital text environments (Cho & Afflerbach, 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 seconds), 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 without placing students at a disadvantage based on their particular social and cultural context. 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. 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. In accordance with the Board's legislative mandate to "ensure that all items selected for use in the National Assessment are free from racial, cultural, gender, or regional bias," experts should represent diverse cultures and languages in order to identify texts that reflect the broad range of student readers' knowledge and experiences. The passages that are selected should themselves be drawn from texts that reflect a diverse range of cultures, regions, 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 the following:

- ***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 A.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 straightforward language and be 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 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). Items may require readers to draw on information contained in audio or visual features.
- ***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 in 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 be answerable by readers only drawing upon text-independent domain knowledge without even reading the passage.

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 Comprehension 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 A.7 in Appendix A presents guidelines for distributing items mapped to Comprehension Targets across grade levels 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 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 item 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 (i.e., the *type* of inference [a local, straightforward inference within a paragraph versus 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]). These factors 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 Assessment 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 A.8 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 obscure words of limited application across grade-appropriate texts 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 and 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 by having all test components 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).

To undertake these aims, the 2026 NAEP Reading Assessment is grounded in 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 et al., 2002; Thompson et al., 2004). See Exhibit 3.5 for an overview of UDA principles that are relevant to all assessments. The NAEP 2026 Reading Assessment employs UDA (Johnstone, Altman, Thurlow, & Thompson, 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 informed interpretations of assessment scores for all test takers. Such digital assessment features include the purposeful selection of item response formats, UDEs, 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" (Thompson 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., 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, Nonbiased 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 "nonbiased"). For example, if a passage contains a highly culturally situated term that might be more familiar to some subpopulations of test takers (e.g., to boys more than to girls), this might result 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.

Exhibit 3.6. Alignment of the 2026 NAEP Reading Assessment 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:*** The NAEP Reading Assessment 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 (see, for example, Lee, 2020; Solano-Flores & Nelson-Barber, 2001). 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 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. • Informational UDEs preview untested topic knowledge and provide definitions for obscure vocabulary not intended to be assessed. This maximizes the extent to which the assessment can measure the same intended construct 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 just as readers would be in 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 at measuring 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, Nonbiased Items	***Regular NAEP 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:*** 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 experience, knowledge, language use, and 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 engage with the 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. UDEs are "Universal Design Elements."

Item Response Formats

Central to the development of the 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; in some cases, they also drag them into the target fields.

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

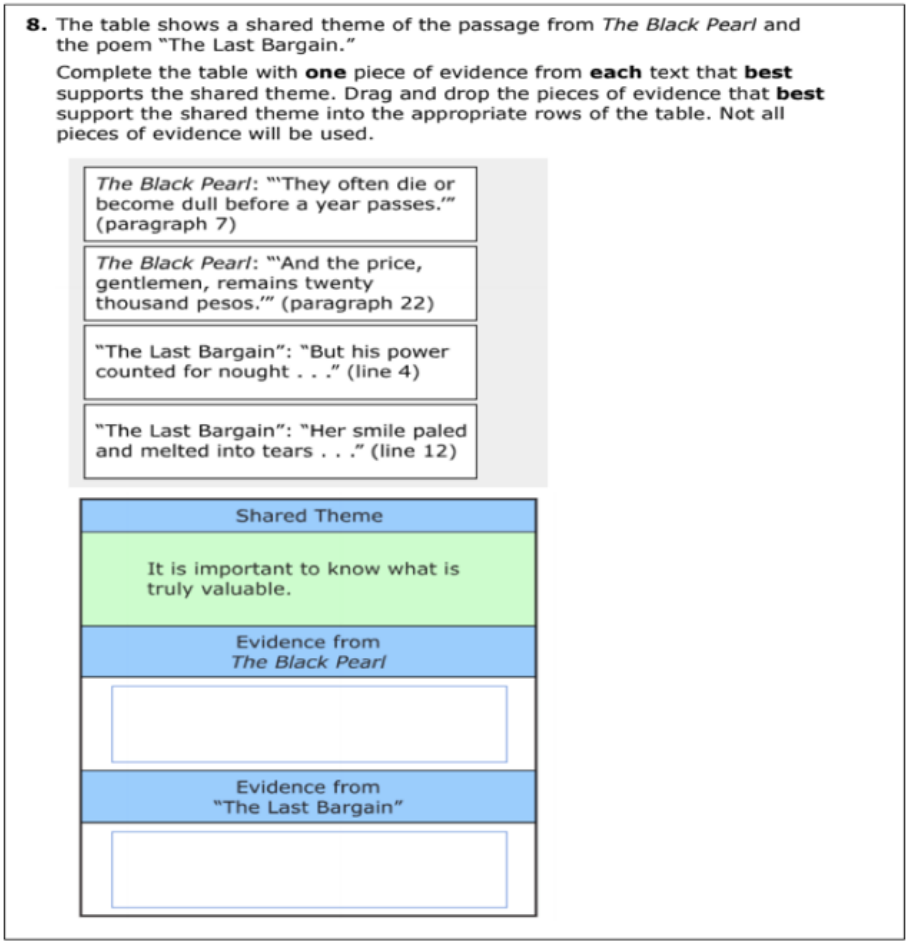
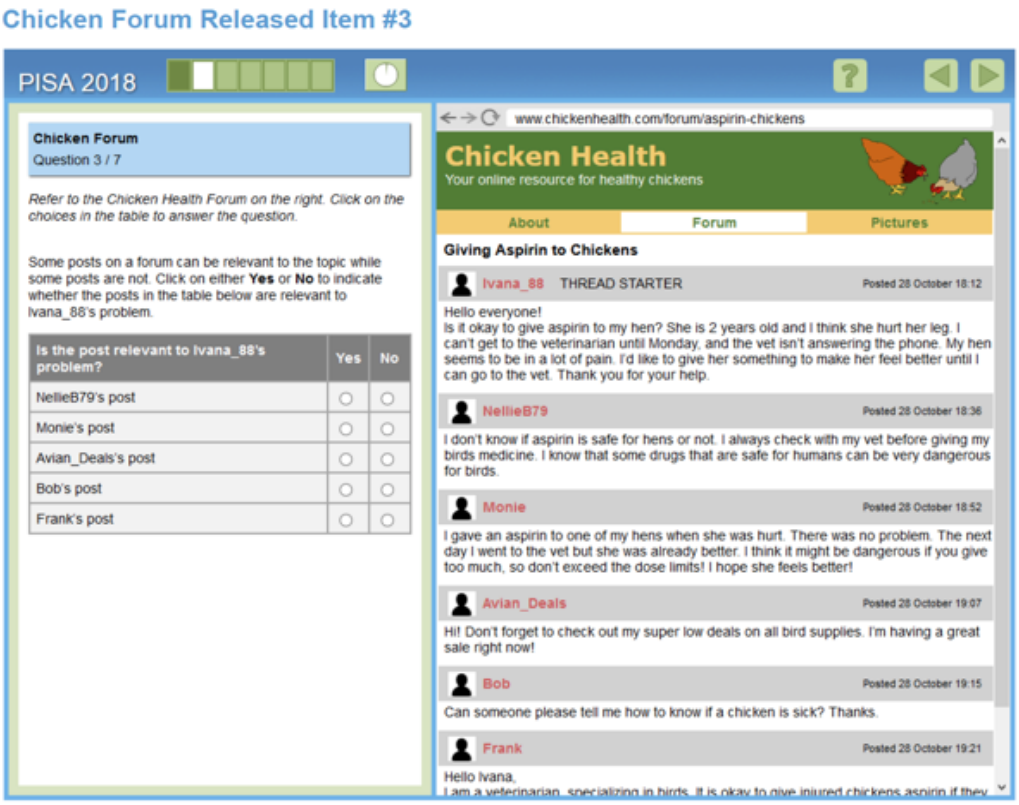


Exhibit 3.8. Example of Grid Response Format From PISA



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 the following:

- **short constructed response** – Students respond by entering a short text in a response box that consists of a phrase or a sentence or two. The fill-in-the-blank (FIB) item type is also considered a short constructed response format.
- **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 one or more choices that meet the condition stated in the stem of the item. Then they write a short explanation about their choices.

  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

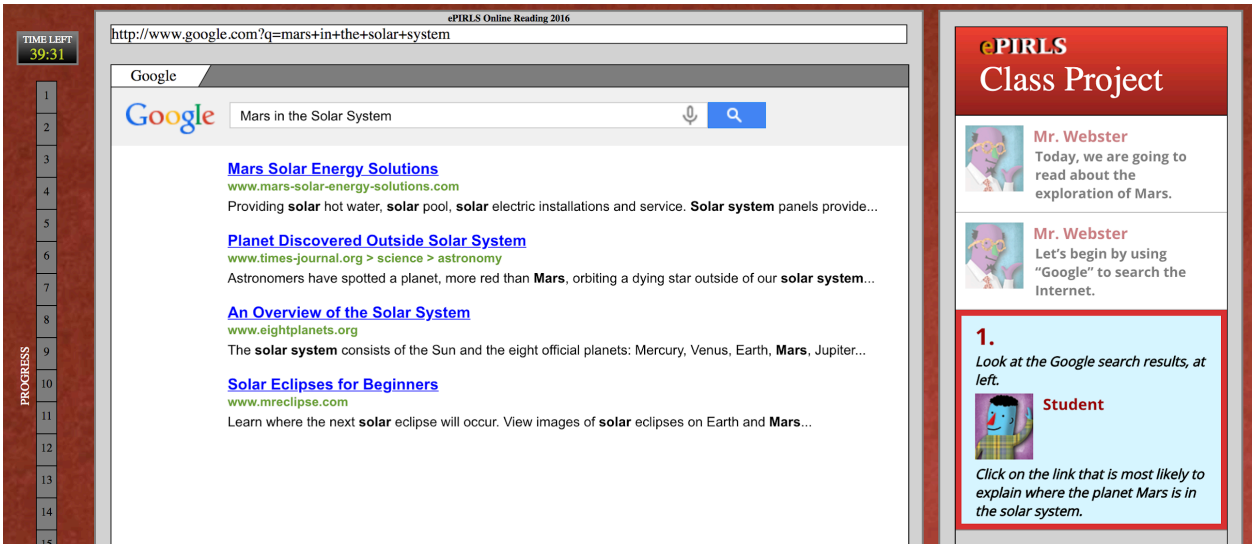
	Selected Response Items	Short Constructed Response Items	Extended Constructed Response Items
Grade 4	40–50%	40–45%	10–15%
Grade 8	40–50%	40–45%	10–15%
Grade 12	40–50%	40–45%	10–15%

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 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, selecting, and ordering);
- 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 UDA (Johnstone, Altman, Thurlow, & Thompson, 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 (nontesting) reading situations to improve the validity of the assessment.

All Readers Have Access to UDEs. UDEs, or the "built-in features of computer-based assessments," have been included in NAEP since the introduction of the digital platform in 2017 and are available for *all* students (NCES, 2021). 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 (NCES, 2019a). Examples of accommodations available on some assessments include extended time, options for responses in braille or American Sign Language, or having test items read aloud. UDA 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 the operational NAEP Reading Assessment (e.g., highlighters and look-back buttons) to reflect real-world experiences and how readers use technology. Amidst the use of these digital supports by all test takers, NAEP has effectively maintained the ability to capture trends over time (NCES, 2021). There are increasingly complex reading purposes and more dynamic texts in today's society. The 2026 NAEP Reading Framework includes three broad categories: task-based UDEs, motivational UDEs, and informational UDEs. The three categories of UDEs are designed to accomplish three different yet sometimes overlapping functions, as described below. 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 *Assessment and Item*

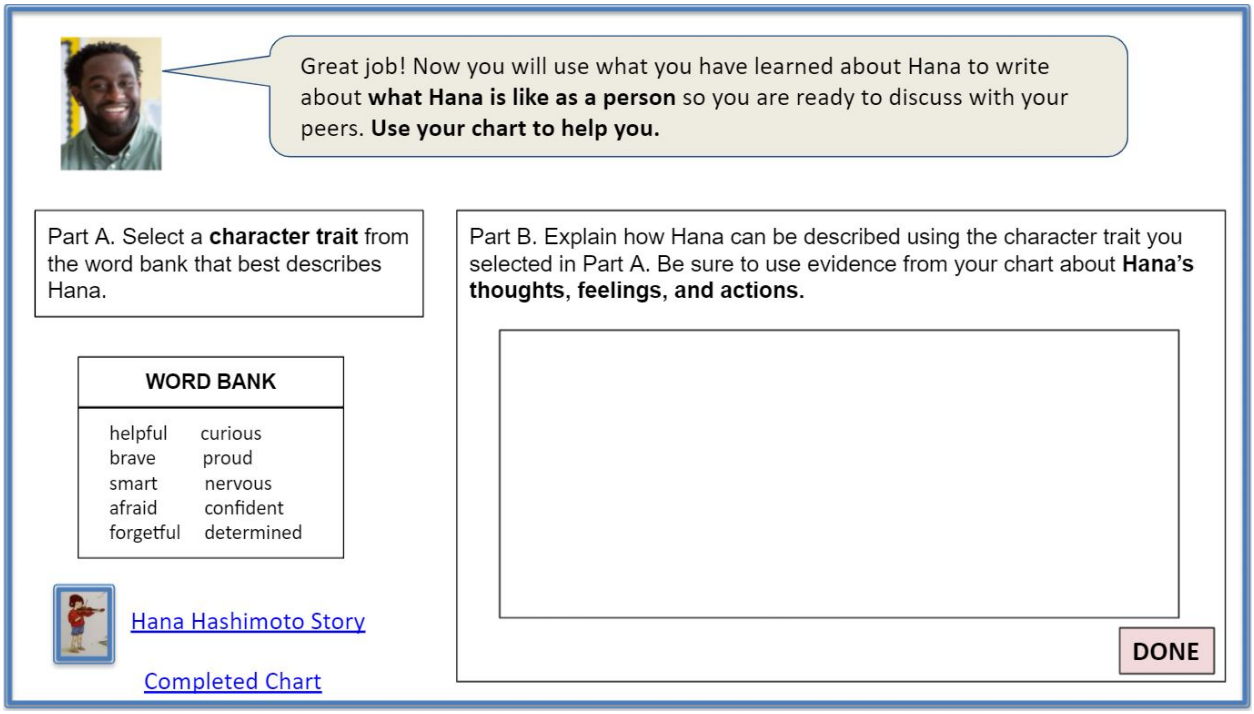
Specifications for the 2026 NAEP Reading Framework. Some examples of UDEs are presented in the next sections. See Appendix E for additional examples of UDEs.

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, 2015).

Exhibit 3.11 illustrates an example of an Analyze and Evaluate 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 a character in depth, 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, all students have an equivalent opportunity to focus more of their time and attention on the Use and Apply construct to be measured rather than on trying to generate a character trait word. 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 item in the relatively short period of time allotted by the assessment. This clarity of expectations also maximizes the likelihood that readers 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., 2002, p. 9). In this case, the word bank decreases construct irrelevance by providing a set of words from which test takers can *select*, rather than *generate*, a relevant character trait. 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.

Exhibit 3.11. 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 as Part of Their Constructed Response



The photograph of Mr. Obas is sourced from https://images.all4ed.org/male-sixth-grade-math-teacher-with-protractor (photographer Allison Shelley for EDUimages).

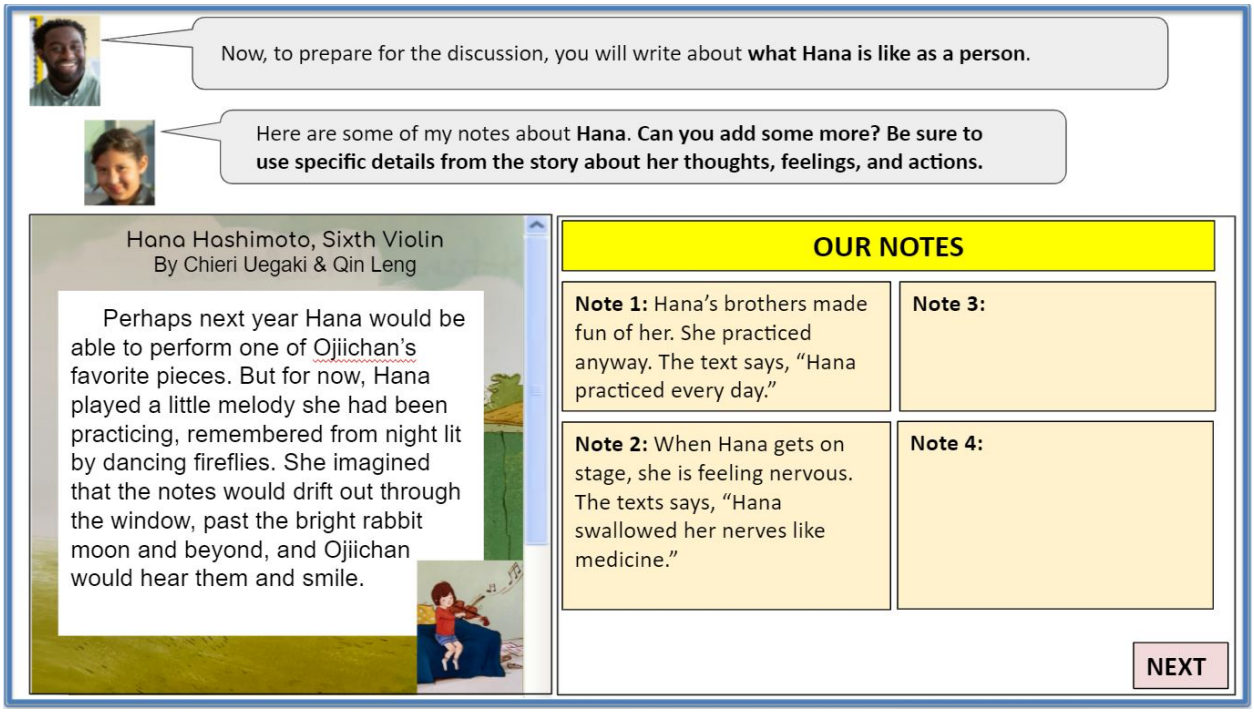
Motivational UDEs. In the 2026 NAEP Reading Assessment, motivational UDEs are designed to facilitate students' interest in assessment content and persistence with challenging tasks (Dalton & Proctor, 2008; Buehl, 2017; CAST, 2020; Guthrie & Klauda, 2015). Motivational UDEs might, for example, provide an engaging prereading preview that helps to generate a minimal amount of interest in an assessment block.

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 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 block.

Motivational UDEs may also maintain readers' interest by communicating explicit connections between the broader purpose for completing a block and the subtasks 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 3.12). 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 et al., 2011).

Exhibit 3.12. Teacher and Student Task Characters Remind the Reader of the Task Goal for the Second Task



The photograph of Mr. Obas is sourced from https://images.all4ed.org/male-sixth-grade-math-teacher-with-protractor (photographer Allison Shelley for EDUimages). The photograph of Gia is sourced from https://images.all4ed.org/elementary-boy-with-backpack-and-girl-with-notebook/ (photographer Allison Shelley for EDUimages).

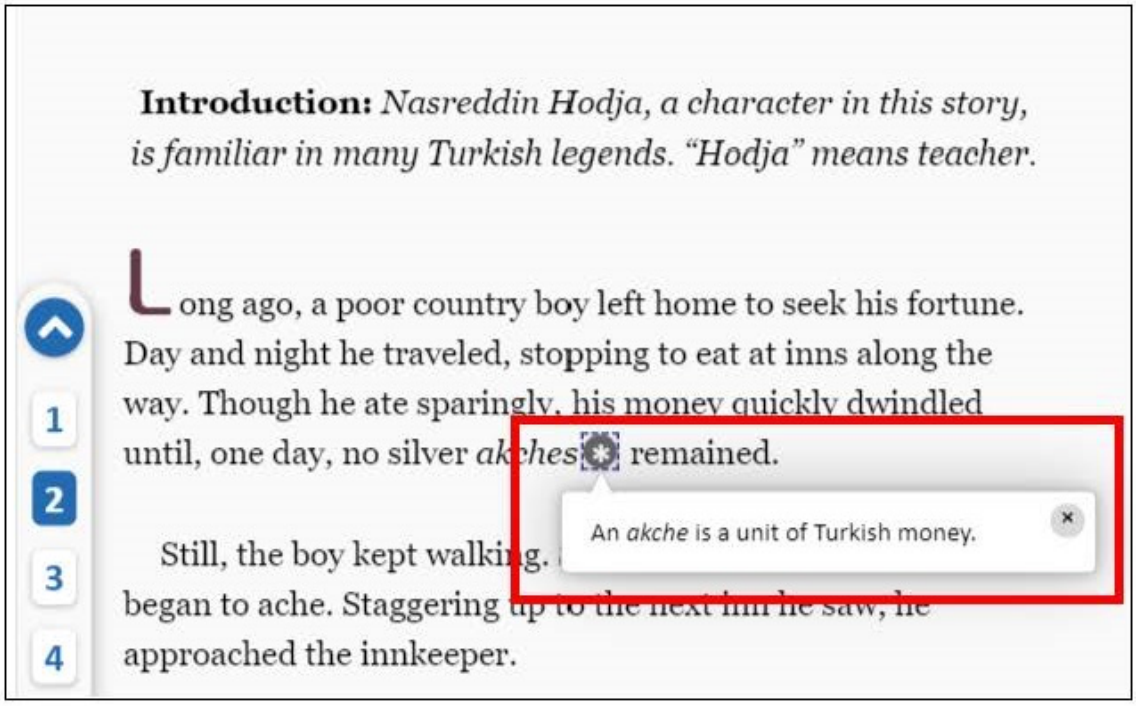
Informational UDEs. In the 2026 NAEP Reading Assessment, informational UDEs will provide two types of information: (a) topic previews in the form of short introductions to either the entire block or to a specific task and text, and (b) definitions or examples for obscure vocabulary unless a word is explicitly tested in a comprehension test item. Obscure vocabulary refers to words of very limited application, such as highly technical terms or non-English referents. In most cases, obscure words will already be defined in the authentic texts, but occasionally the assessment developer may consider whether an additional definition is necessary. Topic previews may take the form of written texts only, unless video, image, or other kinds of introductions are already part of an authentic source text. Topic previews should be offered as appropriate any time additional context about the author or text is needed to orient students to the passage. A determination must be made by assessment developers about whether a UDE is construct relevant. Finally, as noted in chapter 2, blocks without UDEs, including those without informational UDEs, are part of the current Assessment and will continue to exist in the 2026 NAEP Reading Assessment.

Importantly, informational 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 3.13, from a NAEP Grade 4 block, illustrates two informational UDEs. The first informational 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 informational UDE appears in the form of a vocabulary pop-up box defining the Turkish word "akche."

Exhibit 3.13. Example of Two Informational UDEs from NAEP's "Five Boiled Eggs" Block



Selecting appropriate locations for UDEs. Developers decide on appropriate locations at which to insert UDEs into each block of the assessment. Because some 2026 NAEP Reading Assessment 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 flowcharts, 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 variables relate to 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 and pages within passages, using hyperlinks, using the Next button to move through a block) (see Exhibit 3.14),
- 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 3.14. 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 students in Grades 4, 8, and 12. Additional research by NCES can inform decisions about the continued use of UDEs.

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. A reason for this is to ensure that design features yield their intended outcomes for all 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 interpretation 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.

Reporting Results

Historically, NAEP Reading Assessments have 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 other student characteristics (race/ethnicity, gender, English learner status, socioeconomic status, and disability status, i.e., supported by an Individualized Education Program), as required by law. 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, 2015) 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 Assessment 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 in Grades 3 through 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 Assessment 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. Exhibit 4.1 presents the generic policy definitions. 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:

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

To provide more nuanced reports and useful data to key stakeholders, the NAEP reporting system will

- disaggregate scores for demographic subgroups in greater detail to provide a more accurate and dynamic description of student performance,
- expand the number of categories for reporting the achievement of English learners to better reflect the variability of English language proficiency within this population, and
- provide information on research-based contextual variables (derived from demographic, questionnaire, and process data) that can contribute to more nuanced interpretations of group results.

Reporting Categories

The Framework reporting system described below provides opportunities to interpret findings from NAEP Reading Assessment 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, whenever feasible, 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 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 Assessment 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, as required by the law: gender, race/ethnicity, SES, disability status, and English learner 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 SES within the student subgroup of race/ethnicity. In the NAEP Reading Assessment, as in other large-scale assessments, lower levels of achievement are historically correlated with poverty. Disaggregating results by SES within subgroups will reveal subgroup differences in reading achievement that are associated with SES. At the same time, the success of many schools in supporting high levels of achievement among students from low-SES backgrounds suggests that SES alone does not offer a sufficient explanation for reading performance and that additional contextual variables are crucial to better understand variability in reading (Mullis & Martin, 2019; OECD, 2019). Enhanced reporting can help policymakers and stakeholders better understand reading performances in context. For example, these data may allow policymakers to consider how access to resources that support rich literacy opportunities 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

ELs are defined by NAEP as students "who are in the process of acquiring English language skills and knowledge" (NCES, 2019b). 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; thus, 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, et al., 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:

- *Current English Learners* – students designated as English learners at the time of the assessment
- *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
- *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. Questions are intended to be non-intrusive; free from bias; secular, neutral, and non-ideological; and they 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, SES, 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 et al., 2000) to understand reading achievement (Solano-Flores, 2011; Solano-Flores & Nelson-Barber, 2001). Contextual variables are measurable, and some are also malleable (i.e., they can be influenced). These include *reader*

characteristics (e.g., students' self reports about engagement and motivation, knowledge, agency, effort, and interest in reading) 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 by researchers to try to predict or account for variance in the outcome of interest: reading comprehension scores on NAEP. The 2026 NAEP Reading Framework's emphasis on 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 measuring 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 will support efforts by researchers, educators, and policymakers to interpret students' differential performance on the NAEP Reading Assessment.

The 2026 NAEP Reading Framework 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 is a complex process shaped by many factors. Factors may include how social and cultural practice influences how readers approach, engage with, and make meaning from texts (Mislevy, 2019; Moje et al., 2020; Moje & Luke, 2009; NASEM, 2019; Pacheco, 2015, 2018). Readers' values, beliefs, experiences, and ways of communicating and thinking are all shaped by their everyday experiences (Lee, 2007, 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; 2020; Phillips Galloway et al., 2020).

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, sociodemographic student characteristics, and student interests and experiences. Taken together, the network of contextual variables is intended to (a) correlate with performance on the outcome measure of reading comprehension, (b) be malleable (that is, influenced by differences in school and community settings), and (c) comply with the provision of the NAEP law that prohibits assessment of personal or family beliefs and attitudes. 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 et al., 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 & Schrauben 1992; Eccles et al., 2005; Valencia et al., 2014).

Reader characteristic data to be collected from questionnaires and process data include the following:

Cognition and Metacognition

- **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, 2018).
- **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).
- **Topical knowledge** refers to students' use of their preexisting knowledge of the reading topic to enable them to understand text information and construct new knowledge (O'Reilly et al., 2019).

Engagement and Motivation

- **Volume of reading** refers to the amount of reading a student does for personal interest, pleasure, or learning (Schaffner et al., 2013).
- **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).
- **Motivations for reading** refer to students' attention, effort, interest, and value for reading a particular text with a unique set of tasks and questions related to it (Educational Testing Service, 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, 2020). 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 (e.g., 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, 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:

Self-Reports of School and Community Resources

- **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 et al., 1986).
- **Belonging in school** refers to the extent to which students perceive themselves to be accepted members of the school community (Faircloth & Hamm, 2005).
- **Participation in out-of-school reading/literacy activities** refers to the degree to which students have access to resources (e.g., books, computers, media centers, camps, community organizations) that utilize literacy for enjoyment, communication, learning, and the pursuit of a variety of activities (Bowen et al., 2002).

Self-Reports of Teacher, Instructional, and Classroom Supports

- **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 et al., 2020).
- **Teacher support for motivation** refers to the degree to which students perceive their teacher(s) to support their interests and reading goals (Wigfield & Wentzel, 2007).
- **Teacher support for students' background experiences** refers to the students' perceptions that their teacher(s) recognizes and uses students' cultural, language, and social knowledge during reading instruction (Shin et al., 2007).
- **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 policymakers and stakeholders to gain more actionable insights regarding the variables' potential correlations with 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 Assessment performance (Guthrie et al., 2012). Students' positive perceptions of their teachers' support and the classroom climate could also be associated with higher NAEP Reading Assessment 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 (Durlak et al., 2015; Guthrie & Klauda, 2016; Katz et al., 2019; Shin et al., 2007; Skerret, 2020), recognizing that the causal nature of these variables cannot be demonstrated with NAEP cross-sectional data.

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 are 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 & von Davier, 2019. Exhibit 4.2 provides a list of variables along with their sources 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			
Reports of School and Community Resources			
School social support	✓	✓	
Belonging in school	✓	✓	
Participation in out-of-school reading/literacy activities	✓		
Reports 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 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 at both the national and state levels. When the NAEP Report Card for Reading is issued every two years, policymakers and the public pay attention, particularly to trend data. Yet, NAEP results have also been subject to misinterpretation (Linn & Dunbar, 1992; Jaeger, 2003; NASEM, 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 for either reader characteristics (e.g., student motivation) or environmental characteristics (e.g., opportunity to learn factors). 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.** The Assessment reflects the field's evolving understanding of reading comprehension, cognitive processes, and the changing nature of reading demands in today's society (American Educational Research Association, American Psychological Association, and National Council of Measurement in Education, 2014; International Test Commission, 2019; International Reading Association & National Council of Teachers of English Joint Task Force on Assessment, 2010). Importantly, it optimizes readers' opportunities to demonstrate reading comprehension that reflect the changing demands of our increasingly complex world (Mislevy, 2016; NASEM, 2018). Reframing and expanding the reporting system is as important as the assessment construct itself in enhancing the appropriateness of inferences based on NAEP results.
2. **Revise Questionnaires**. To increase the capacity to examine the relationships between readers and their environments, NAEP seeks to revise and refresh questions. 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 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, most 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, 2019 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 that, for 4th graders in the nation, breaks down the performance of low- versus high-SES students on the cognitive targets of Locate and Recall, Integrate and Interpret, and Analyze 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 reports beyond the standard ones offered by the Nation's Report Card.

Conclusion

Reading comprehension performances vary depending on the combination of individual and contextual factors at the time of the assessment. Thus, NAEP Reading Assessment 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 provides opportunities to examine malleable contextual variables that may be correlated with comprehension scores. The identification of malleable factors by the 2026 NAEP Reading Assessment reporting system also provides information that may eventually lead to policies and practices that improve students' reading comprehension instruction and performance. Moreover, the disaggregation of reporting that examines heterogeneity within groups (e.g., race/ethnicity, SES, gender, English learners) will also be important. Efforts to disaggregate scores beyond what has been done in past iterations of the NAEP Reading Assessment provide opportunities for further understanding 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. 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.

Accessible: 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 Grades 4, 8, and 12)

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: nonstatic digital format; involves movement or navigation across modes (e.g., print, images, 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 is 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 the literal meaning of words (e.g., metaphors, hyperbole, personification,simile)

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

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

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 competencies 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

Informational UDE: a type of 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 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: 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 3 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

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, 1998); the level where readers make connections between text ideas and their own knowledge

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: nonmoving print, graphics, or images

Student identity: a student's evolving view of self in a given social context influenced by their 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 UDE 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, conflict–resolution)

Text-based inference: logical conclusions or assumptions based on information stated in the text

Topic knowledge: understanding of 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 of Assessment (UDA): principles for creating and administering evaluations or tests so they are 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: an informational 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 A.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 A.1. 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 A.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 A.2. Illustrative Examples of Texts and Other Media Across Single Static and Dynamic Texts and Multilayered Digital Text Environments

<table>
<tr><th>SINGLE STATIC TEXT</th><th>SINGLE DYNAMIC TEXT</th></tr>
<tr><td>Single static genres and forms of continuous prose, noncontinuous prose, and everyday reading materials from which designers might sample as readers read to engage in literature, science, or social studies.</td><td>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</td></tr>
<tr><th colspan="2">MULTILAYERED DIGITAL TEXT ENVIRONMENT</th></tr>
<tr><td>• Augmented reality text
• Blog
• Database
• Digital creation/composition tool
• Dynamic simulation
• Email
• Interactive model</td><td>• Google document or Google folder
• Role-play simulation
• Search engine
• Social media (e.g., Facebook, Instagram, Twitter)
• Threaded discussion
• Webpage or website</td></tr>
</table>

Exhibit A.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 A.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: • should be 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 regarding form and purpose • vetted for accuracy, authenticity, and appropriateness by experts in the discipline, NCES's text selection panel, and the Assessment Development Committee • no use of items asking students to evaluate source credibility of such commissioned texts • meets the same complexity and other criteria for text selection as all texts for the NAEP Reading Assessment

Exhibit A.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 A.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 A.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 A.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 • Coming-of-age stories • Satires • Science fiction • Magical realism • Fantasy • Comic books • Graphic novels • Manga • Fan fiction Poetry • Haiku, sonnet, ballad, dirge, epic, etc. Related Nonfiction • Memoirs • Biographies and autobiographies • Literary analyses • Reviews and recommendations • Author profiles	• Plot types • Character types • Narrative elements (character, setting, plot, conflict, rising action, climax, resolution) • Figurative language (symbolism, imagery, simile, metaphor, personification, satire) • Point of view • Theme • Soliloquy, dialogue, and monologue • Diction, word choice • Repetition, exaggeration • Flashback • Foreshadowing • Mood, tone, irony, paradox, and sarcasm • Visual and graphical elements such as illustrations and photographs • Multimodal elements such as narrative soundscapes • Description • Narrative and expository text structures

Science	• Science reports • Press releases • Science news and features • Science magazine articles • Reference materials and field guides • Discovery narratives • 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 • Procedures • Published research articles • Personal communications	• 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) • Embedded 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 (era, time, sequence, tense) indicators • 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 simulation, time-lapse photography, and animations
Social Studies	• Historical and contemporary documents such as newspaper articles, editorials, political cartoons, broadsides, blogs, census data, diaries, letters, speeches, inventories and records of sale, advertisements, archival documents • 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: “this stance of brinkmanship...”) • Rhetorical markers of persuasion • Historical expressions and terminology • 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

	• Interpretive explanations 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 (era, time, sequence, and tense) indicators

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 A.6 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 A.6. Text Structures and Features Within and Across Single Static and Dynamic Texts and Multilayered Digital Text Environments

SINGLE STATIC TEXT	SINGLE DYNAMIC TEXT
Text structures are comparable to those in a printed format for texts designed to inform, entertain, and/or persuade. **Text features** may include visual media elements in a single text comparable to those in a printed format that convey meaning primarily through static words, numbers, and/or visual graphics, such as those in a still photograph, diagram, or table.	**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). **Text features** include one or more **multimodal** elements (words, moving images, animations, color, music, and sound) embedded in a single text or other media element

MULTILAYERED DIGITAL TEXT ENVIRONMENT
In multilayered digital text environments (Cho & Afflerbach, 2017)**, text structures** may include one or more static or dynamic texts, with a strong likelihood of containing |nonlinear elements 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 that 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 A.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. Therefore, 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 A.7. Distribution of Cognitive Comprehension Targets Across Grade Level and Broad Purposes

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 Integrate and Interpret, 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. Thus, 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	15–40%
Integrate and Interpret	10–40%

Analyze and Evaluate	10–25%
Use and Apply	0–30%
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	10–25%
Integrate and Interpret	20–35%
Analyze and Evaluate	20–35%
Use and Apply	0–30%
Grade 12	
Locate and Recall	10–25%
Integrate and Interpret	25–35%
Analyze and Evaluate	25–40%
Use and Apply	0–45%

Exhibit A.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.

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

Language Structures & Vocabulary Included/Excluded 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 15%–20% of items in any assessment block will assess passage-relevant Language Structures and Vocabulary knowledge while concurrently measuring a specific comprehension process.

The NAEP Reading Framework 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 Framework 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 the NAEP Reading Assessment. 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 2026 NAEP Reading Framework does not provide reporting ALDs; those will be constructed using empirical data during a later stage in the NAEP cycle (i.e., an operational 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 (National Assessment Governing Board, 2018b).

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 2026 Framework has identified three (literature, science, and social studies). In the ALDs below, all three disciplinary contexts are described within each performance level.

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 their 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:* (a) Locate and Recall, (b) Integrate and Interpret, (c) Analyze and Evaluate, and (d) 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

Reading activities in an assessment block are situated within a disciplinary context as well as a broad reading purpose. Each assessment block is designated as having one of two *broad* purposes: Reading to Develop Understanding or Reading to Solve a Problem. RDU blocks ask students to *read and comprehend deeply* (analyzing, inferencing, interpreting, and critiquing) in or across disciplinary contexts. By contrast, RSP blocks ask students to demonstrate understanding across multiple texts and related perspectives in order to solve a problem. RSP 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 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 in the real world occurs. The subsections below describe how specific reading purposes map to disciplinary contexts.

Literature 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 literature 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 flowcharts, 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 when reading and viewing science and health news. NAEP simulates these Contexts of Reading to Engage in Science by providing test takers 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:

- 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 takers 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, recall, and/or record specific pieces of information; identify relationships between explicitly stated pieces of information; make simple inferences and interpretations in static, dynamic, and multimodal texts; determine the accuracy of summaries; and show understanding of vocabulary in the disciplinary contexts.

When engaged in reading literature texts such as fiction, drama, film, poetry, and literary nonfiction, 4th-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 (e.g., text features) among literature subgenres appearing in specific task texts; and show understanding of vocabulary and simple figurative language. Readers should be able to determine the accuracy of 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), 4th-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 form an evidence-based opinion about a text. 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 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, 4th-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 form an evidence-based opinion about a text. Readers should be able to describe text structures as they pertain to the presentation of content in a specific text and compare and contrast explicit information found in a firsthand and secondhand account of the same event or topic. Readers should be able to determine the accuracy of a simple summary of a text and integrate information from lower complexity sources to apply to a new context.

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 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 literature texts such as fiction, drama, film, poetry, and literary nonfiction, 4th-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 information from a multimedia source contributes to understanding of a printed text and to show understanding of nuances in word meaning. Readers should be able to apply understanding of a character to an interpretation of another character's point of view.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including investigations), 4th-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 based upon content in the text and 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 determine missing steps in a procedure (e.g., a simple investigation, craft-making related to a scientific concept) based on 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, 4th-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 adopt the persona of a historical figure when applying information learned to a new context.

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 judgments based upon evidence within and across static, dynamic, and multimodal texts.

When engaged in reading literature texts such as fiction, drama, film, poetry, and literary nonfiction, 4th-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 common human experiences. Readers should be able to apply knowledge acquired about the 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), 4th-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 based upon content in the text, interpret an author's point of view or purpose, and 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, 4th-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 based upon content in the text, interpret an author's point of view or purpose, and argue for or against a particular interpretation. Readers should be able to use and apply information from texts in a new context, such as proposing a caption for an illustration or cartoon, or to create a set of recommendations.

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 based upon content in the text; determine the accuracy of summaries; analyze word choice; and show understanding of vocabulary in the disciplinary contexts.

When engaged in reading literature texts such as fiction, drama, film, poetry, and literary nonfiction, 8th-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 should be able to determine the accuracy of a summary of a text and construct an argument 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), 8th-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; and show understanding of academic and domain-specific vocabulary, key terms, and symbols. Readers 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, 8th-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. They should be able to demonstrate an understanding of the purpose/function of specified text features (e.g., introductions, sidebars, headings, illustrations, charts). 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 should be able to use information from multiple sources to apply to a new context.

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 literature texts such as fiction, drama, film, poetry, and literary nonfiction, 8th-grade readers performing at the *NAEP Proficient* level should be able to 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 information from a multimedia source contributes to understanding of a printed text and how text structure contributes to meaning and style. They should be able to analyze how word choice impacts a text's meaning and tone. Readers should be able to apply analysis of multiple texts to an explanation of 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), 8th-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 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, 8th-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. Readers should be able to present an argument 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 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 literature texts such as fiction, drama, film, poetry, and literary nonfiction, 8th-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. They should be able to determine how the text structure contributes to the development of theme, setting, or plot. Readers should be able to describe how a story might change if written from the perspective of another character.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts, 8th-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 should be able to construct an argument or explanation 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, 8th-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 should be able to trace and connect 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 based upon content in the text; determine the accuracy of summaries; analyze word choice; and show understanding of vocabulary in the disciplinary contexts.

When engaged in reading literature texts such as fiction, drama, film, poetry, and literary nonfiction, 12th-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 should be able to determine the accuracy of a summary of a text and apply a common theme or central idea culled from multiple texts to common human experiences.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), 12th-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 a text structures information to serve an author's purpose and help readers organize their thinking. 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 apply findings described in a text to a new context or situation.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, 12th-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. They should be able to use information from multiple sources to construct an explanation or argument.

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 literature texts such as fiction, drama, film, poetry, and literary nonfiction, 12th-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 should be able to analyze how word choice impacts a text's meaning and tone. Readers should be able to present an opinion regarding a universal problem that is elicited from an analysis of the text.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts, 12th-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, the analysis including providing an explanation or describing a procedure and 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 should be able to construct an argument or an explanation 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, 12th-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 historical documents and evaluate the effectiveness of the structure in the text's exposition or argument. Readers should be able to evaluate multiple sources of information presented in different media or formats (visually or in words) in order to construct an argument with evidence to 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 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 literature texts such as fiction, drama, film, poetry, and literary nonfiction, 12th-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) compared to the source text. Readers should be able to use or apply information gained from a literary text or a poem to analyze a new text.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), 12th-grade readers performing at the *NAEP Advanced* level should be able to delineate and evaluate the argument, claims, and reasoning in a text and analyze 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 should be able to construct an argument, explanation, or recommendation that requires the application of scientific content from a text.

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, 12th-grade readers performing at the *NAEP Advanced* level should be able to delineate and evaluate the 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 should be able to construct an argument, explanation, or recommendation that utilizes an understanding of legal and ethical principles to address a societal matter of debate (e.g., indigenous peoples' land rights).

This 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 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 C.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.

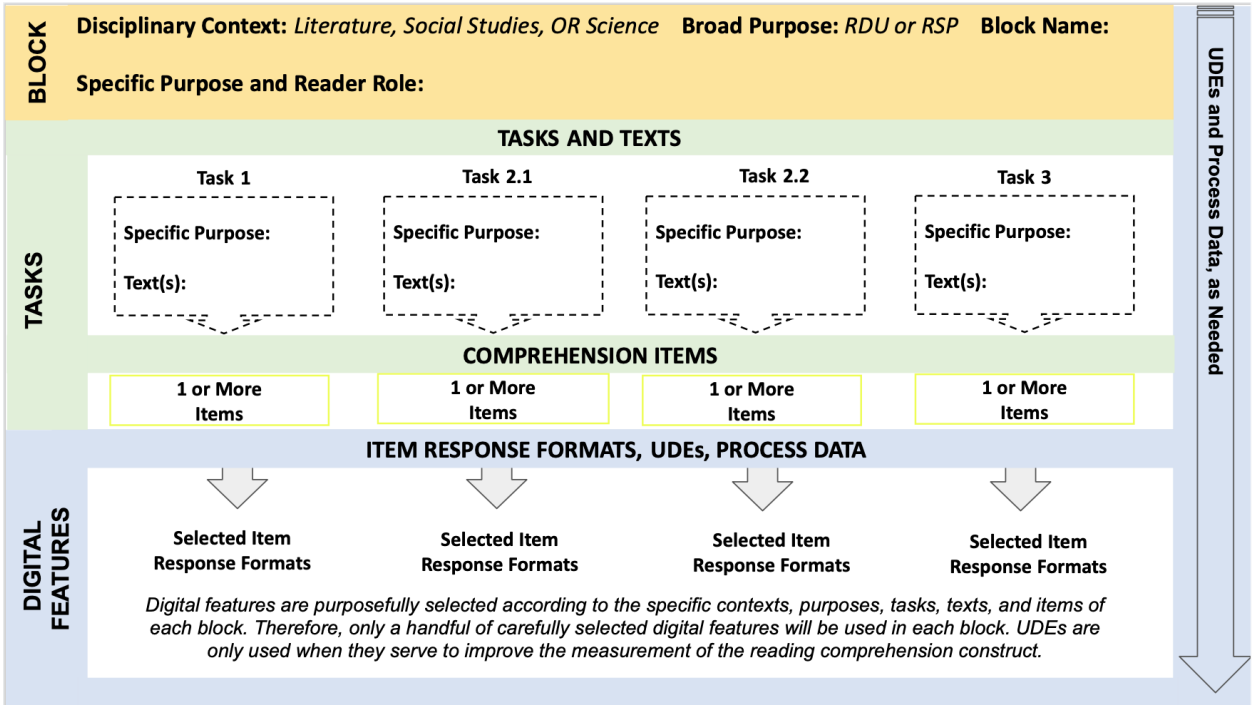
First, students learn what disciplinary context and broad purpose they are working in, and then they learn the specific purpose and their role. Second, students are given a text or texts to read and tasks to work on as they read that text. As students engage with the texts and tasks, they complete comprehension items, which are situated within the tasks, as illustrated in Exhibit C.1. Third, design features such as item formats, UDEs, and process data are used to leverage the digital assessment environment to measure 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 these, 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 C.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 subtasks, 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 C.1. This exhibit provides one sample approach to an assessment block; other approaches are possible that would have variations in the components (e.g., the number of tasks and texts). 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 C.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 C.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 C.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 C.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 toward 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, with less probability of readers moving back and forth		Purpose-driven tasks are situated in line with disciplinary context, but tasks are more tightly structured so that one task builds on the previous; there is more probability that tasks are interdependent; there may be

	across items within tasks; there is less need for resetting. There is a less involved culminating task or no culminating task. The task is not necessarily a determinant of all items in the block.		more need for resetting. There is a more involved culminating task at the end of an activity that directly addresses the question or problem; it is a 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 to do so 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. There are 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. There are more dynamic item formats to support more complex/multilayered tasks and items.
Universal Design Elements (UDEs)	There are fewer cumulative reading purposes that may require UDEs for knowledge or motivation and potentially		There are more cumulative reading purposes that may require UDEs for knowledge or motivation and potentially greater need for task-based UDEs.

	less need for task-based UDEs.		
Process Data	There are 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.		There are 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:

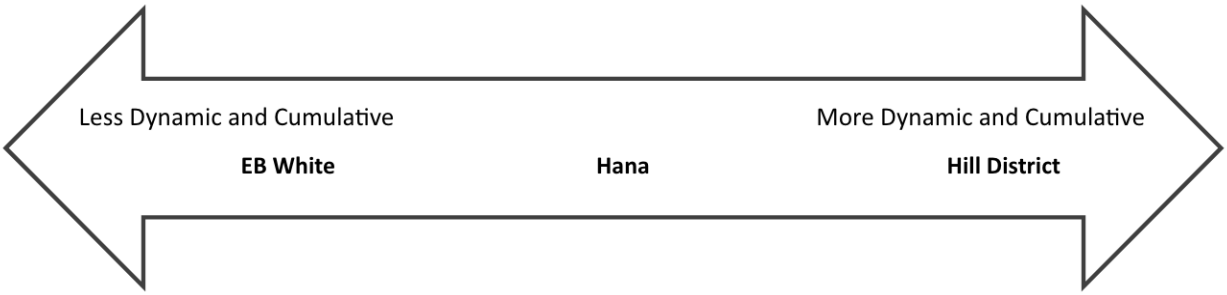
1. Students deserve to know the tasks that lie ahead of them in the block. Guidance in the form of task-based UDEs is essential.
    a. Both block-specific purpose and reader role need to be made apparent at the outset of a block.
    b. Students should be reminded of purpose and role as appropriate 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 are 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 (Sabatini et al., 2020) and in the current 2019 operational NAEP SBT blocks offer an existence proof that these guidance features are manageable by students in Grades 4, 8, and 12. 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[5]) illustrates a block developed for the broad purpose of RDU. 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 RSP. And the third (labeled *E. B. 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. Referring to the underlying continuum of variation for assessment components within blocks as detailed in Exhibit C.2 above, these three block sketches are situated on three hypothetical points along that continuum, as illustrated in Exhibit C.3.

Exhibit C.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, 4th-grade readers read and interpret story excerpts from the short story, "Hana Hashimoto," by Chieri Uegaki and Qin Leng, in preparation for a book discussion with three peers. First, students are asked to read to develop an understanding of the characters, key events,

5 Material from *Hana Hashimoto, Sixth Violin* written by Chieri Uegaki and illustrated by Qin Leng is used by permission of Kids Can Press Ltd., Toronto, Canada. Text © 2014 Chieri Uegaki. Illustrations © 2014 Qin Leng.

and authors' craft. Second, they apply their insights to describe what Hana is like as a person so that they are ready to contribute to the discussion.

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 an RSP block with texts situated in a social studies context. In this block, students engage in more 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 multiple texts to solve a problem.

E. B. White illustrates a second RDU block, but it is 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:

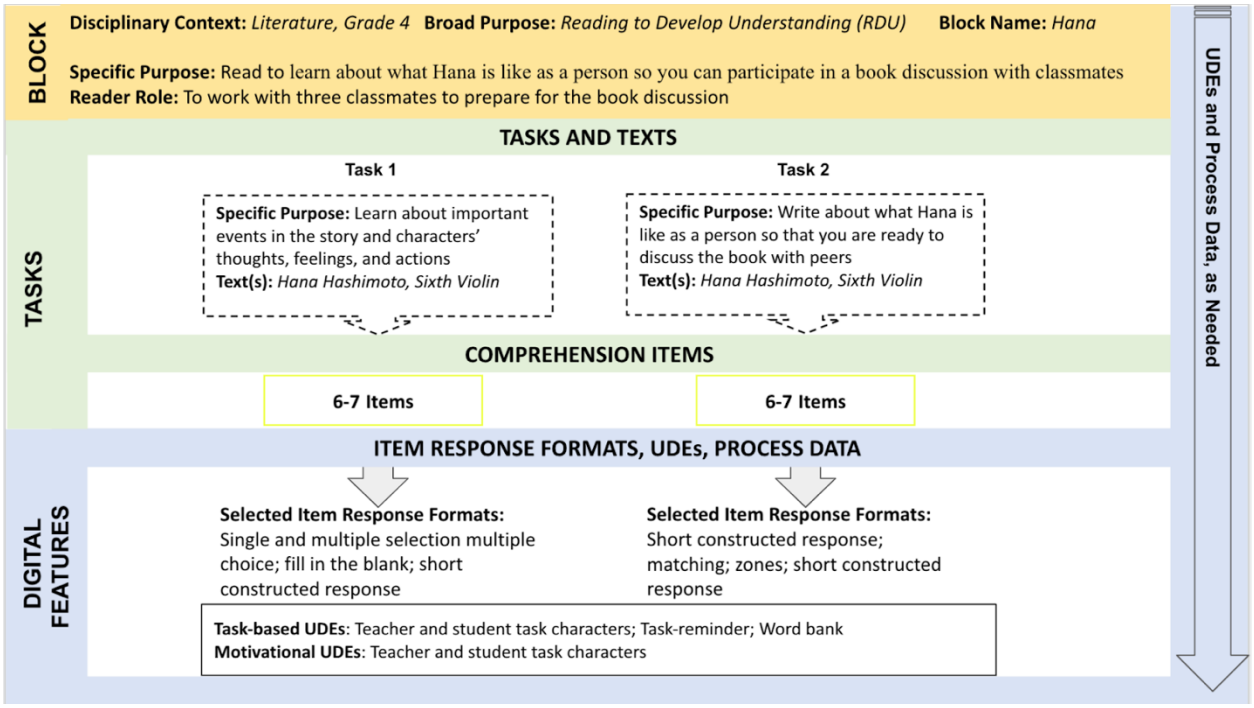
- The purpose of these block sketches is to help readers of this 2026 Reading Framework develop an understanding of the range of comprehension activity and assessment components students might experience when they participate in the NAEP Reading Assessment.
- None of the examples is complete in the sense 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 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 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 A.7), 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 *E. B. White* block is otherwise a classic RDU block but lends itself to a Use and Apply culminating task (which is more characteristic of RSP blocks).
- The *E. B. 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 C.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 C.4. Block Design for *Hana*



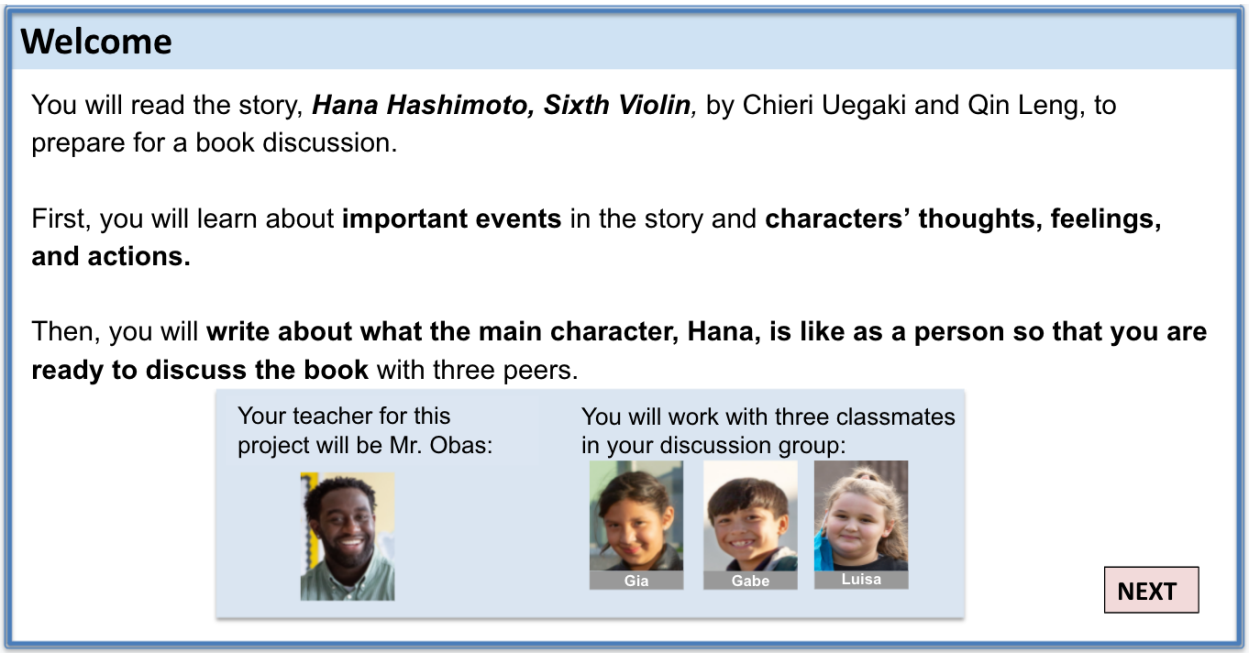
Block Components (Disciplinary Context, Purposes, and Reader Role). This block is designed to assess how 4th-grade readers develop understanding within a single printed text in a literature context. In this block, readers identify important events in the story and analyze how characters' thoughts, feelings, and actions describe the kind of people they are. Then, readers use and apply what they have learned to form an overall interpretation of the main character, Hana. They choose a character trait from a word bank and then 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 C.5), readers are 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 with three 4th-grade classmates (represented in the assessment by task characters Gia, Gabe, and Luisa). They are also introduced to their teacher for the project (represented by the task character Mr. Obas).

Then, a **task-based UDE** in the form of two statements informs students what tasks will be expected of them. Here, students are told that, to prepare for the book discussion, they will read the story and (a) learn about important events in the story and characters' thoughts, feelings, and actions and (b) use what they have learned about Hana to describe what she is like as a person. **Motivational UDEs** (here, student and teacher avatars) serve to motivate readers to engage with the block.

Exhibit C.5. Specific Purpose, Reader Role, and 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



Throughout Appendix C, the photograph of Mr. Obas is sourced from https://images.all4ed.org/male-sixth-grade-math-teacher-with-protractor (photographer Allison Shelley for EDUimages). The photograph of Gia is sourced from https://images.all4ed.org/elementary-boy-with-backpack-and-girl-with-notebook/ (photographer Allison Shelley for EDUimages). The photograph of Gabe is sourced from https://images.all4ed.org/third-grade-boy-with-backpack-outside/. The photograph of Luisa is sourced from https://images.all4ed.org/fifth-grade-girl-mask-break (photographer Allison Shelley for EDUimages).

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

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 C.6). 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 C.6).

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 toward the specific purpose of the block (in this case, preparing for a book discussion) as well as the goal of each task. Exhibits C.6–C.11 illustrate items that help students accomplish the first task of learning about the events and characters. Exhibits C.12–C.14 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.

Item response types 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 might, for example, include single-selection multiple choice items to assess readers' ability to locate and recall important events and other details (see Exhibit C.6), short constructed-response items that include fill-in-the-blank options (see Exhibit C.7), multiple-selection multiple choice items (see Exhibit C.8), and longer short constructed response items that ask readers to interpret and integrate details about the character's thoughts, feelings, and actions into their understanding of the story (see Exhibit C.10).

Exhibit C.6. A Grade 4 RDU Block Illustrating a Locate and Recall Multiple Choice Item. The Teacher Reminds the Reader of the Specific Purpose (to Prepare for a Discussion) and the First Task (to Learn About Events and Characters)

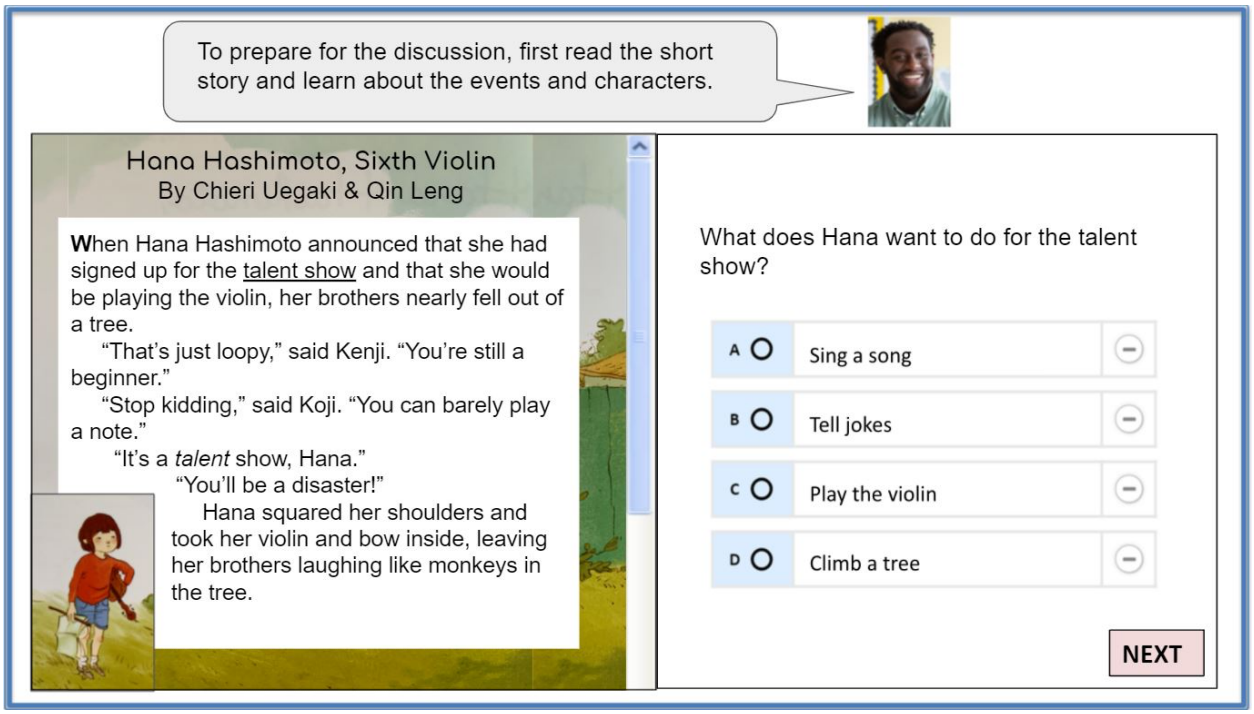


Exhibit C.7. A Grade 4 Locate and Recall Item Illustrating a Fill-In-the-Blank Short Constructed Response Item

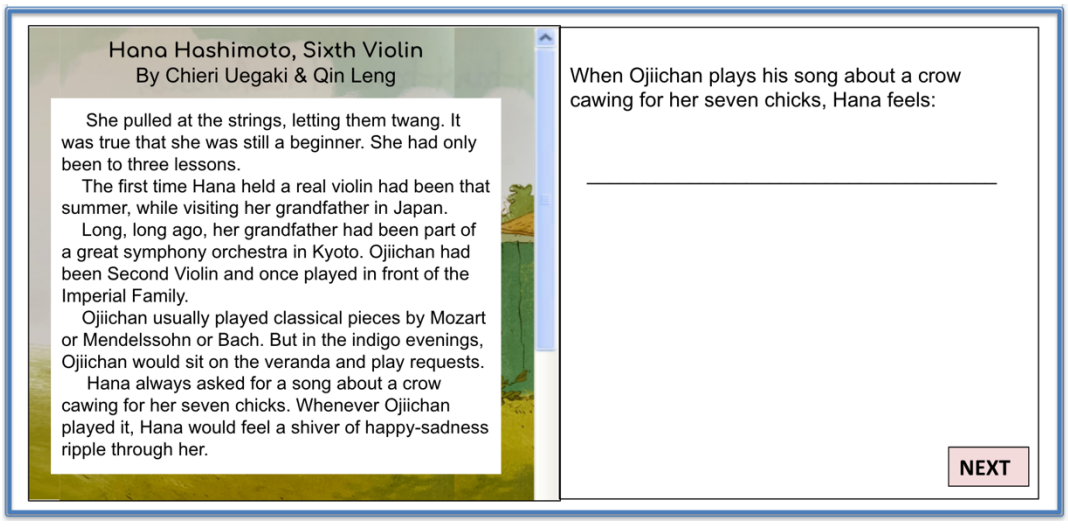
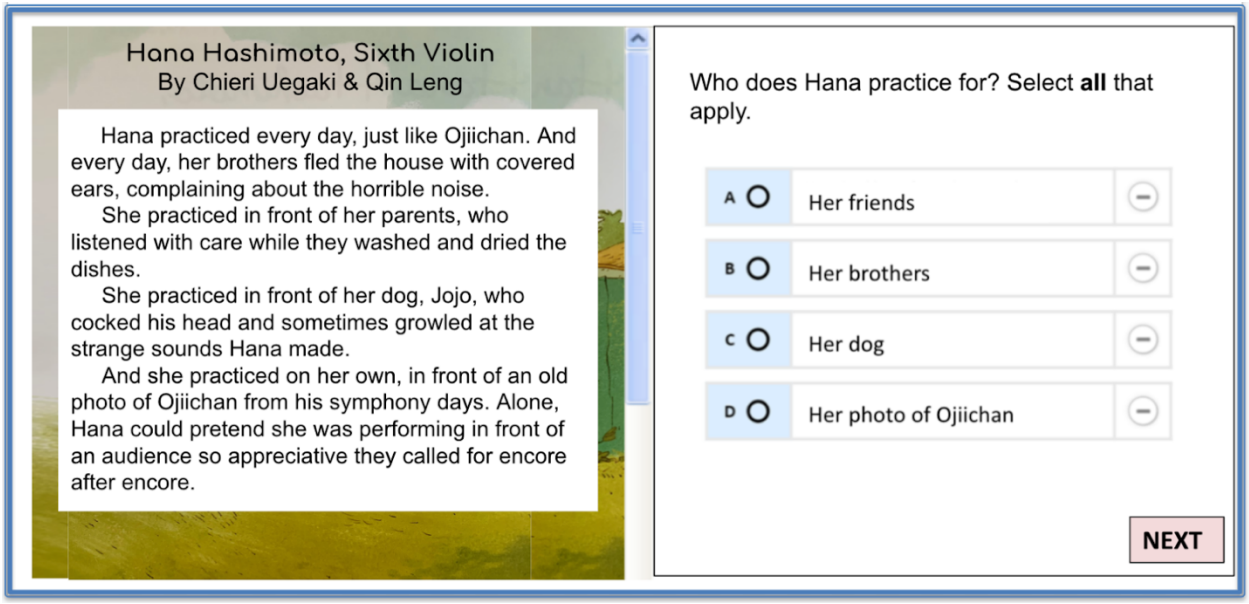
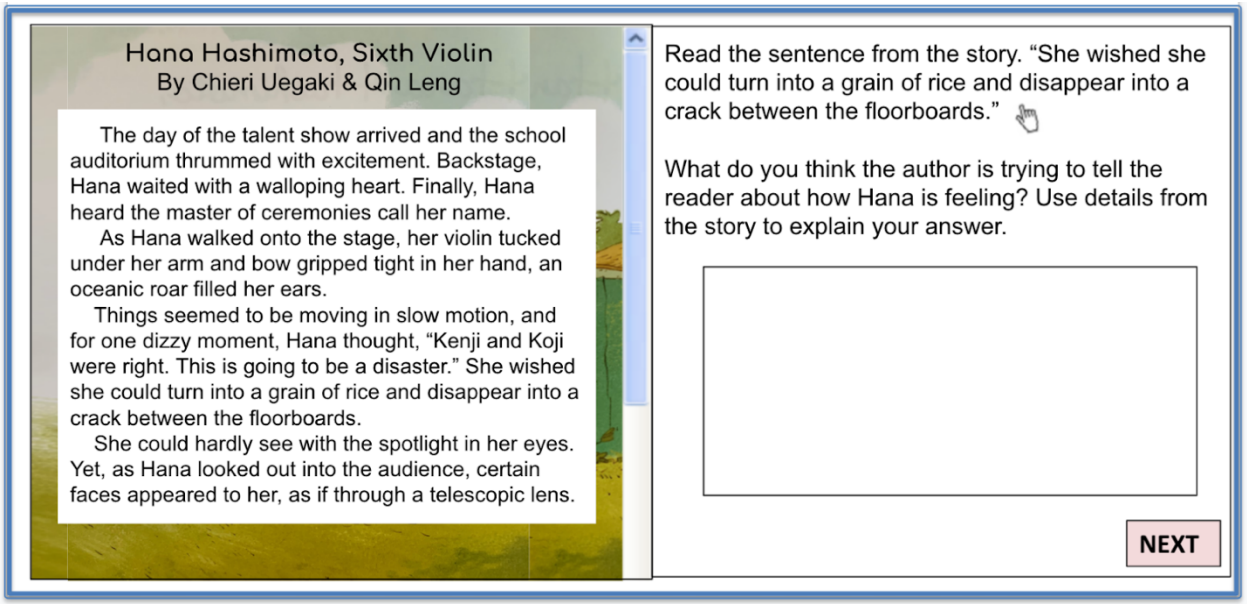


Exhibit C.8. A Grade 4 Locate and Recall Item Illustrating a Multiple-Selection Multiple Choice Response Format



In addition, **a look-back button (a task-based UDE)** is embedded into items with excerpted text (see Exhibits C.9 and C.10). If readers wish, they can click to see exactly where the excerpted text is located in the context of the original story in the assessment space. Multiple choice and constructed response item formats are interspersed throughout the assessment.

Exhibit C.9. A Grade 4 Analyze and Evaluate Short Constructed Response Item Illustrating a Task-Based UDE in the Form of a Look-Back Button That Refers Readers to the Relevant Section of Text



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. Her family loves her performance so much that later that evening, they ask her to play them more musical notes around the dinner table.

Exhibit C.10. 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

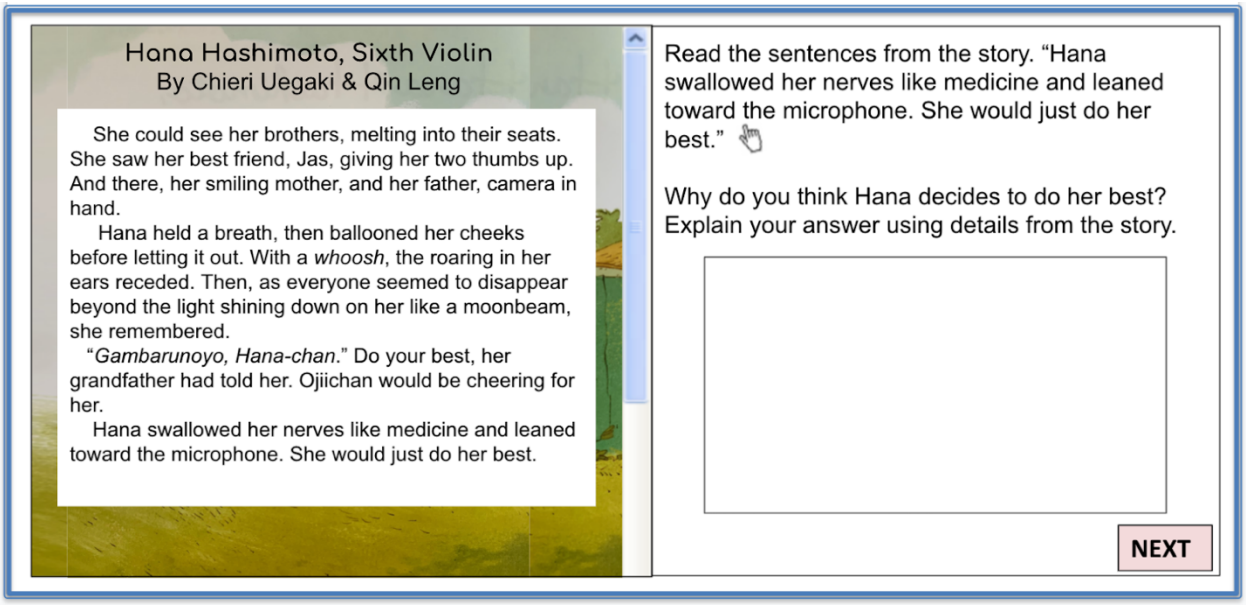
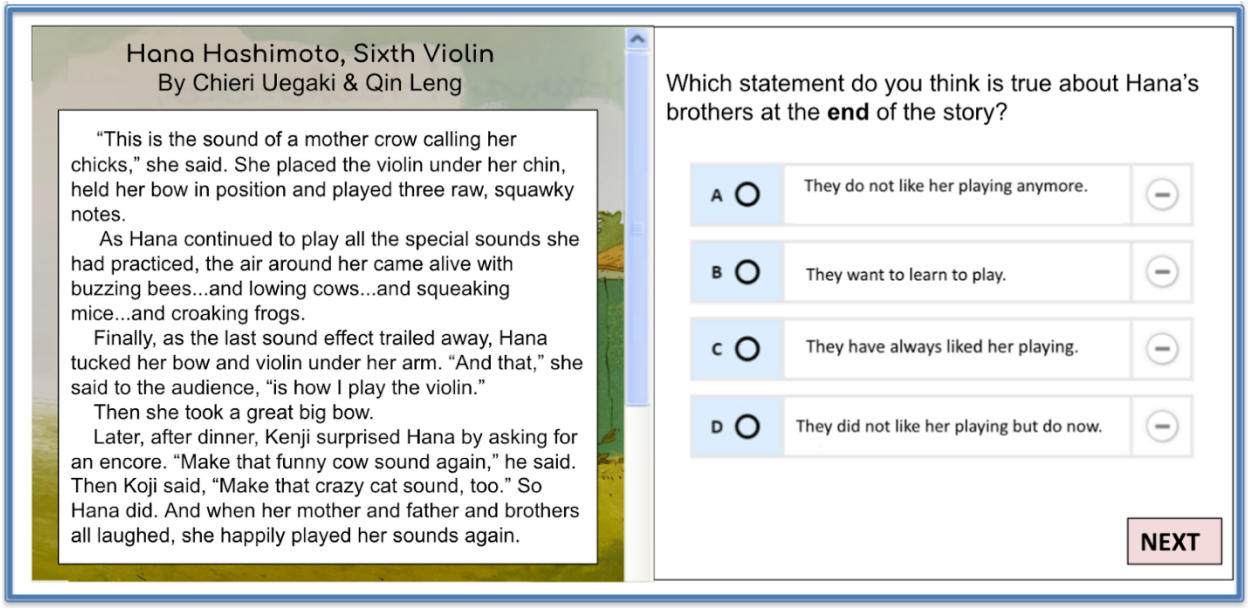


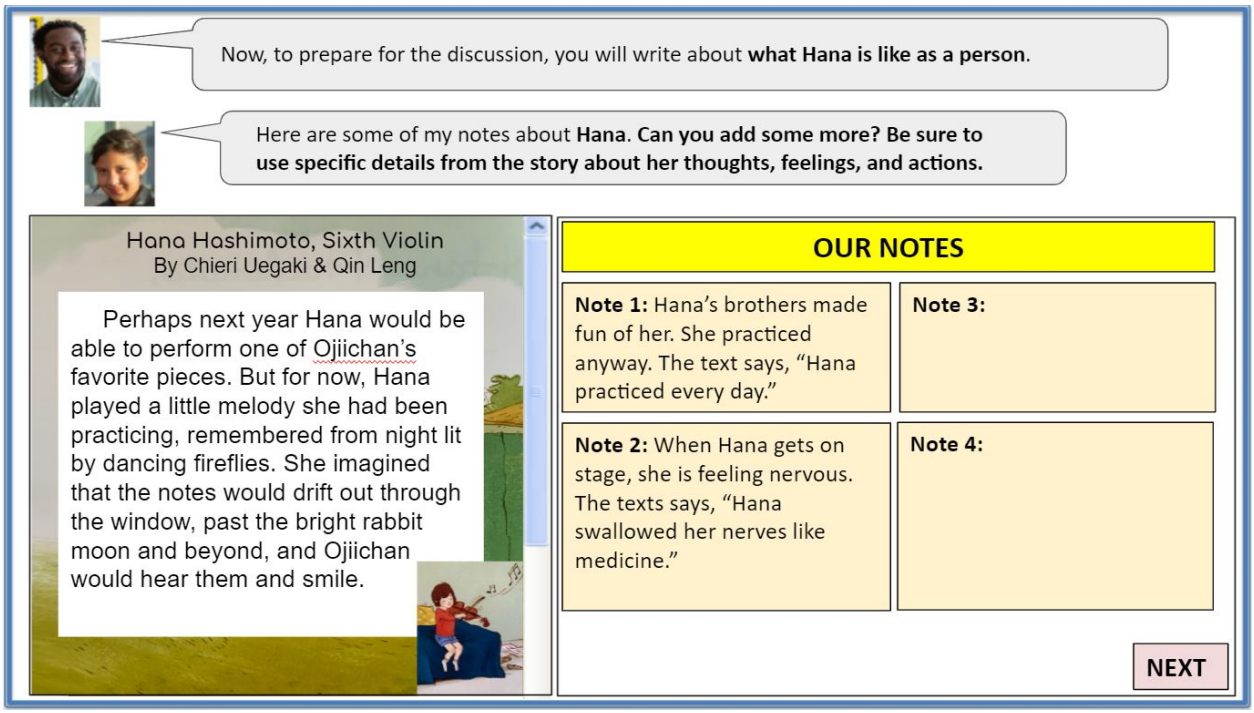
Exhibit C.11. A Grade 4 Integrate and Interpret Item for the First Task Using a Single-Selection Multiple Choice Format



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 C.12).

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 C.13).

Exhibit C.12. Teacher and Student Task Characters Remind Readers of the Second Task Goal in This Integrate and Interpret Item

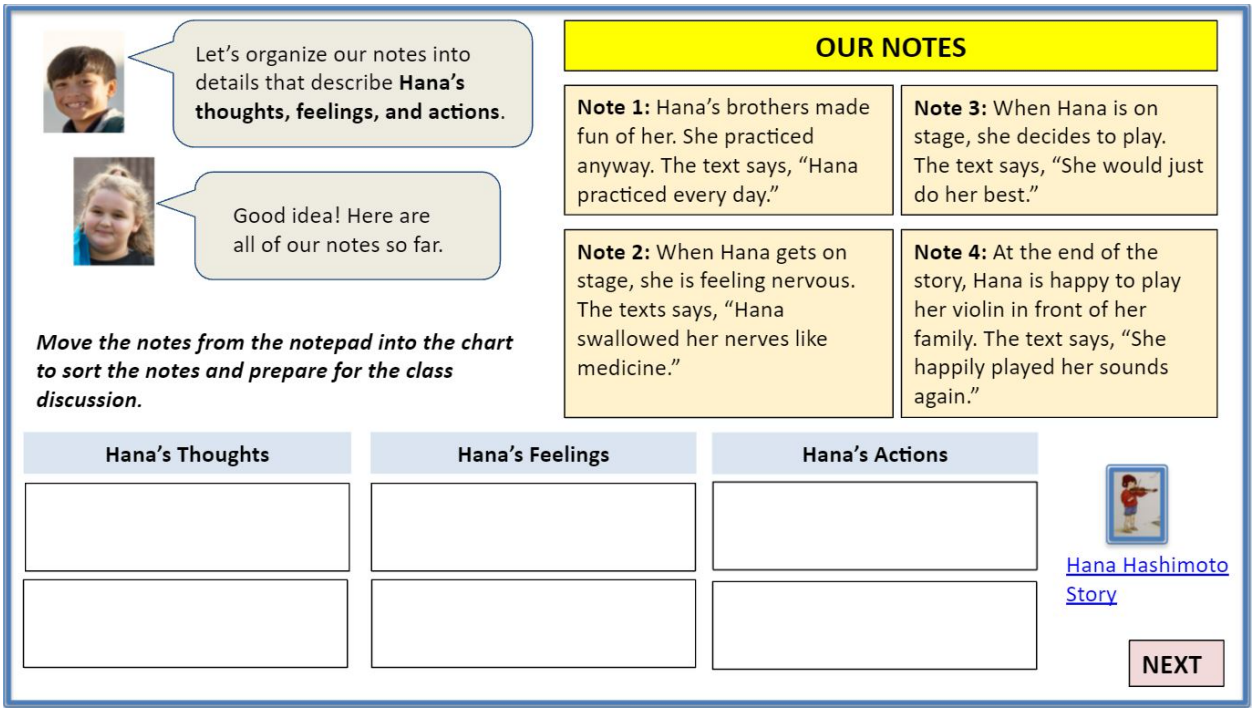


In Exhibit C.13, the other two classmates serve as **motivational and task-based UDEs** to engage students in the task while also reminding them to stay focused on the character's thoughts, feelings, and actions. The student's responses from the previous item are carried over to the next item as the 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 C.13, 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 C.14), 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.

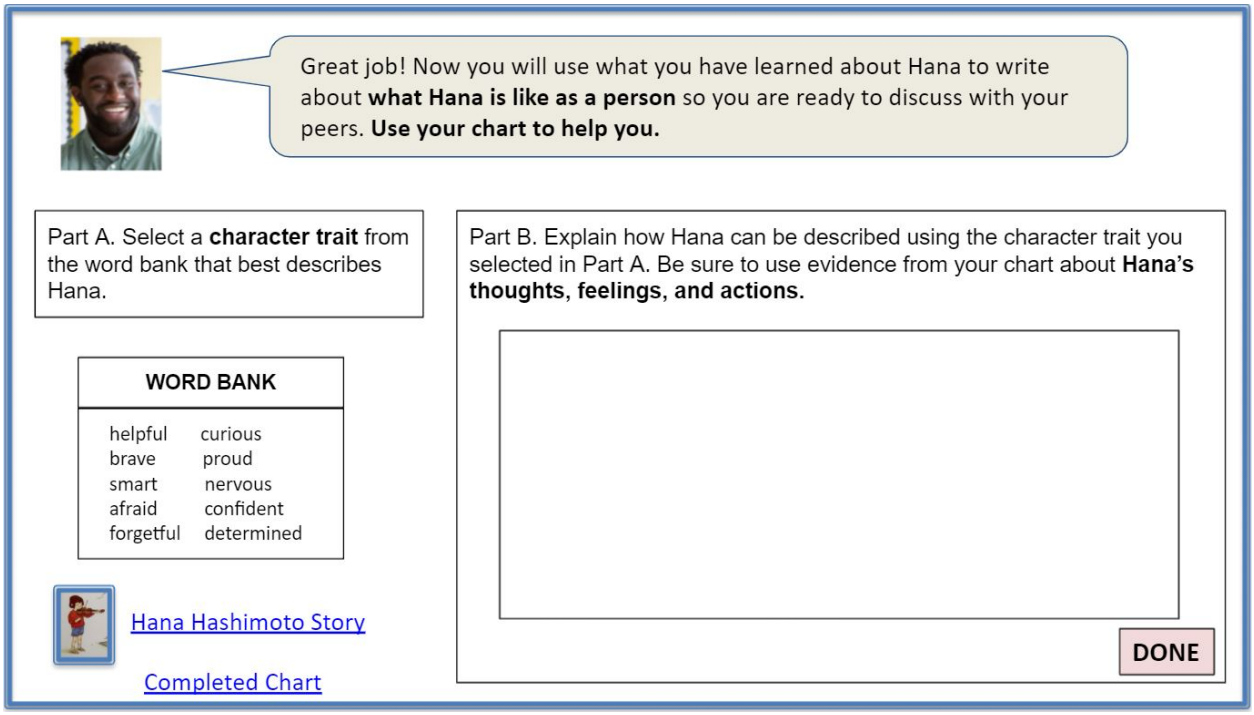
Exhibit C.13. The Student's Responses From Their Completion of the Previous Item Are Carried Over to the Next Item as the Completed Notes. A Graphic Organizer With Drag-and-Drop Features Offers Students an Efficient Way to Demonstrate Their Understanding of How the Text Conveys the Character's Thoughts, Feelings, and Actions in This Grade 4 Integrate and Interpret Item



A longer constructed response item such as the example shown in Exhibit C.14 is designed to assess readers' ability to Use and Apply understandings learned from the story to form a characterization of Hana. As readers engage with this final part of the block, 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 C.14, 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; when readers click on a word, the word is highlighted and is recorded as the student's answer to Part A) when asked to describe 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 C.14. 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 an Extended Constructed-Response Item Format. 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.



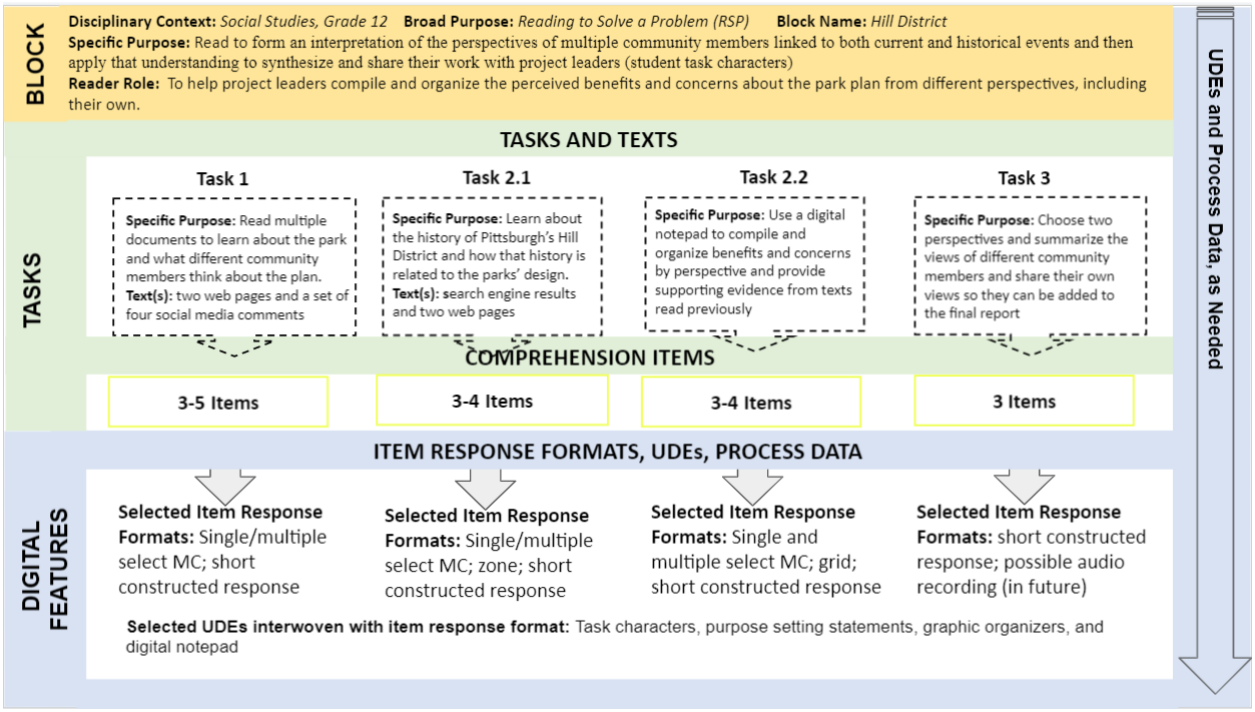
Performance Evidence and Indicators*. When interpreting reading achievement from performance on the 2026 NAEP Reading Assessment, multiple indicators can be used to explain what students are able to do. As indicated earlier in this chapter, each block would be classified with a primary disciplinary context, grade level, and broad purpose. Scores from the Hana Hashimoto, Sixth Violin block, then, 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 students' ability to develop their understanding of a single text and then apply that understanding in a simple culminating event (in this case, describing the kind of person Hana is based on her thoughts, feelings, and actions in the story).*

Test developers keep 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 subgroups 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 12

Block Components (Context, Purposes, and Reader Role). This block is designed to assess how 12th-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 C.15 for the block design and Exhibit C.16 for the introduction to the block).

Exhibit C.15. Block Design for Hill District Sketch



More specifically, readers are invited to engage with three students (represented by task characters in the assessment) who have been asked by the mayor to compile and organize public reactions to an ambitious plan 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.

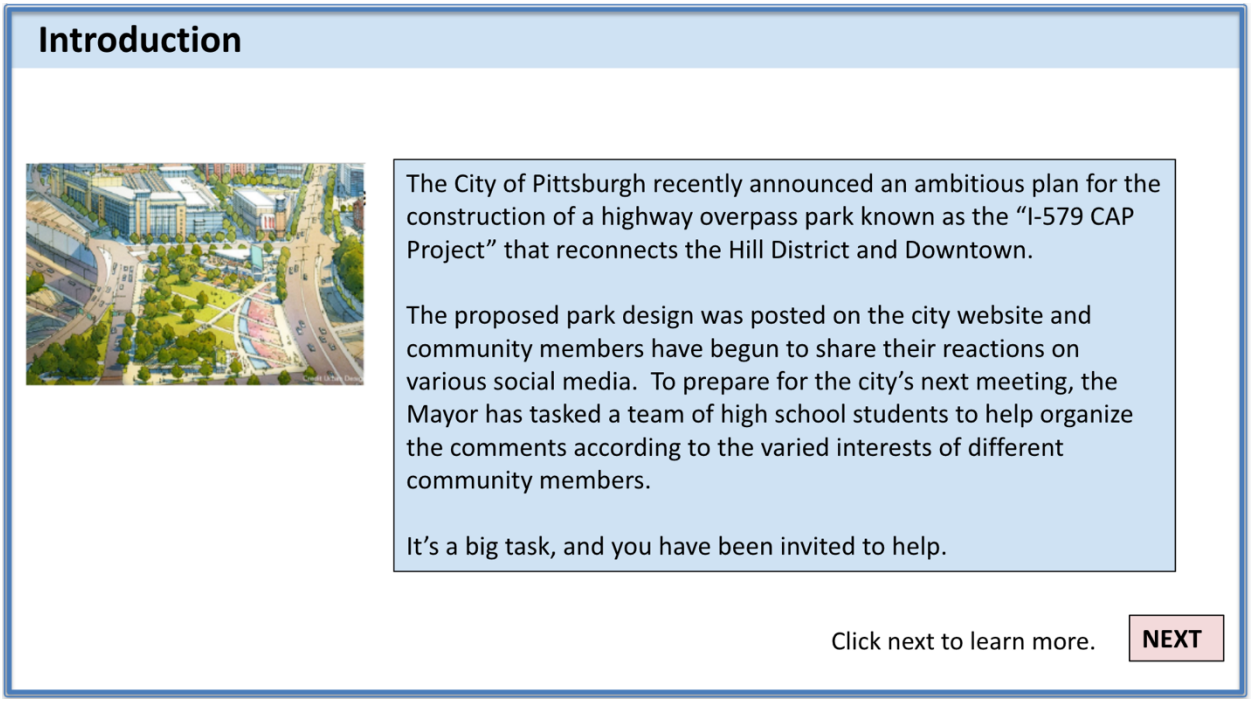
The tasks in this RSP block reflect design features that are more dynamic and cumulative in terms of content and format, as depicted toward the right side of the continuum in Exhibit C.2. For example, readers are constrained by specific purposes and role expectations about how to engage with provided texts. The four tasks (and related subtasks) are tightly structured so that one task builds on the previous task, 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.

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 cut across issues of contemporary and historical relevance. Throughout the block, readers are asked to activate and employ their personal, cultural, and civics 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.

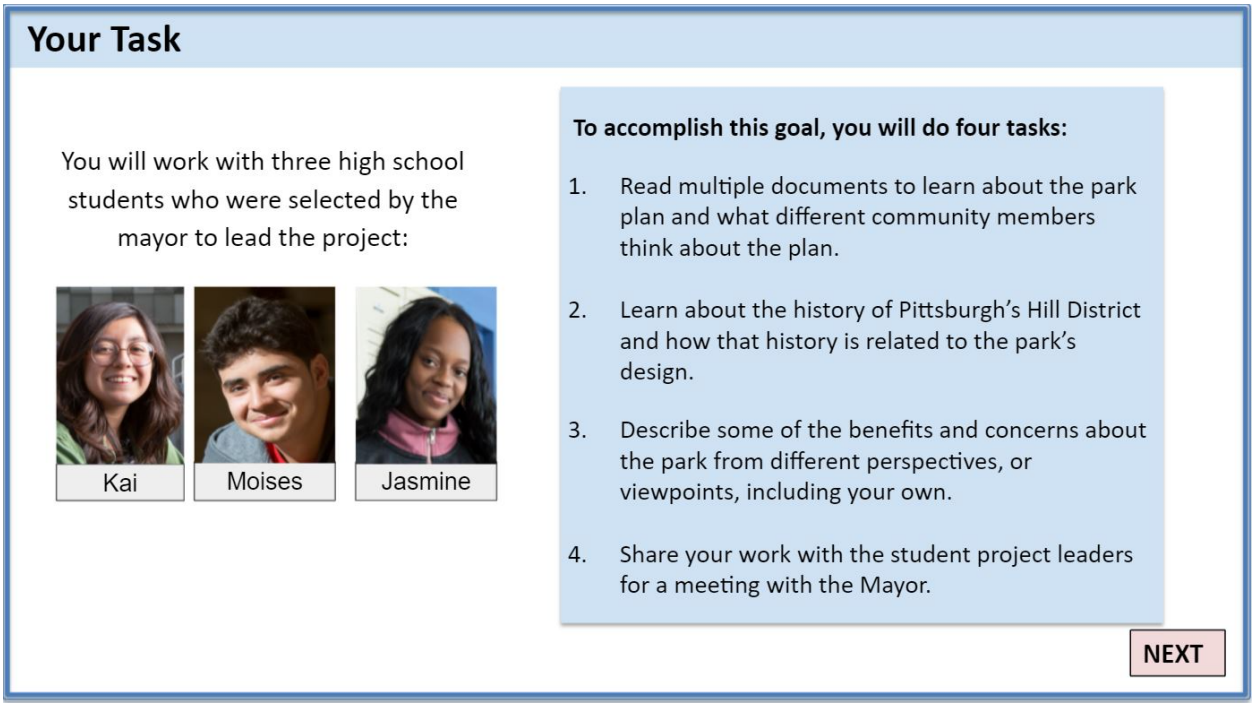
Specific Reading Purpose(s) and Reader Role. At the beginning of the assessment (see Exhibit C.16), students learn that the city has recently unveiled the park plan to the public on its website and city residents have been invited to share their reactions on various social media. Students are also introduced 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 C.17). 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 C.16. A Social Studies Context and Reader Role Serve to Situate Readers in a Grade 12 RSP Block Involving Several Interconnected Digital Texts



Screenshot and text used with permission from https://triblive.com/local/pittsburgh-allegheny/pittsburgh-council-authorizes-32-million-in-spending-on-i-579-cap/. Image of Lower Hill redevelopment used with permission from LaQuatra Bonci.

Exhibit C.17. 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



Throughout Appendix C, the photograph of Kai is sourced from https://images.all4ed.org/high-school-boy-and-girl-near-playground (photographer Allison Shelley for EDUimages). The photograph of Moises is sourced from https://images.all4ed.org/high-school-boy-in-hallway (photographer Allison Shelley/The Verbatim Agency for EDUimages). The photograph of Jasmine is sourced from https://images.all4ed.org/high-school-boy-and-girl-drive-robots (photographer Allison Shelley/The Verbatim Agency for EDUimages).

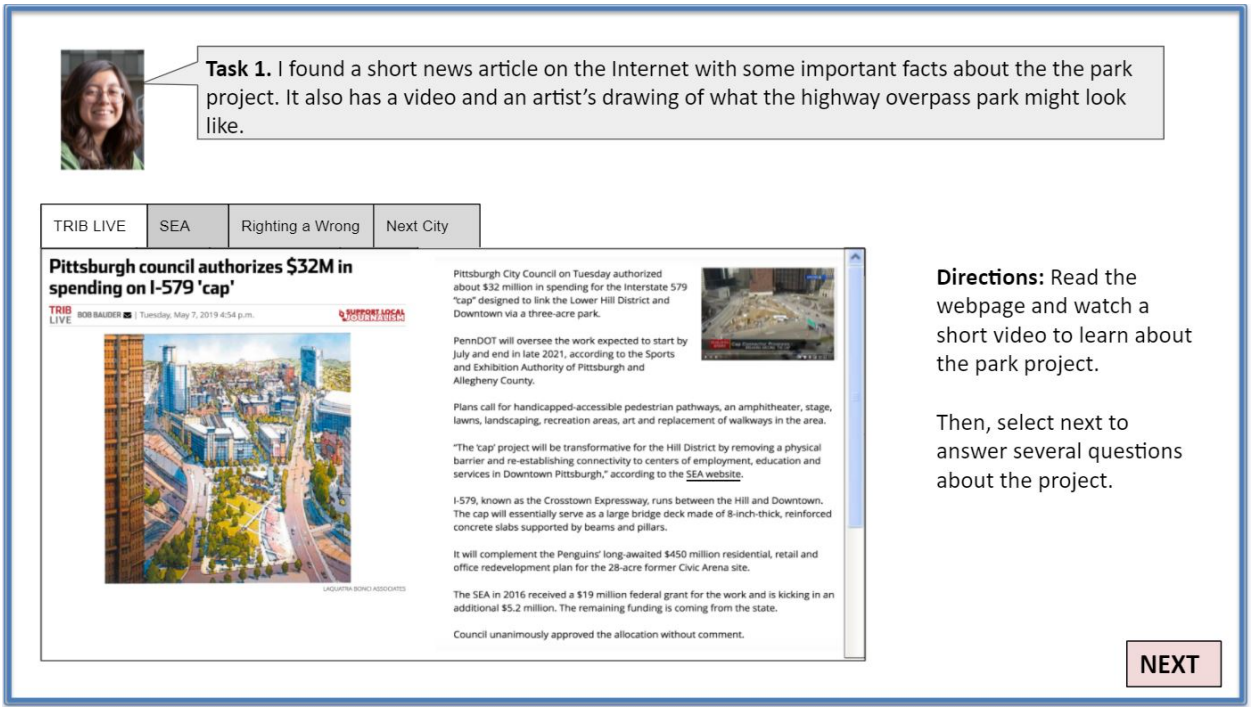
Task Components (Tasks, Text(s), and Items).

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 opportunity to express their own opinions about the project) serve to guide and motivate readers to engage with the 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 plan. These include social media comments from community members; 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 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.

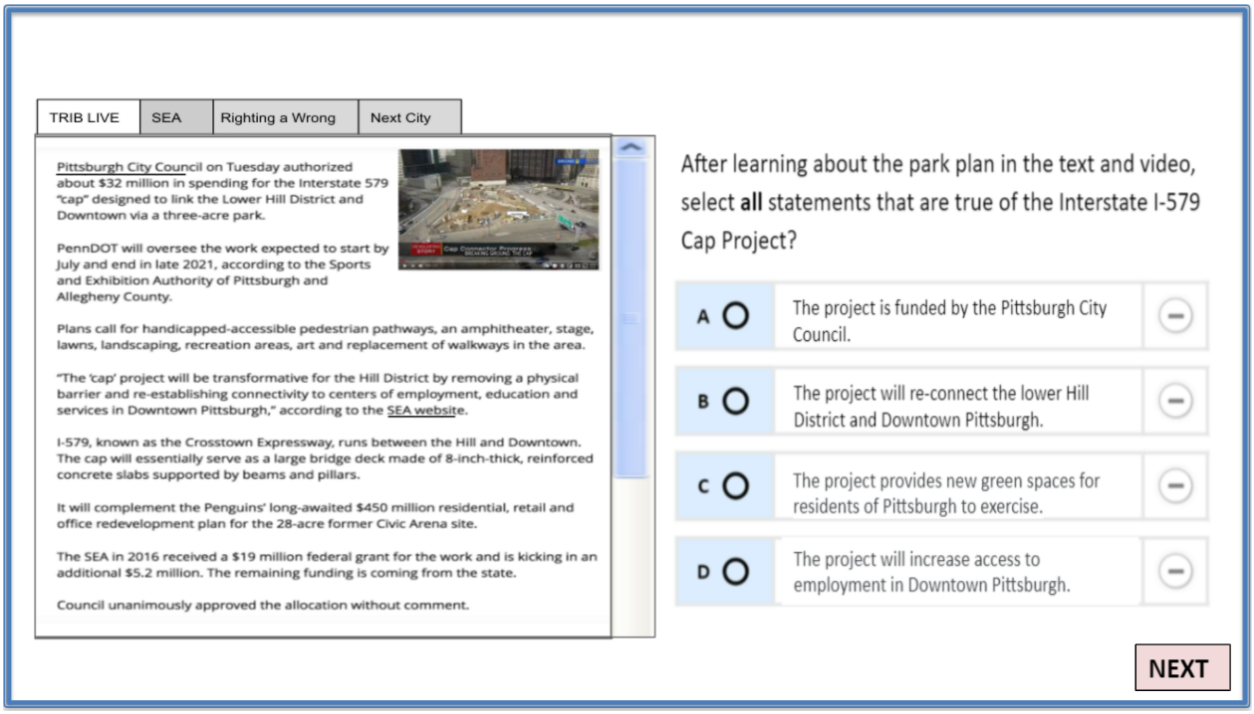
Comprehension Items. Item response types would vary from simple multiple choice to short answer or hybrid constructed response items to give readers different kinds of opportunities to demonstrate their understanding in the block and apply that understanding to solve the problem. While some items give students opportunities to demonstrate their understanding and develop thinking within a specific text, other items are designed to assess how readers navigate and make meaning across sources representing multiple and diverse perspectives. After being asked to read text and watch a short video on a website about the park project (Exhibit C.18), 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 C.19 and C.20). 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 C.20). **Task-based UDEs** (e.g., one of three task characters) provide short prompts (shown at the top of Exhibits C.18 and C.21) designed to cue the reader about the steps they are completing as they read across different sources to solve the problem.

Exhibit C.18. 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



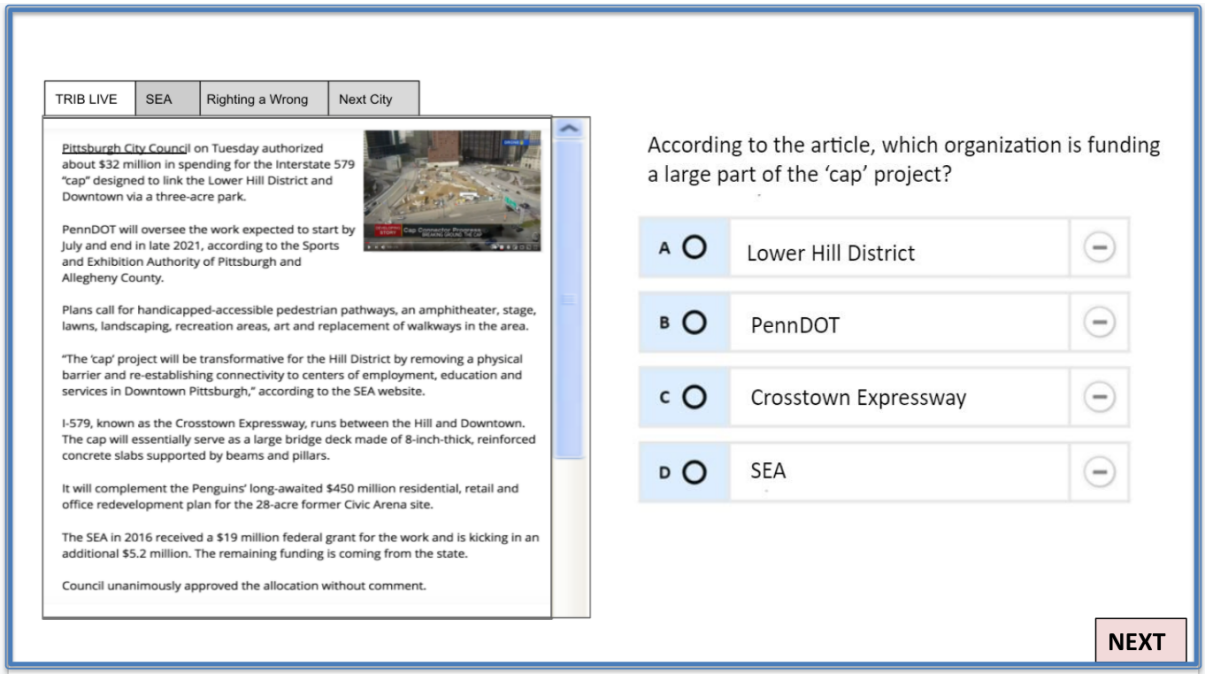
Screenshot and text used with permission from https://triblive.com/local/pittsburgh-allegheny/pittsburgh-council-authorizes-32-million-in-spending-on-i-579-cap/. Image of Lower Hill redevelopment used with permission from LaQuatra Bonci. Image of video (https://www.youtube.com/watch?v=97eySeLPlo4) used with permission from WPXI.

Exhibit C.19. A Grade 12 Locate and Recall Item Illustrating a Multiple-Selection Multiple Choice Response Format



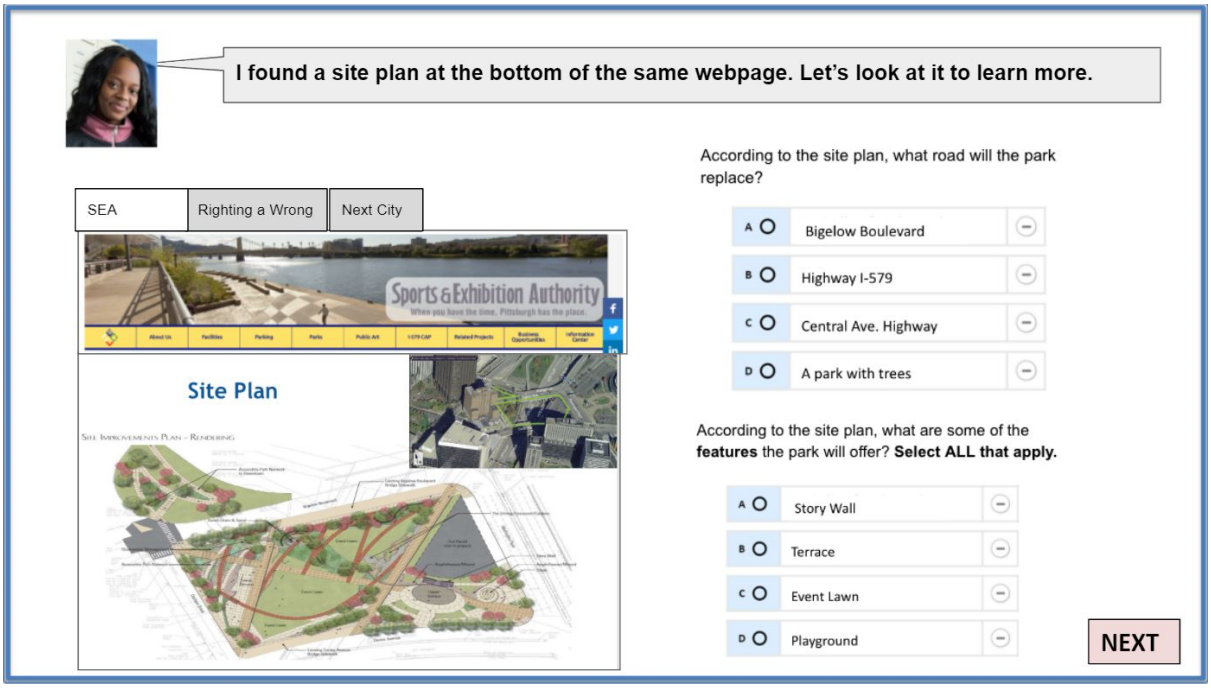
Screenshot and text used with permission from https://triblive.com/local/pittsburgh-allegheny/pittsburgh-council-authorizes-32-million-in-spending-on-i-579-cap/. Image of video (https://www.youtube.com/watch?v=97eySeLPlo4) used with permission from WPXI.

Exhibit C.20. A Grade 12 Locate and Recall Item Illustrating a Single-Selection Multiple Choice Item Response Format



Screenshot and text used with permission from https://triblive.com/local/pittsburgh-allegheny/pittsburgh-council-authorizes-32-million-in-spending-on-i-579-cap/. Image of video (https://www.youtube.com/watch?v=97eySeLPlo4) used with permission from WPXI.

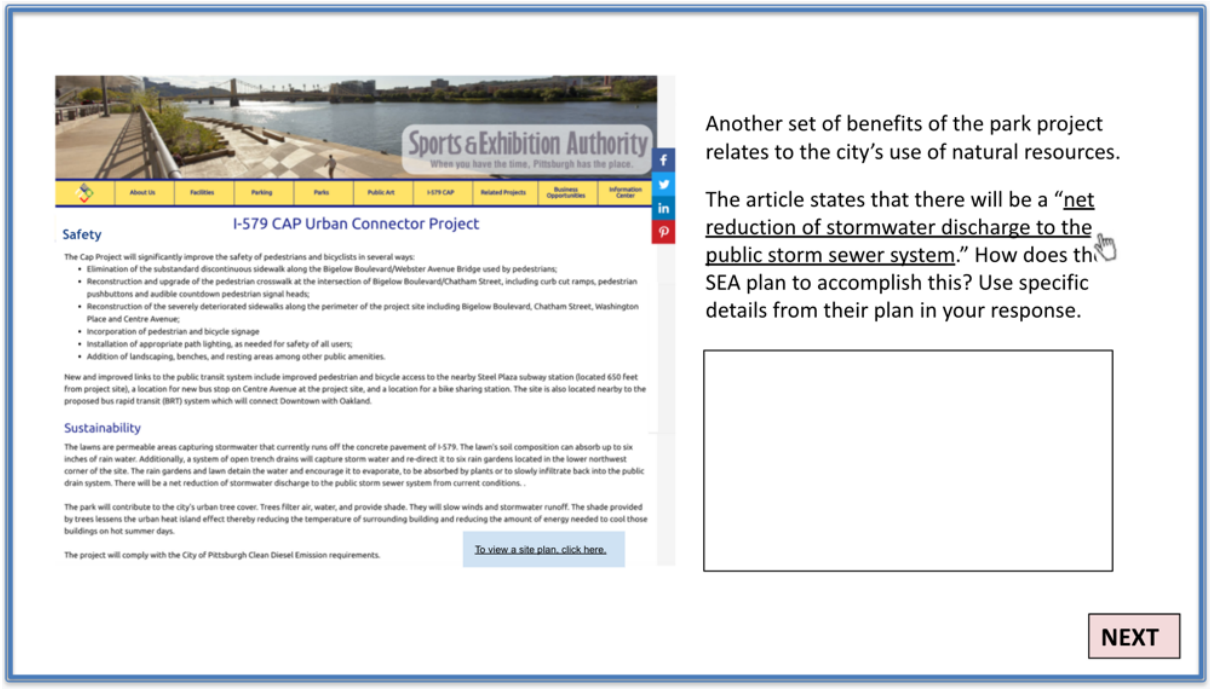
Exhibit C.21. Two Grade 12 Items That Ask Readers to Integrate and Interpret (Item 1) and Locate and Recall (Item 2) Ttextextual and Visual Information From an Artist's Rendering of the Site Improvement Plan Published on a Website



Screenshot and text used with permission from http://www.pgh-sea.com/index.php?path=i5-ucp.

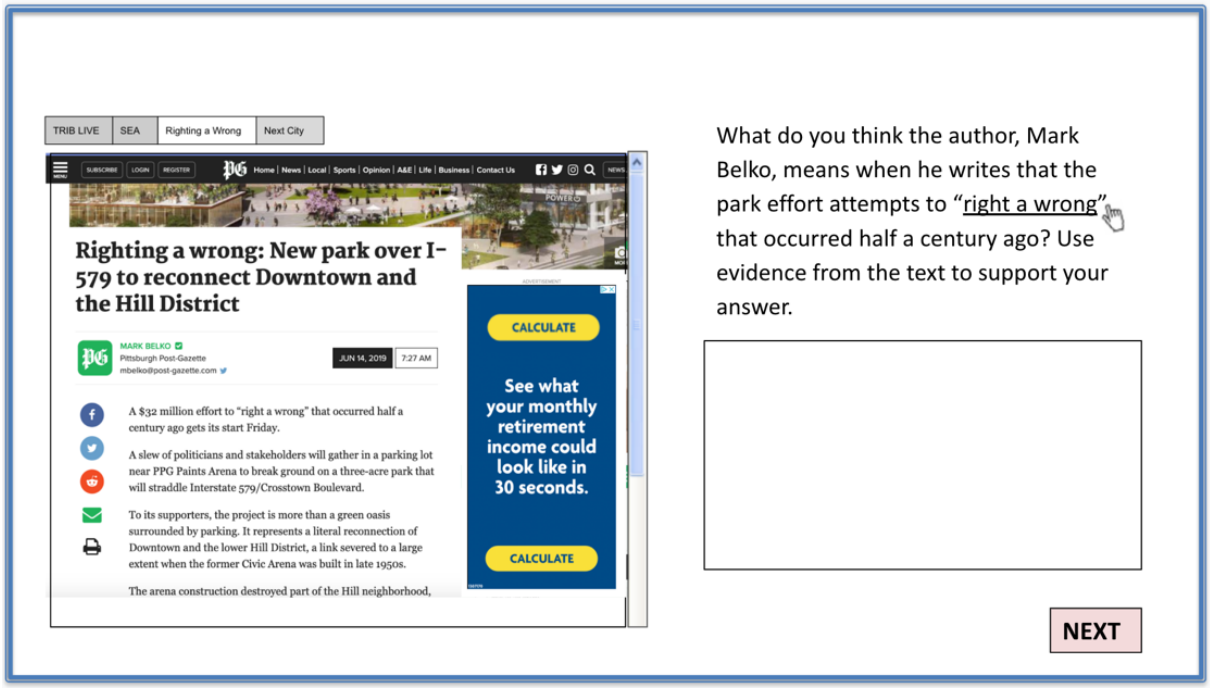
Examples of short constructed response items earlier in the block might ask readers to integrate and interpret information about how park designers plan to modify the city's use of natural resources to address environmental concerns (Exhibit C.22). Later in the block, readers might be asked to integrate and interpret information in an online newspaper article about the historical significance of the park's design (Exhibit C.23) or to analyze and evaluate the requests of some community members to include park features that honor the history of their neighborhood (Exhibit C.24). Also depicted in Exhibit C.24 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 C.22. 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



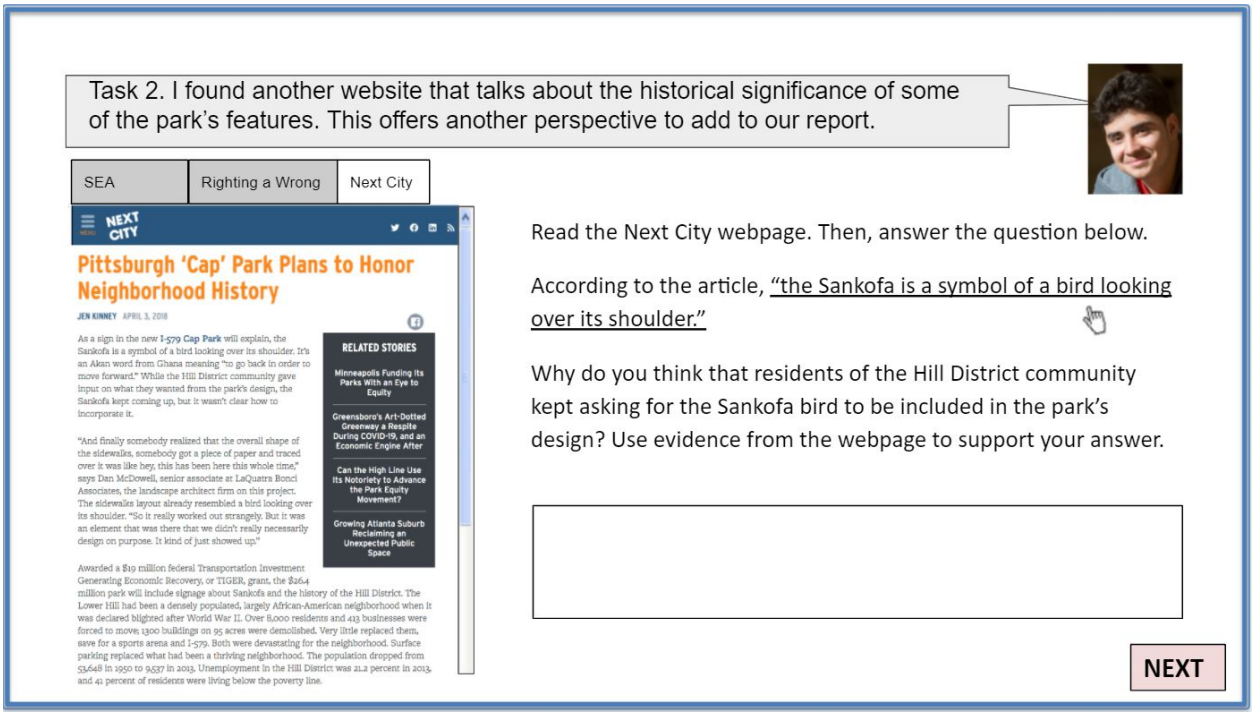
Screenshot and text used with permission from http://www.pgh-sea.com/index.php?path=i5-ucp.

Exhibit C.23. 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



Screenshot and text Copyright ©, Pittsburgh Post-Gazette, 2021, all rights reserved. Reprinted with permission. https://www.post-gazette.com/business/development/2019/06/14/new-park-over-I-579-to-bridge-Downtown-pittsburgh-Hill-District/stories/201906140069. Image of Lower Hill redevelopment used with permission from Gensler.

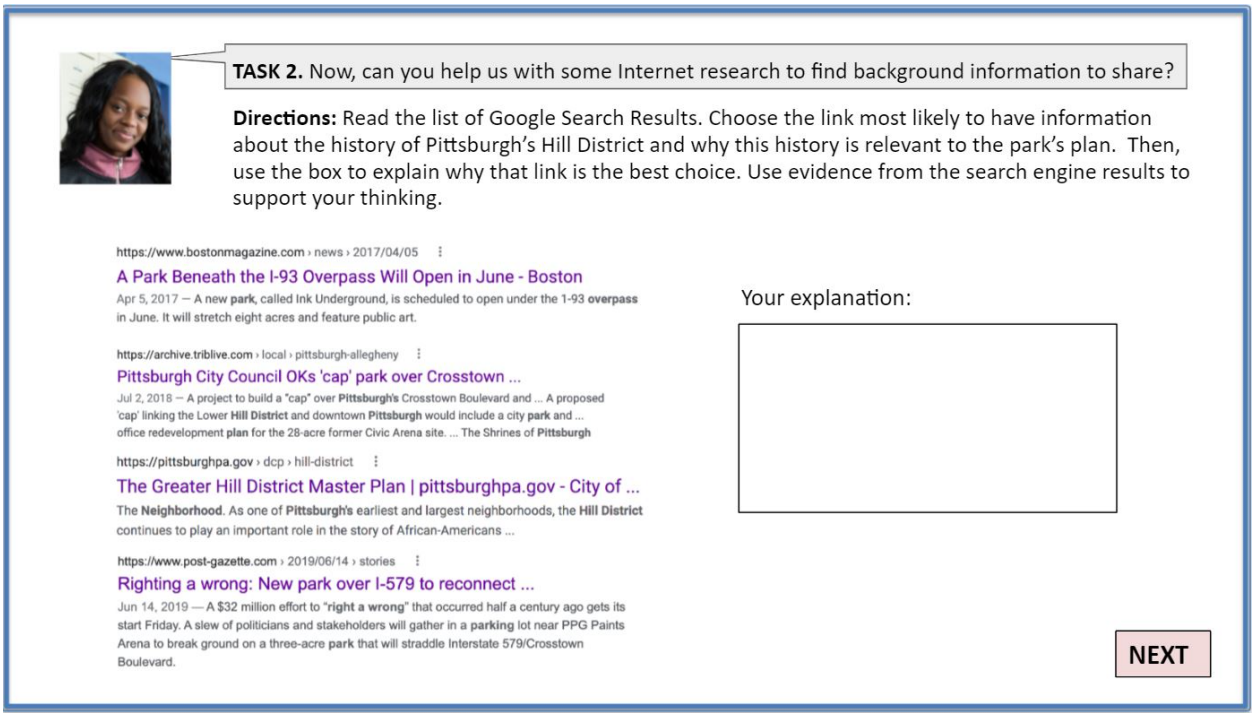
Exhibit C.24. A Grade 12 Short Constructed Response Item That Asks Readers to Integrate and Interpret Information on a Webpage With a Look-Back Button (Task-Based UDE). The Task Character Reminds Readers of the Specific Purpose of the Second Task



Screenshot and text used with permission from https://nextcity.org/daily/entry/pittsburgh-cap-park-plans-to-honor-neighborhood-history.

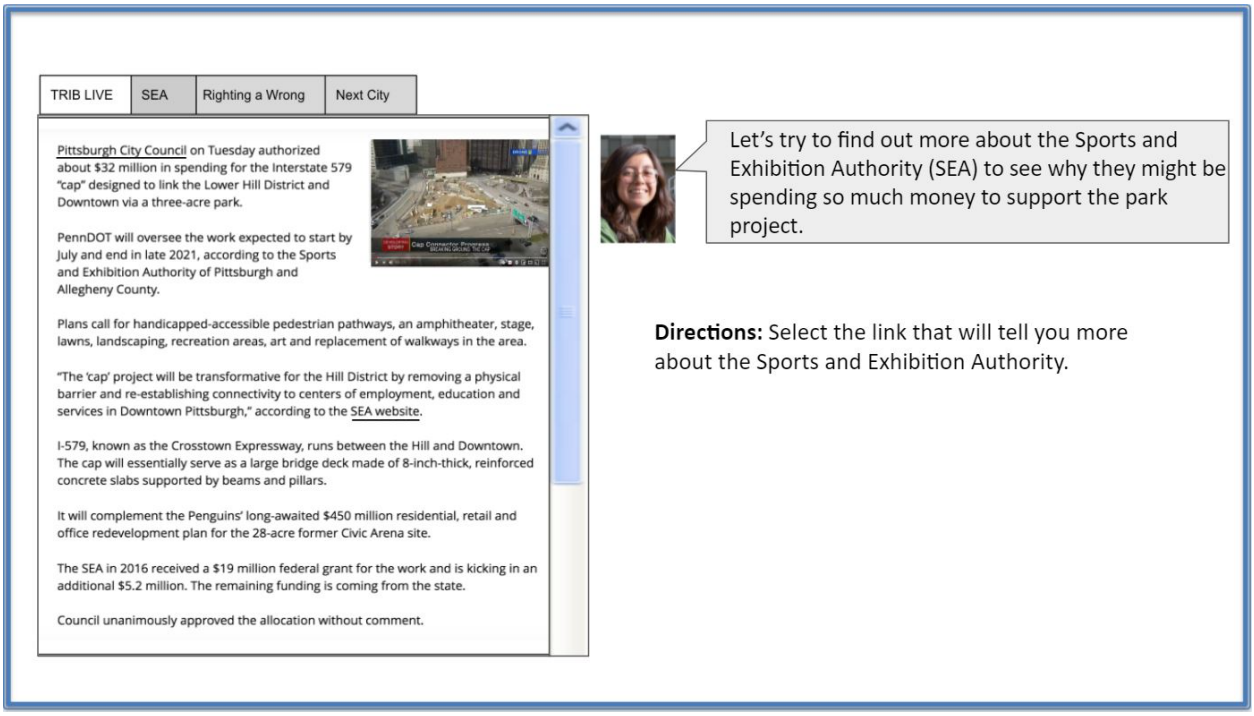
Other potential items might ask readers to locate and evaluate the relevance of search engine results pertaining to the historical significance of some of the park's features (see Exhibit C.25) or to locate (navigate to) and then analyze information from a website's menu to evaluate the expertise of the group responsible for publishing information about the park project (see Exhibits C.26 and C.27, 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 C.25. A Grade 12 Selected Response Zone Item Designed to Capture Process Data About Which Link Is Selected and Paired With a Short Constructed Response Scored Item That Asks Readers to Analyze and Evaluate the Relevance of Their Search Engine Choice



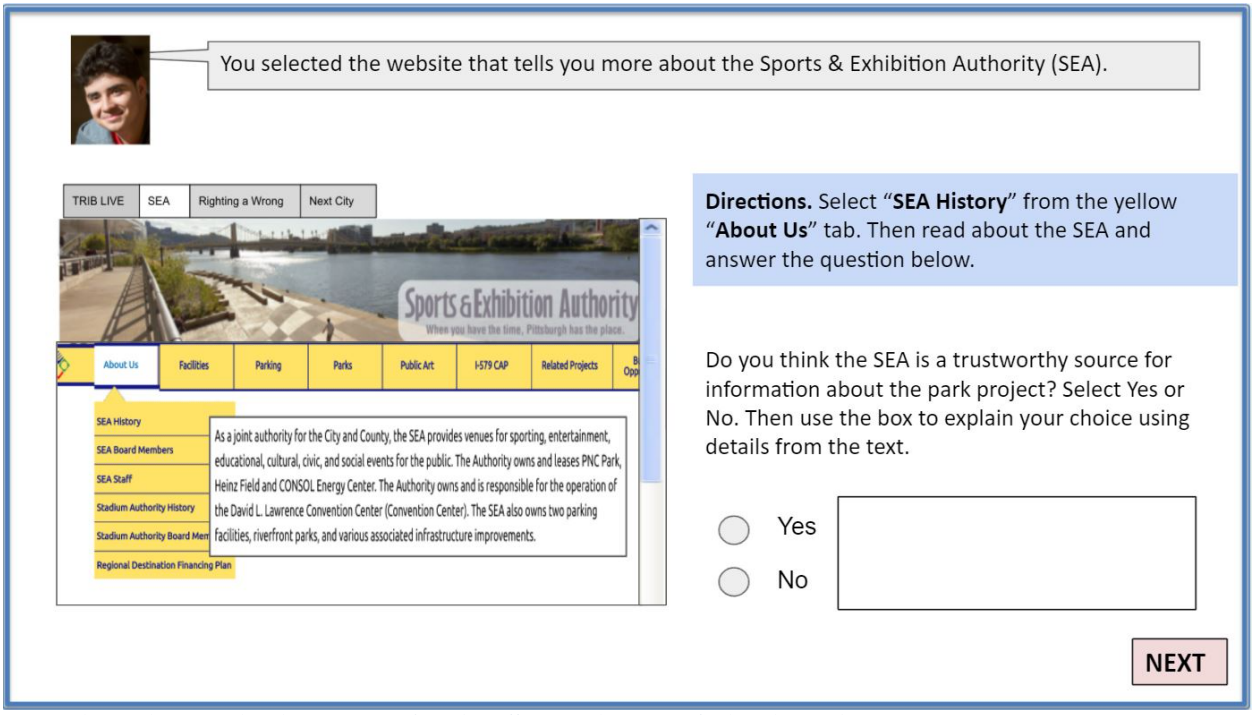
Google is a trademark of Google LLC and this publication is not endorsed by or affiliated with Google in any way.

Exhibit C.26. A Grade 12 Selected Response Zone Item Designed to Capture Process Data About How Readers Navigate Through Hyperlinked Webpages



Screenshot and text used with permission from https://triblive.com/local/pittsburgh-allegheny/pittsburgh-council-authorizes-32-million-in-spending-on-i-579-cap/. Image of video (https://www.youtube.com/watch?v=97eySeLPlo4) used with permission from WPXI.

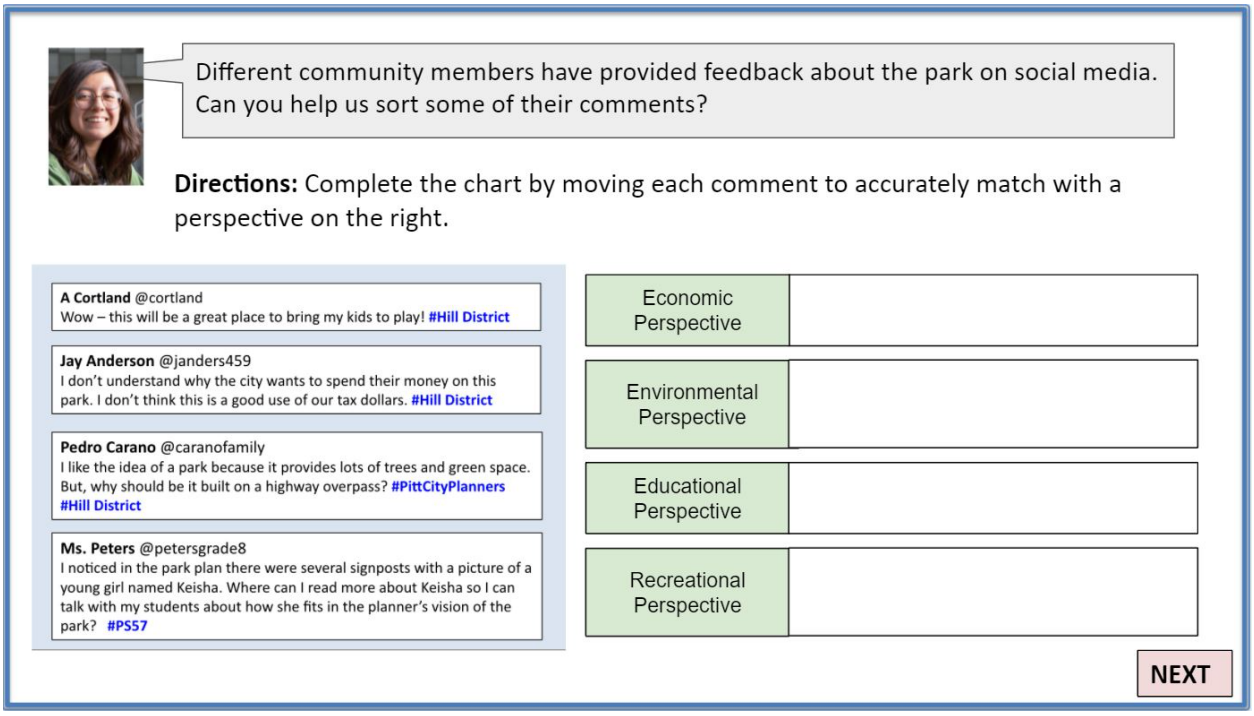
Exhibit C.27. A Grade 12 Critical Online Resource Evaluation Item That Asks Readers to Analyze and Evaluate the Extent to Which an Organization Has the Appropriate Qualifications to Publish Details About the Proposed Park Plan on Their Website Using a Hybrid Constructed Response



Screenshot and text used with permission from http://www.pgh-sea.com/index.php?path=i5-ucp.

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 C.28, 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 and it functions 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 C.28. 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 C.29 and C.30) 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 C.29 applies a multiple-selection response format with a **task-based UDE (table) and a 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 match each comment with one or more specific benefits on the right.

In contrast, Exhibit C.30 engages readers in a similar matching process, but for this item, a task character (**motivational UDE**) asks 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. 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 C.31 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 C.29. 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

I noticed that there are a lot of different opinions and perspectives on the park in this website. I was thinking we could organize these by topic and add these to our summary report for the Mayor.

Directions. The table below lists comments from two community members and columns with three benefits of the proposed plan. **Select <u>one or more</u> benefit that applies to each person's comment.**

Comments from Community Members as Quoted in Website #1 ("Righting a Wrong")	Connects Hill District to Downtown	Offers Green Space	Rights A Wrong
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*)	◯	◯	◯

NEXT

Exhibit C.30. 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

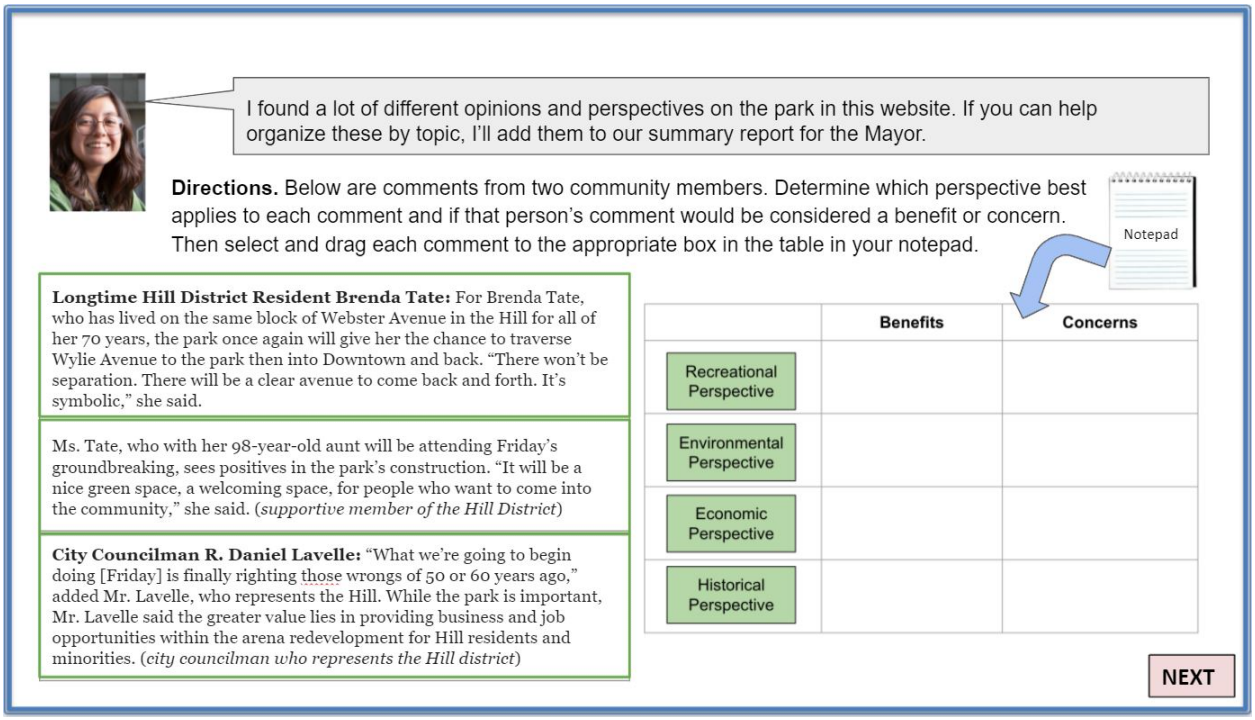
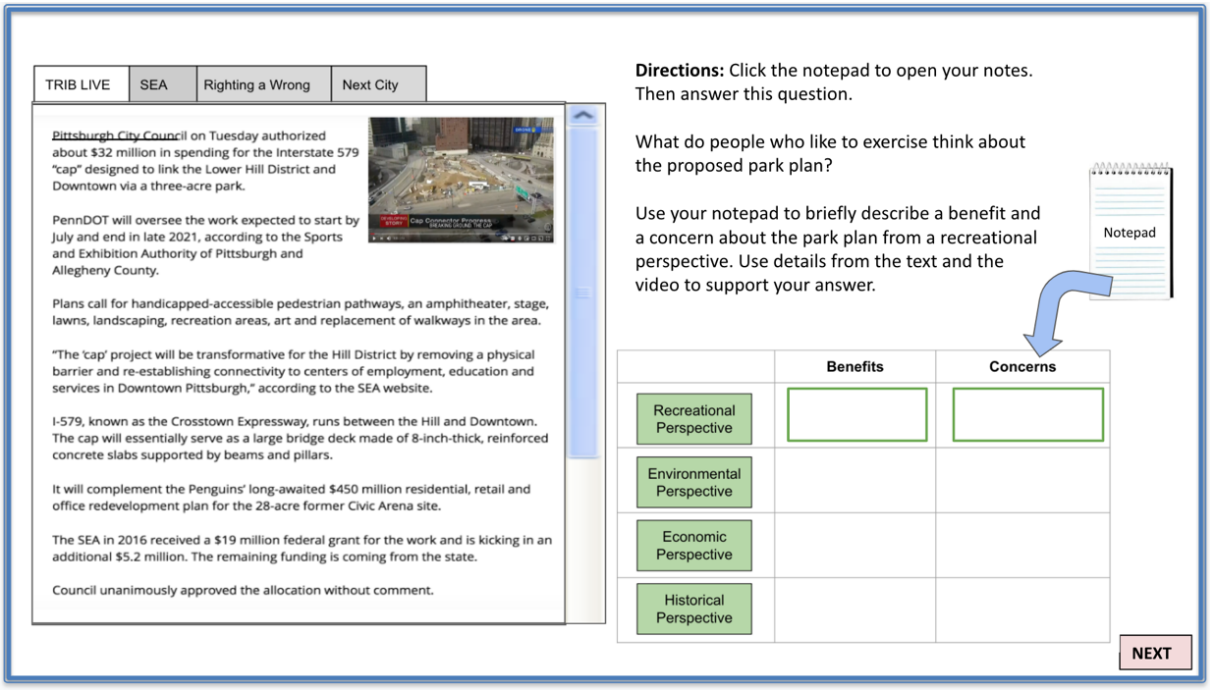


Exhibit C.31. A Grade 12 Dynamic Matching Response Grid Item With a Task-Based UDE (Retractable Digital Notepad) That Serves to Support Readers at Another Point in the RSP Block as They Read Across Multiple Sources to Solve the Problem at Hand



Screenshot and text used with permission from https://triblive.com/local/pittsburgh-allegheny/pittsburgh-council-authorizes-32-million-in-spending-on-i-579-cap/. Image of video (https://www.youtube.com/watch?v=97eySeLPlo4) used with permission from WPXI.

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 C.32), 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 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 have 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 C.33 illustrates how this item might look using a short constructed response format, similar to those in existing NAEP assessment blocks, and

Exhibit C.34 is included to depict what an item might look like in 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 C.32. 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

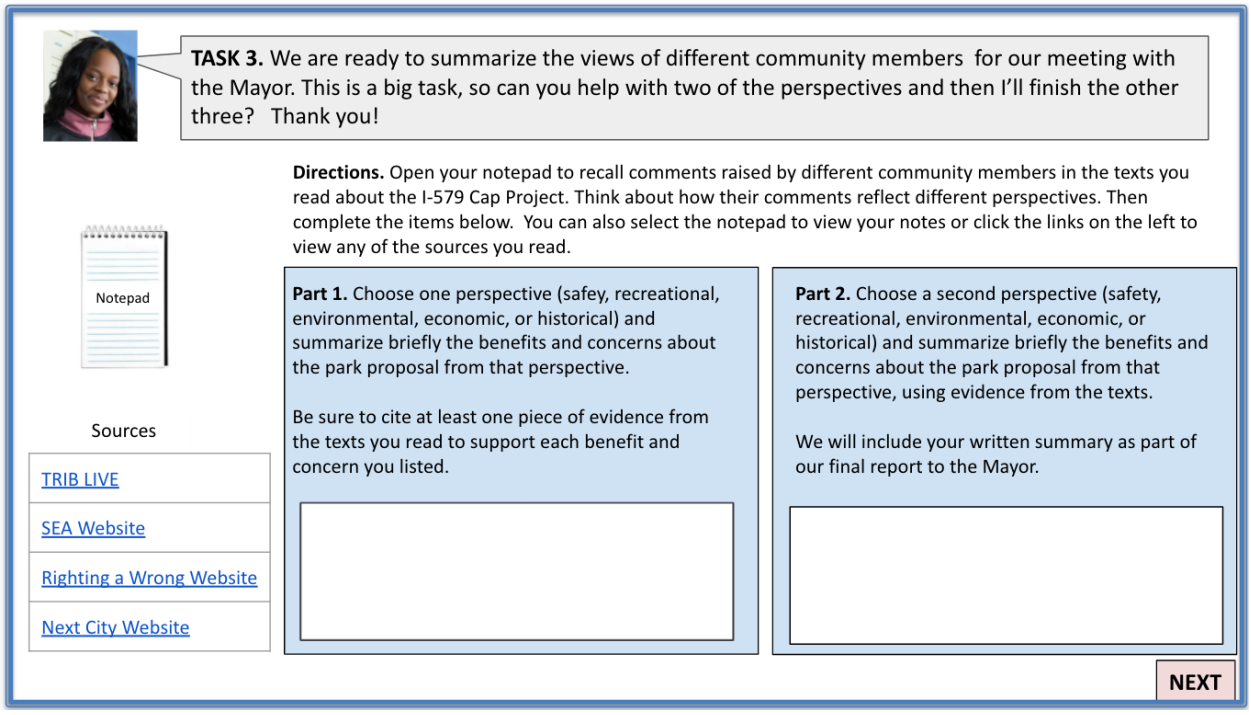


Exhibit C.33. This Final Use and Apply Item With Open Constructed Response Format Illustrates the Use of a Task Character (Motivational UDE) Who Reminds Students They Have Accomplished Their Goal and Validates the Test Taker's Role by Inviting Them to Use What They Learned and Apply That Understanding by Sharing Their Own Opinion

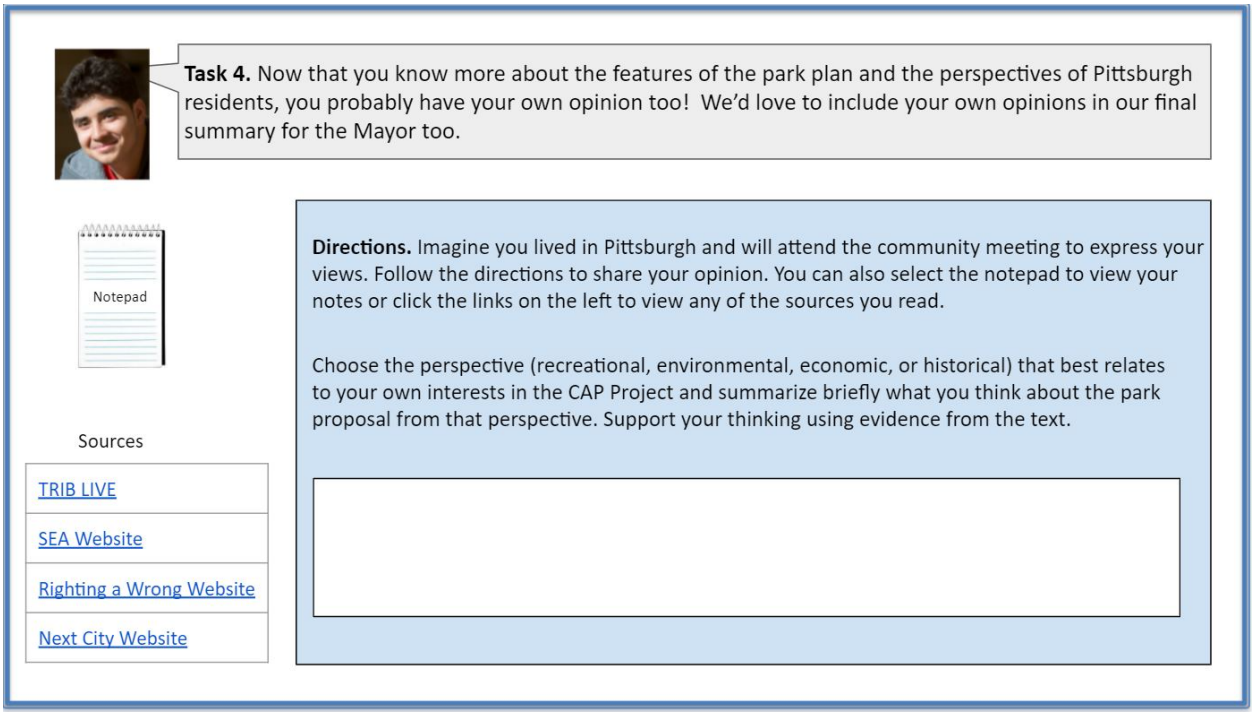
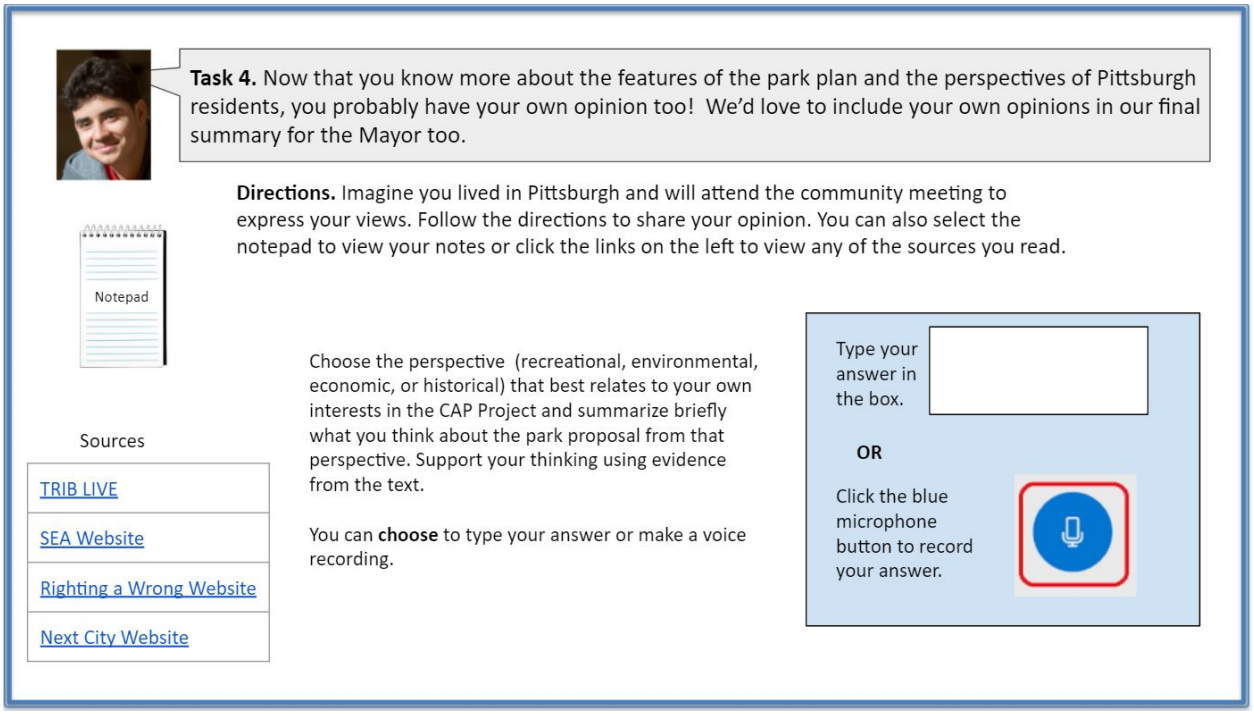


Exhibit C.34. This Alternative Format for the Final Use and Apply Item With open Constructed Response Format Illustrates the Use of Motivational UDEs for Two Purposes: A Task Character Who Invites Students' Own Opinion Paired With an Opportunity to Choose Their Preferred Format (Text or Audio) for Expressing Their Opinion



Performance Evidence and Indicators. *Scores from the Hill District block reveal what Grade 12 students can do when Reading to Solve a Problem in a social studies context. Ultimately, NAEP produces descriptions of what 12th graders (or subgroups of 12th graders) can do in each disciplinary reading context. Thus, from students' participation in the Hill District block (and other assessment blocks designated as RSP in social studies contexts), it is possible to characterize how well Grade 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.*

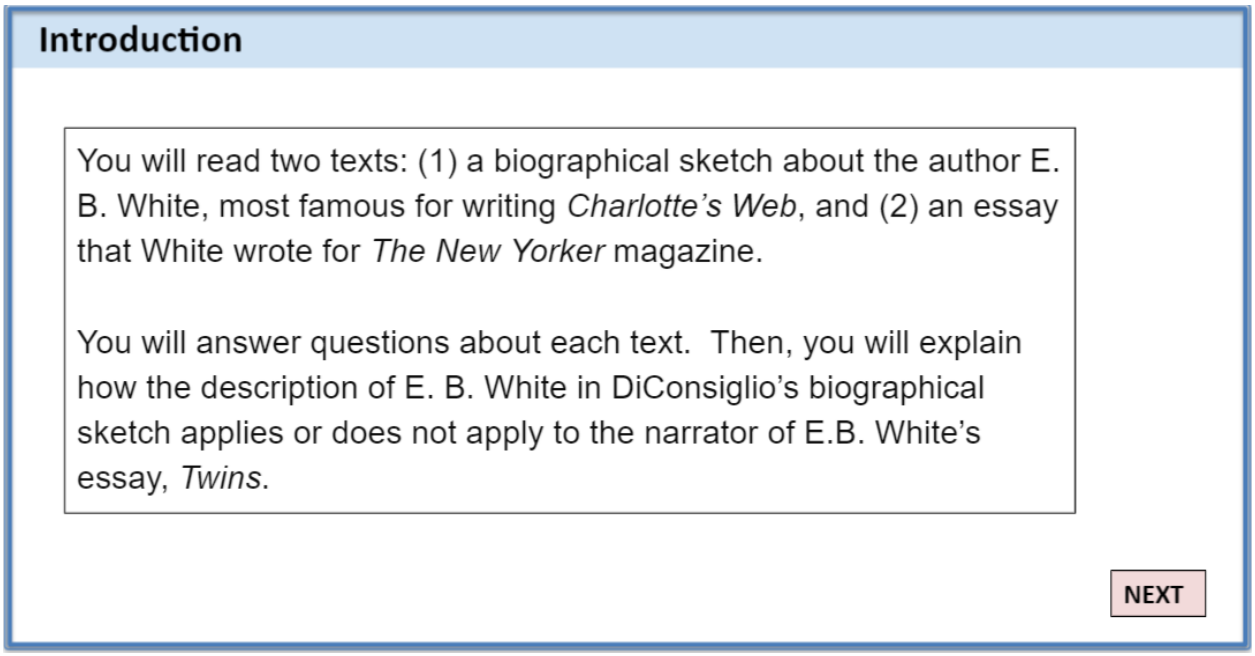
E. B. White, Grade 8

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 C.2. Here, students have more time to develop deep understanding of the texts. Tasks are relatively simple, 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 RDU 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 literature 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 C.44 and C.45), as they did in that assessment.

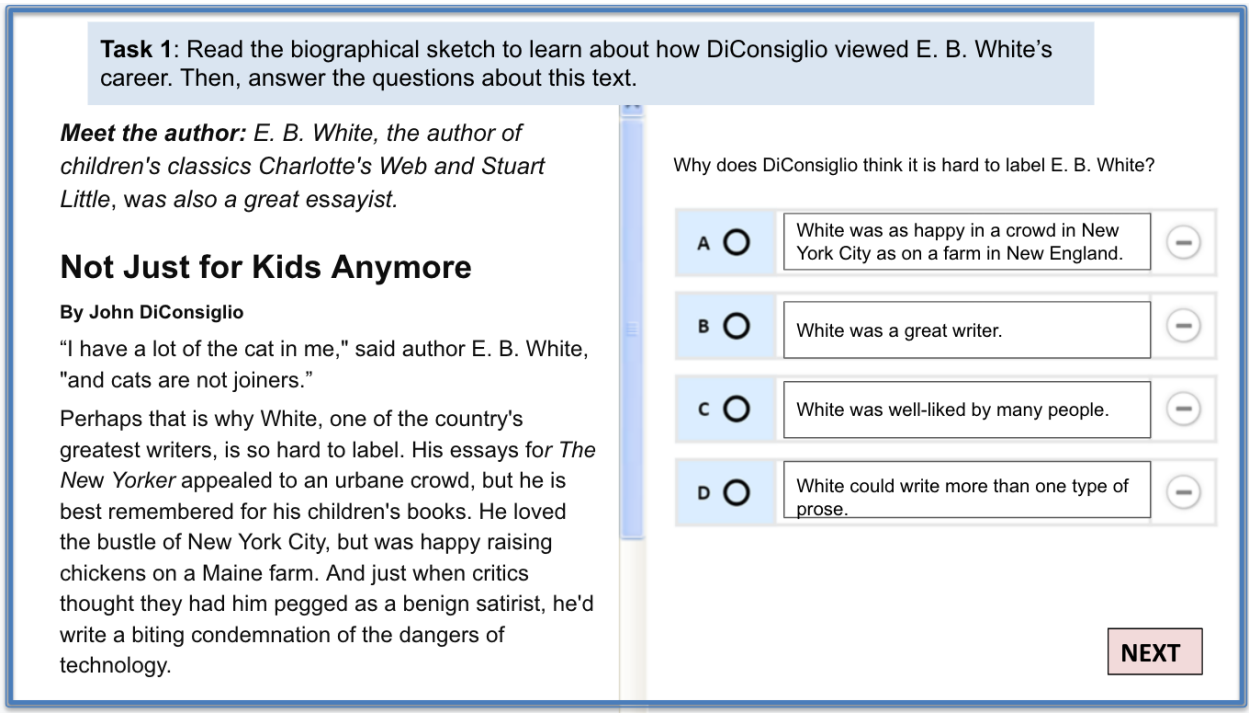
At the outset, readers are provided a specific reading purpose and are informed about the role (working on their own) they will be asked to assume during the block, which is composed of two common literature genres—a biographical sketch and a human-interest essay (see Exhibit C.35).

Exhibit C.35. Introduction to E. B. White



Task Components: Tasks, Text(s), and Items. This E. B. White block has three tasks that include (a) reading and answering questions about the biographical sketch *Not Just for Kids Anymore*, (b) reading and answering questions about the essay *Twins*, and (c) 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 Exhibit C.36, which shows Task 1.

Exhibit C.36. Introduction to the Grade 8 E. B. White Literature Block

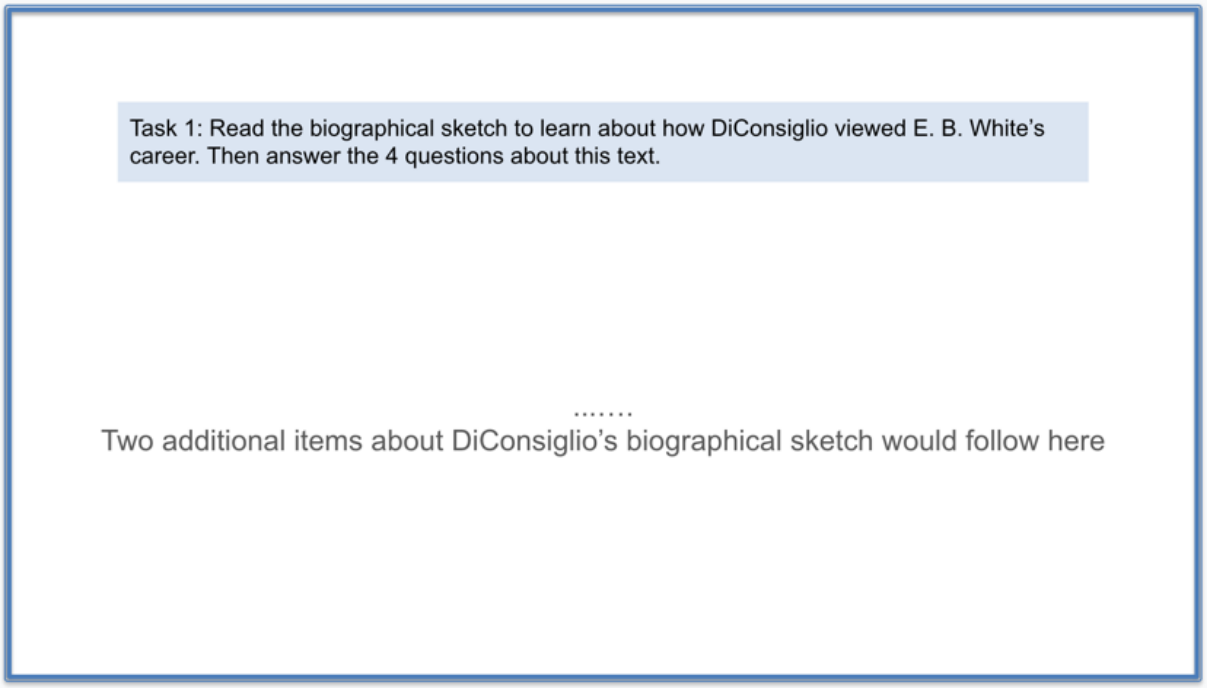


The comprehension items for Task 1 could help the reader develop understanding on segments of the biographical sketch that focus on characteristics of White that might be useful in Task 3 (see Exhibit C.37). Plausible segments for focus could be

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

In terms of UDEs, note that there is an informational 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 C.37. Task 1 Would Involve Additional Items



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 C.38–C.40). Candidates for items include the following:

- cases in which we get more than we bargained for and the sighting of the doe and her twins
- White's characterization of the doe being 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 an informational 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 C.38. Task 2 for the Grade 8 E. B. White Block Illustrating an Integrate and Interpret Item With a Short Constructed Response Item Format

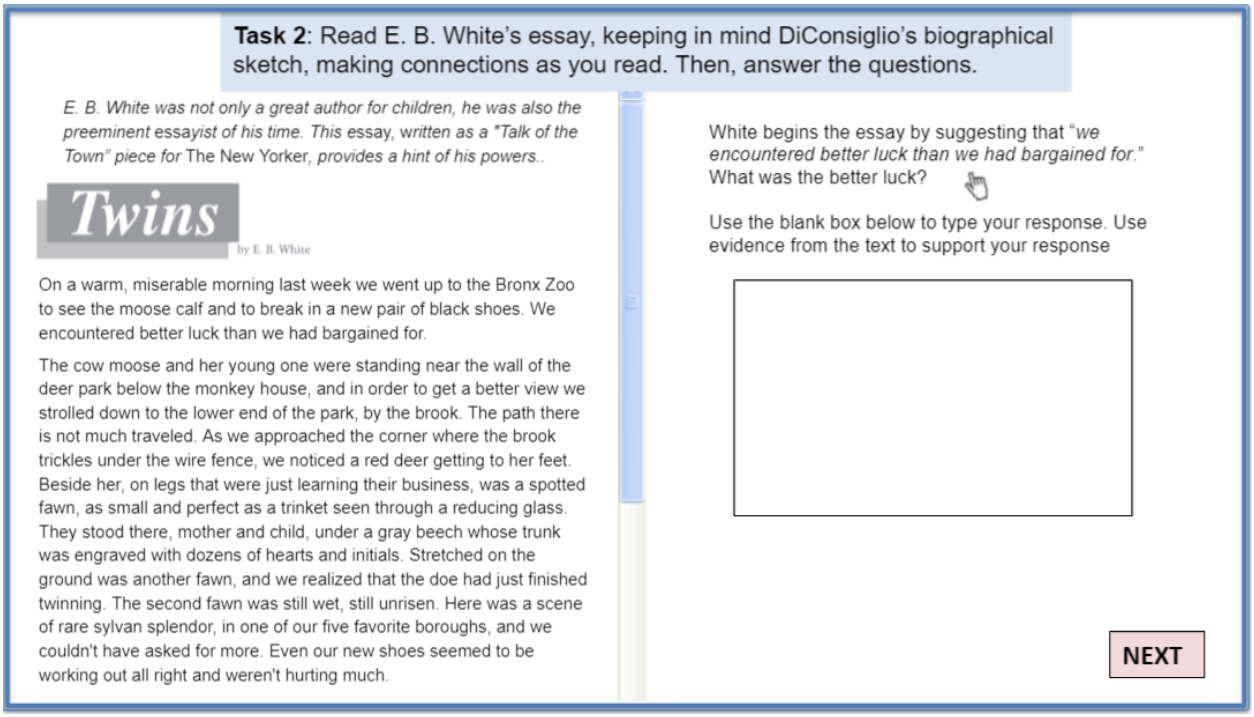


Exhibit C.39. Task 2 Continues for the Grade 8 E. B. White Block, Illustrating an Analyze and Evaluate Item With a Multiple Choice Item Response Format

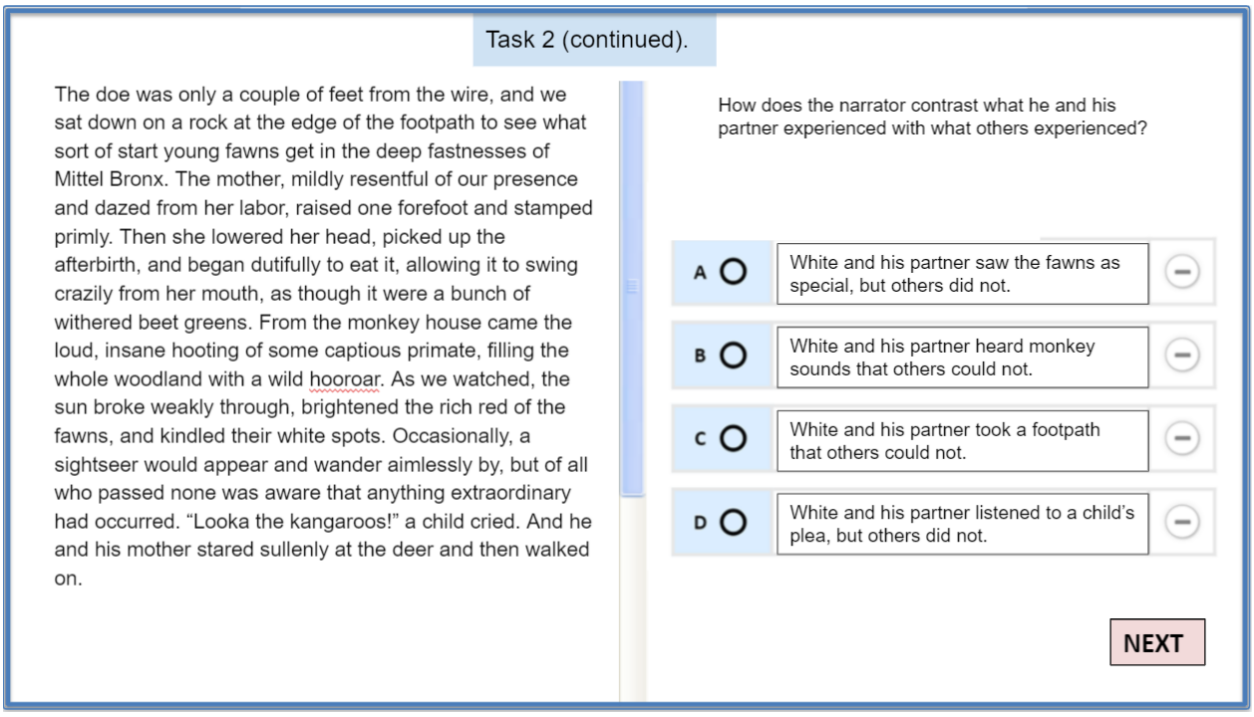
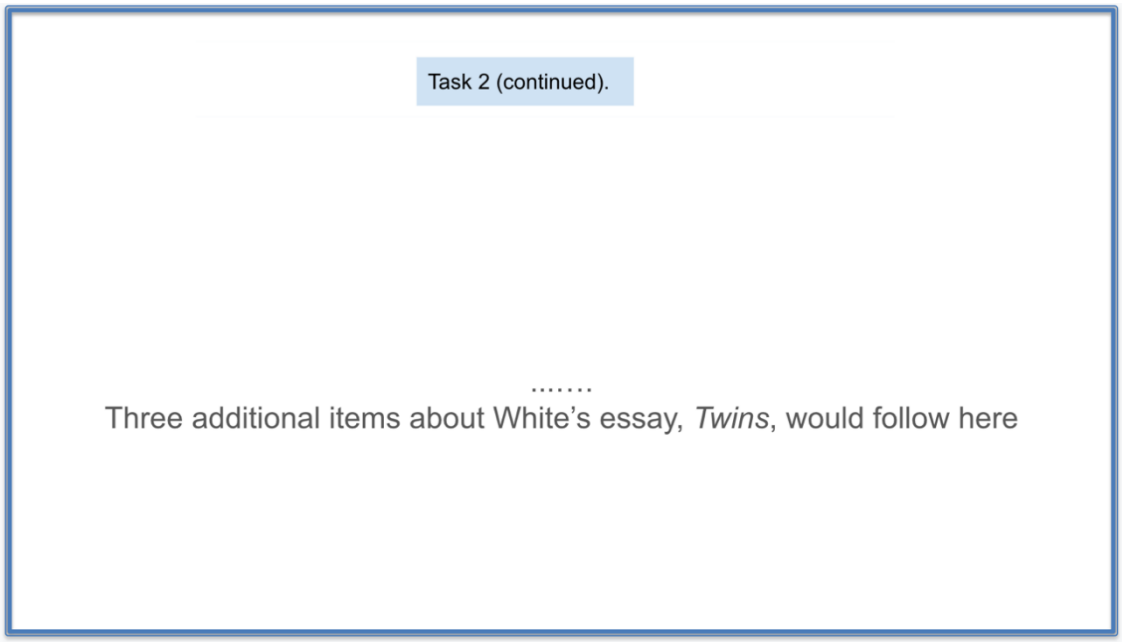
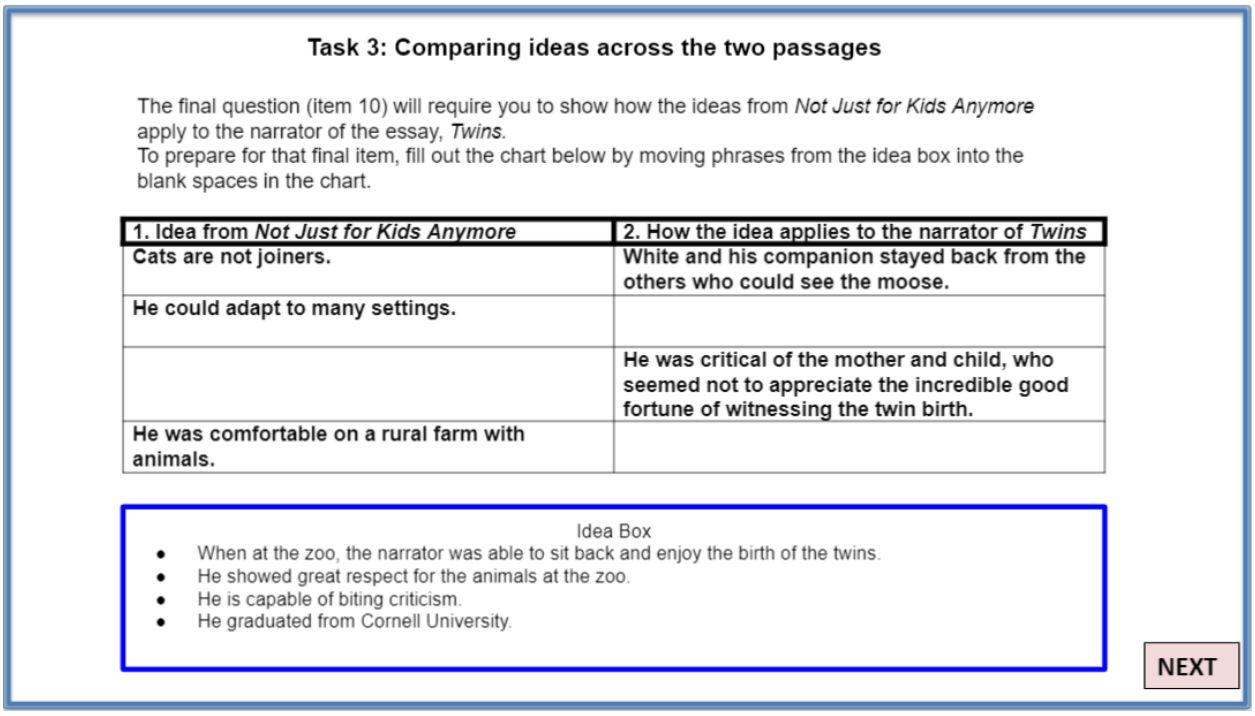


Exhibit C.40. Additional Items Accompany Task 2



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 C.41 and C.42), might be provided to assist students in organizing their response to a final Use and Apply extended constructed response item (see Exhibit C.43).

Exhibit C.41. 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 blank spaces in the chart.

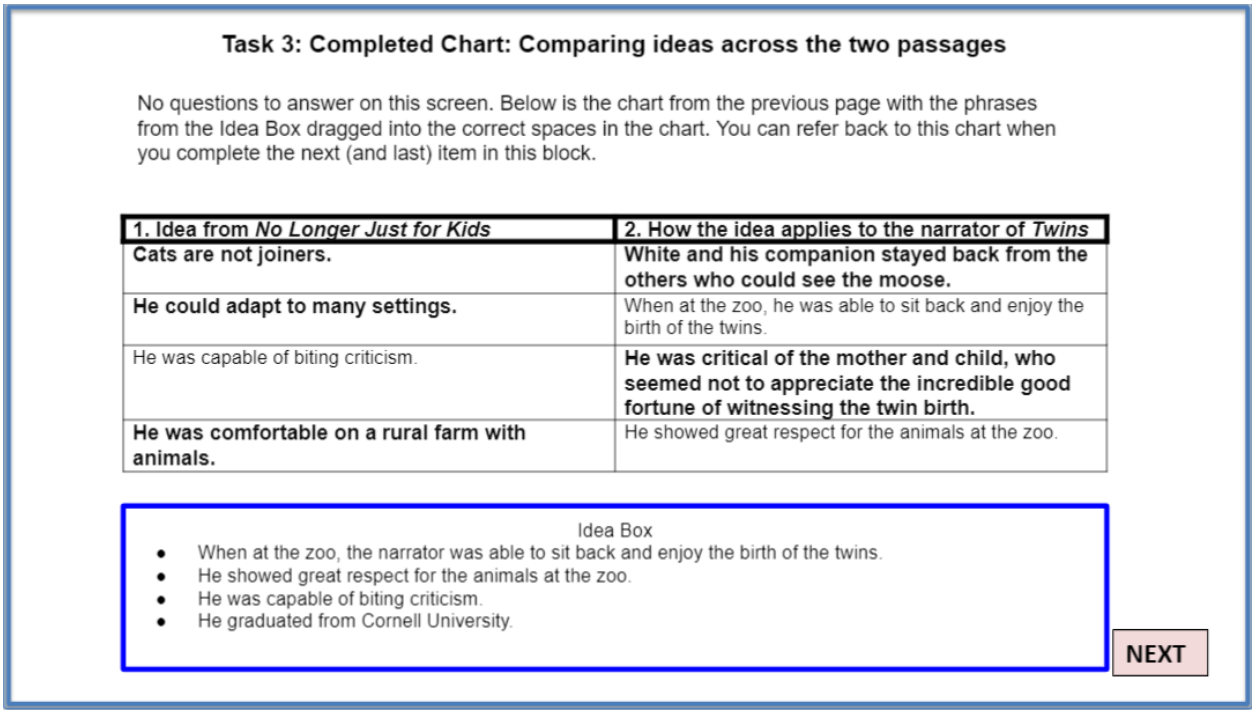
1. Idea from *Not Just for Kids Anymore*	2. How the idea applies to the narrator of *Twins*
Cats are not joiners.	**White and his companion stayed back from the others who could see the moose.**
He could adapt to many settings.	
	He was critical of the mother and child, who seemed not to appreciate the 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 C.42. Integrate and Interpret Item Illustrating Resetting of Item Responses From Prior Item



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

No questions to answer on this screen. Below is the chart from the previous page with the phrases from the Idea Box dragged into the correct spaces in the chart. You can refer back to this chart when you complete the next (and last) item in this block.

1. Idea from *No Longer Just for Kids*	2. How the idea applies to the narrator of *Twins*
Cats are not joiners.	**White and his companion stayed 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.

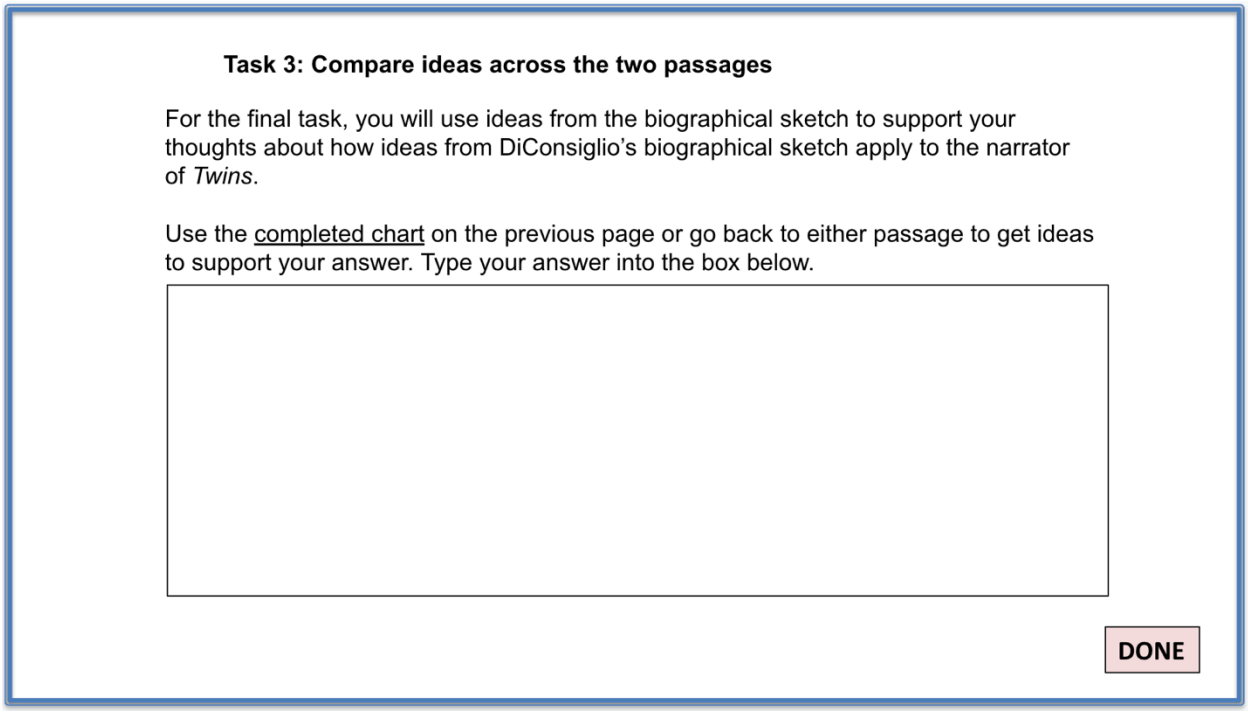
Idea Box

- 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 C.41), students receive feedback about how the chart might best have been completed in Exhibit C.42. The task-based UDE, called resetting, is provided so that students do not carry misconceptions into the final item in Exhibit C.43.

Exhibit C.43. A Final Use and Apply Item Asks Students to Use Ideas From the First Text to Develop Ideas About the Second Text



As suggested earlier, the E. B. White block sketch provides an example of how blocks might look under the auspices of the 2026 Assessment 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 C.44 and C.45.

Exhibit C.44. 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 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.

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 Stuart 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 Times 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

Exhibit C.45. The Second Text for the E. B. White Task: An Essay From the *New Yorker*

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.

Exhibit D.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.

Exhibit D.1. Example of a Matching Selected Response Item for a Webpage Text From PISA's Rapa Nui Block

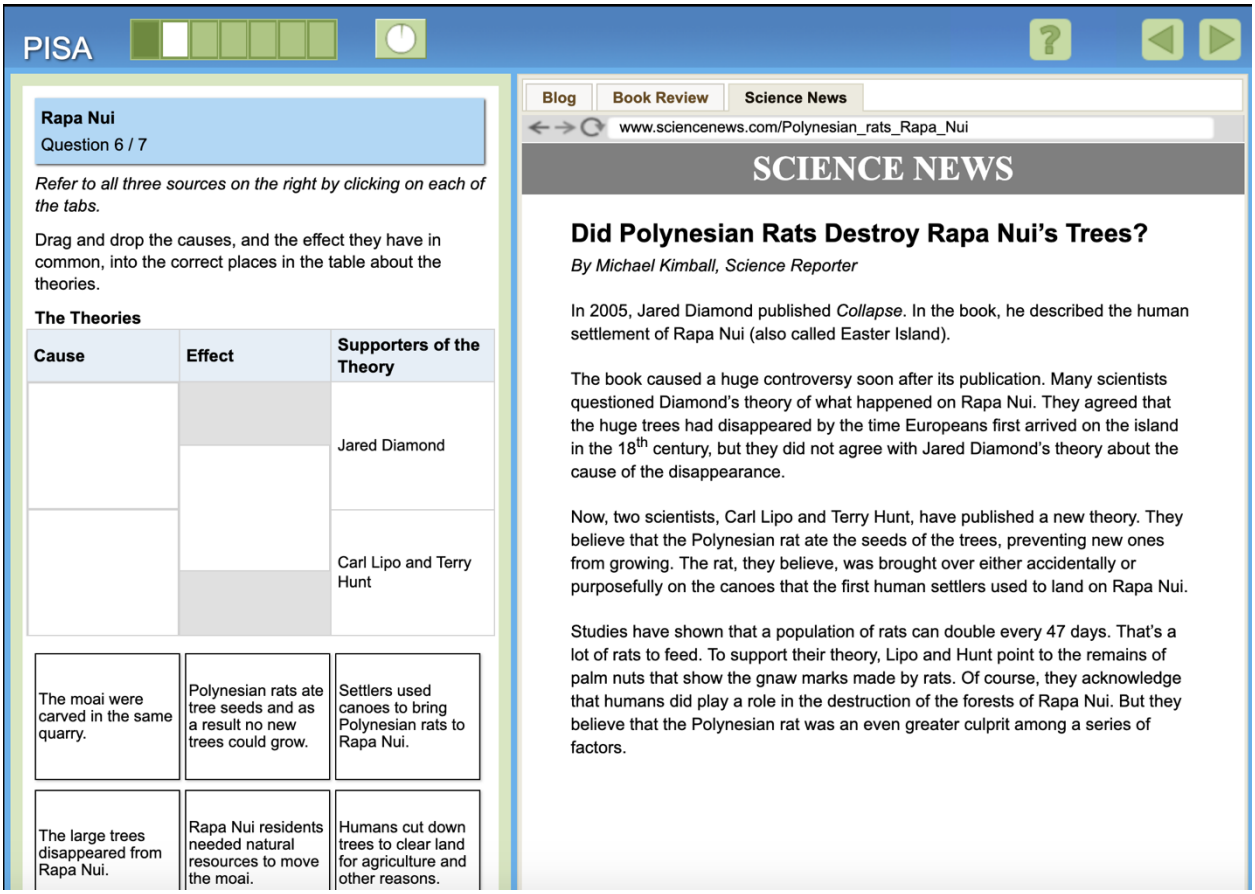


Exhibit D.2, from a PARCC Grade 12 task, illustrates a matching format. Students are asked to "drag" the ideas into the Venn diagram.

Exhibit D.2. Example of a Matching Selected Response Item From a Grade 12 PARCC Block

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)

"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)

"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)

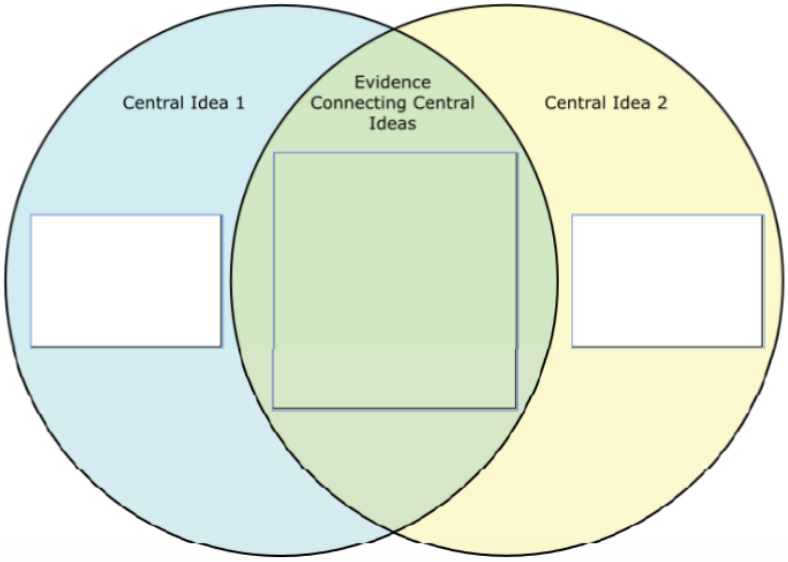


Exhibit D.3, from an ePIRLS task for Grade 4 students, illustrates a zones item format. The item asks students to "click on the website tab 'Rover Called Curiosity.'" To do so, students must click on the tab of the webpage with the same title. This item also illustrates the use of task characters, or avatars. An animated icon of a teacher shows "Mr. Webster," and another one shows the "Student," who is the test taker.

Exhibit D.3. Example of a Zones Selected Response Item Format and the Use of Task Characters From ePIRLS' Mars Block

Exhibit D.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 D.4. Example of a Grid Selected Response Item From PISA's Rapa Nui Block

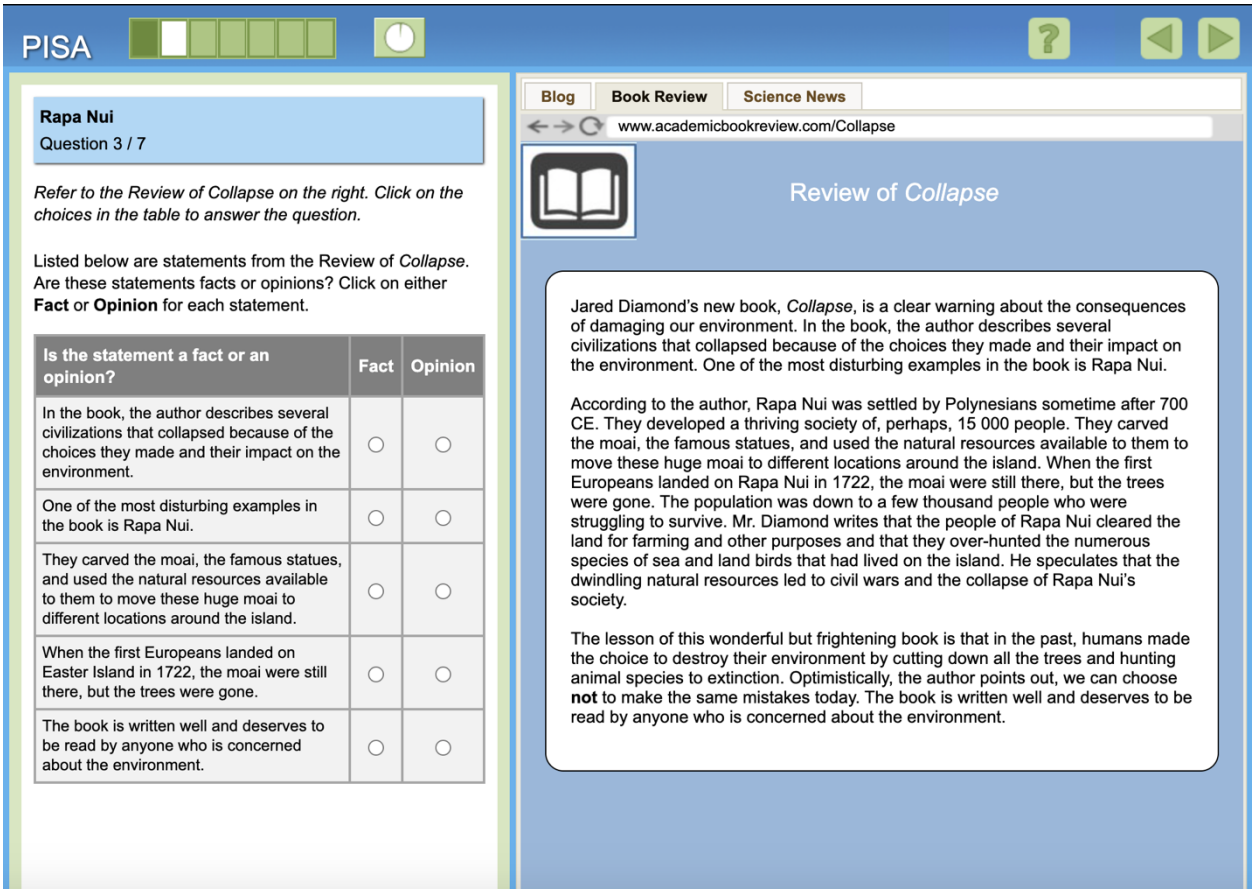


Exhibit D.5, from ePIRLS' assessment for Grade 4 students, provides an example of the use of a zones item format. Here, students are asked to "click on the link that is most likely" to have the requested information—in this case, "information about the life and achievements of Doctor Elizabeth Blackwell." This exhibit also illustrates the use of an internet text in the form of a search engine results page.

Exhibit D.5. Example of a Zones Item for an Internet Text From ePIRLS' "Elizabeth Blackwell" Block

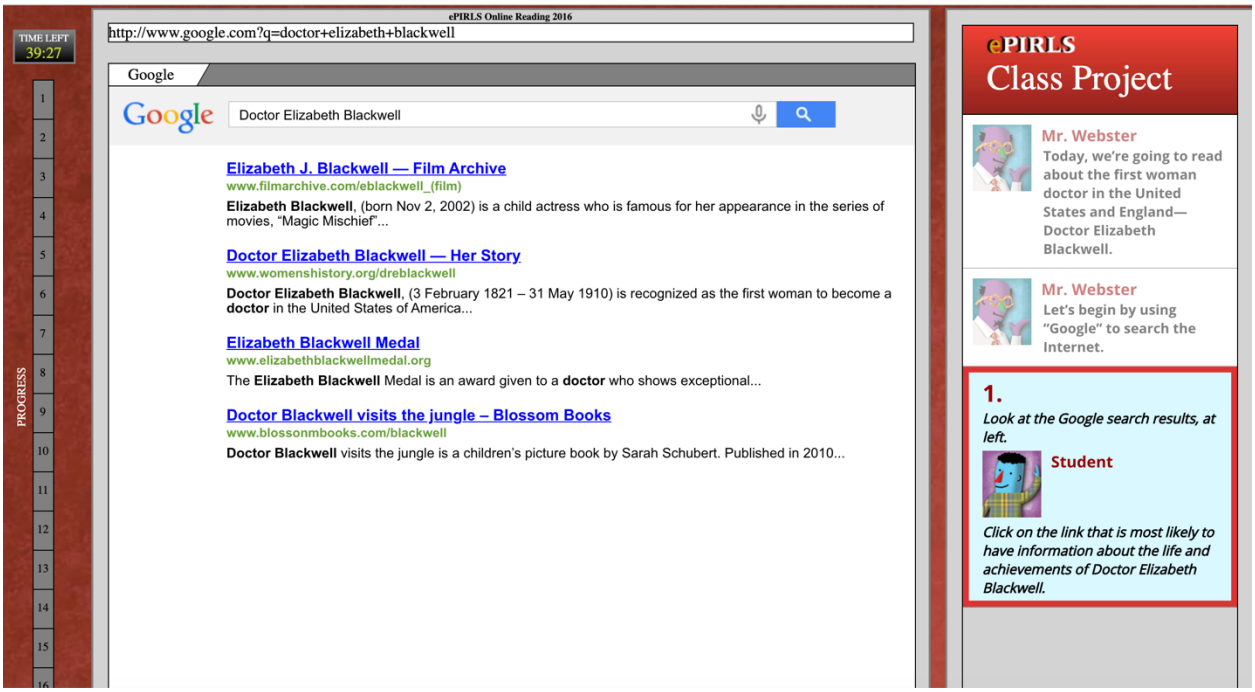


Exhibit D.6, from ePIRLS' assessment for Grade 4 students, asks students to use the digital diagram to answer questions by selecting responses from a drop-down menu (an in-line choice item). This item also collects **process data** of where on the graphic stimulus students click. While the clicks are 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 D.6. Example of an In-line Choice Item from ePIRLS' Mars Block That Also Collects Process Data on Where Students Click on the Webpage

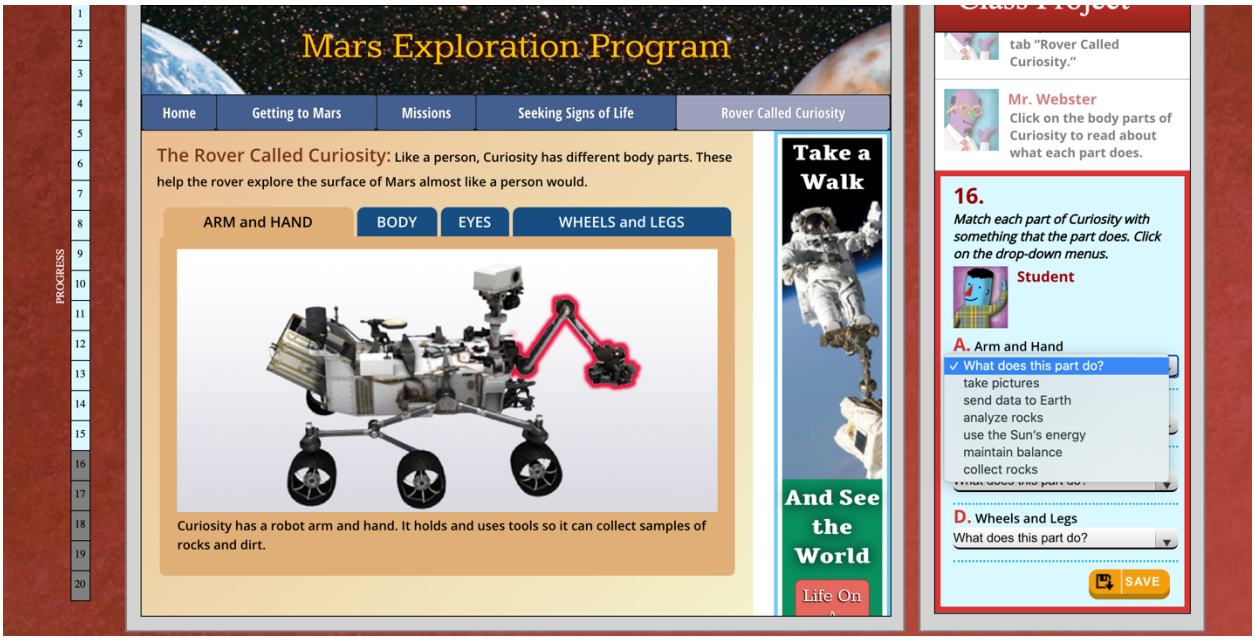


Exhibit D.7, from PISA's Reading Literacy test for 15-year-olds, illustrates a short constructed response. Here, students are given a small text box and asked to write about a key difference they read about in the approach taken to two different conservation programs.

Exhibit D.7. Example of a Short Constructed Response Item From PISA's Galapagos Islands Block

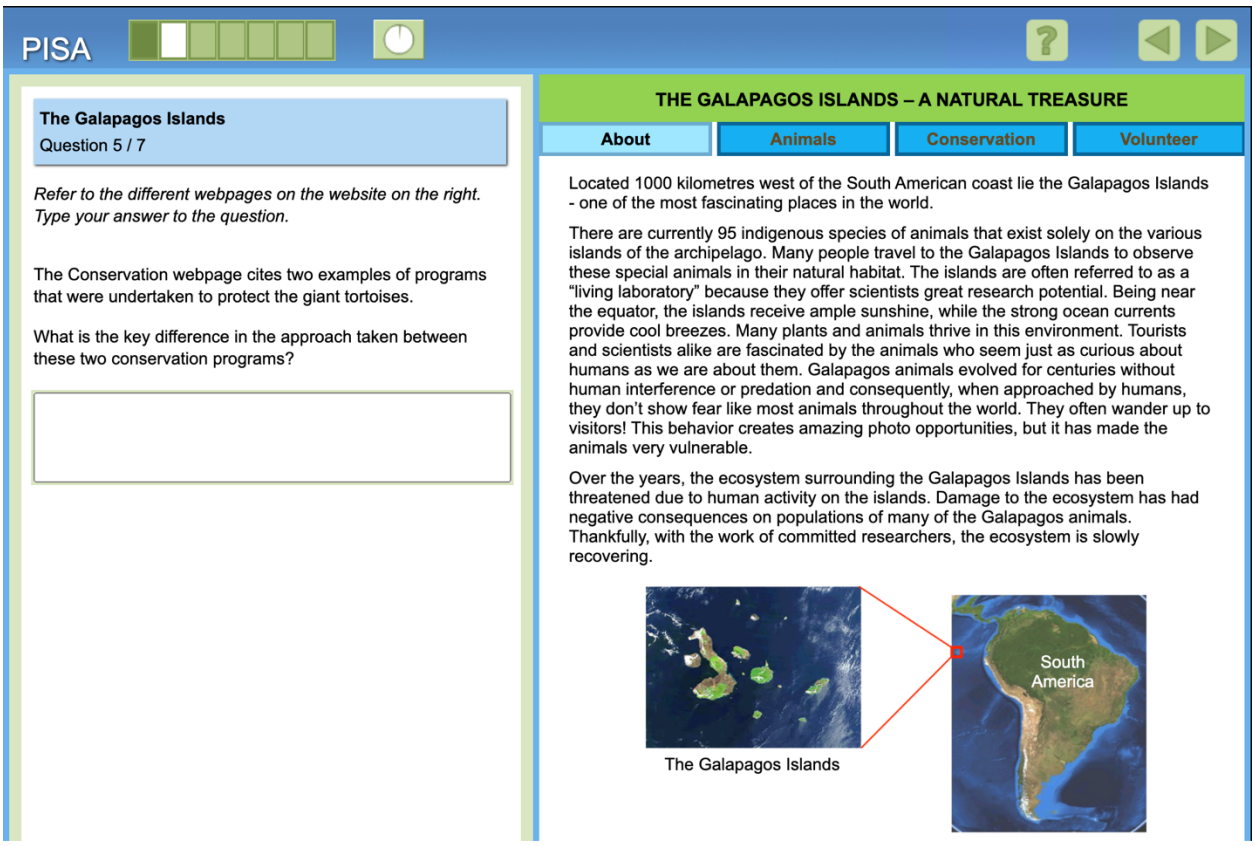


Exhibit D.8, from ePIRLS for Grade 4 students, illustrates the use of a fill-in-the-blank item response format for a digital website text that is a graphic. Here, students are asked to use the graphic to identify the "names of the three planets between Mars and the Sun." To give their answers, students type each name ("Mercury," "Venus," and "Earth") into three separate text fields.

Exhibit D.8. Example of a Fill-in-the-Blank Item Response Format From ePIRLS' Mars Block

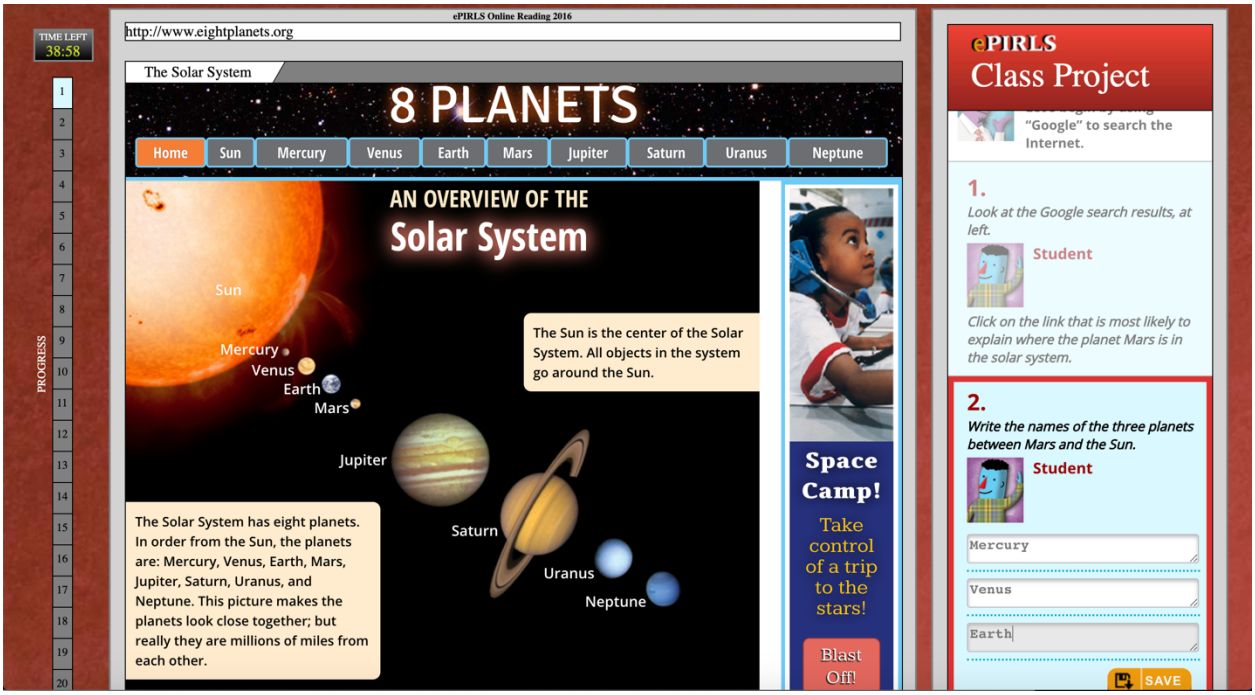


Exhibit E.1, from PISA's Reading Literacy test for 15-year-olds, illustrates how readers are situated, at the beginning of the block, within a specific reading purpose: to conduct research on the history of Rapa Nui in order to prepare for a lecture at a local library. This example also illustrates an informational 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 E.1. Example of a Specific Reading Purpose and an Informational UDE From PISA's Rapa Nui Block

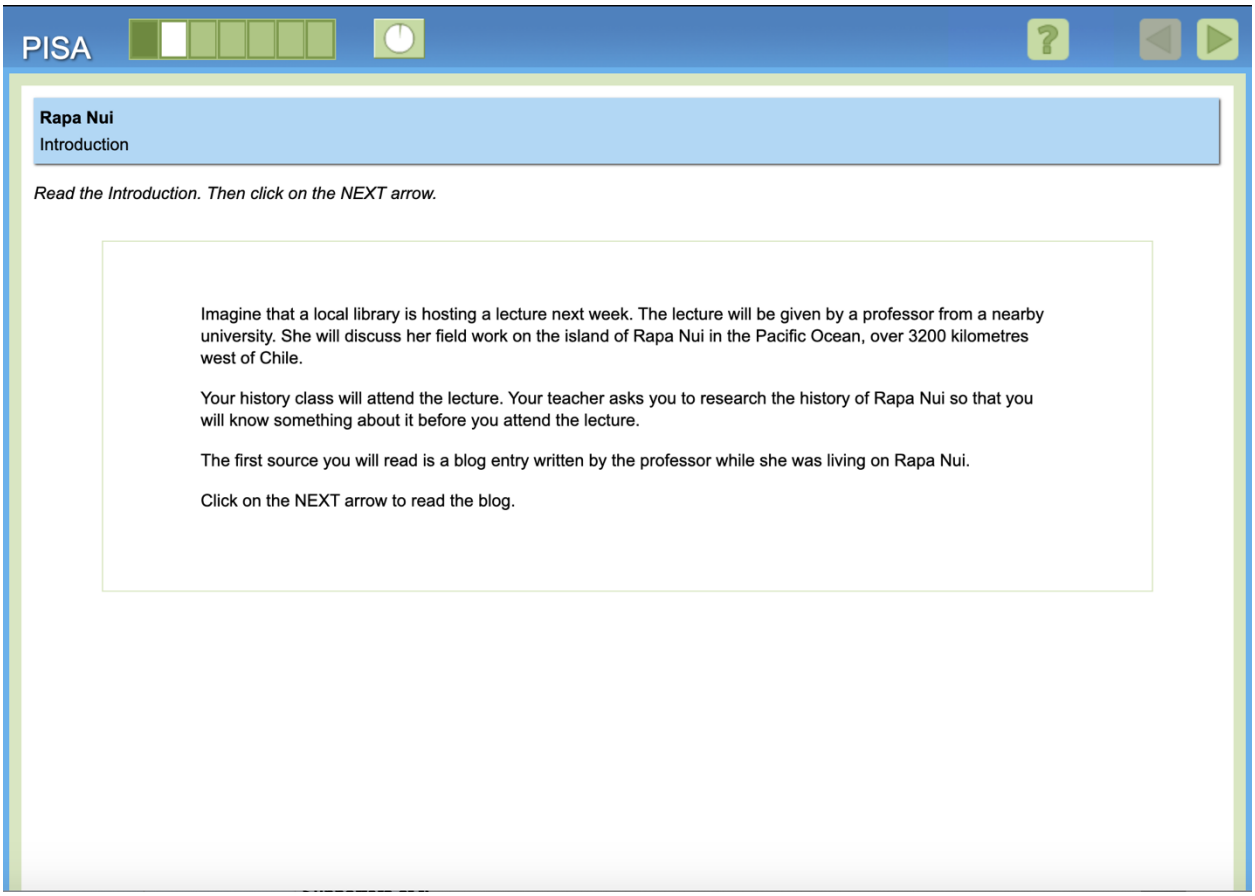


Exhibit E.2, from the Smarter Balanced test for Grade 8 students, illustrates a task-based UDE in the form of scoring criteria and steps for writing an explanatory article. Additionally, the example illustrates the use of an extended constructed response item in the form of what would be a Use and Apply Comprehension Target in the 2026 NAEP Reading Assessment.

Exhibit E.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.

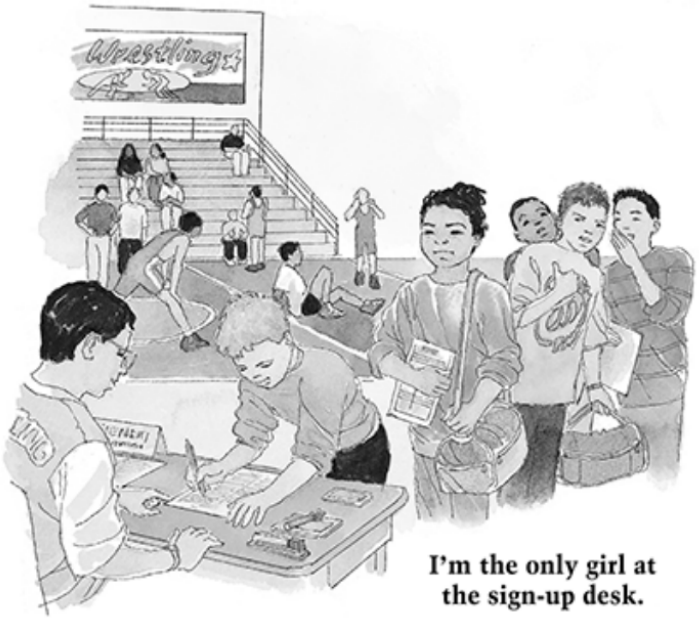
Word-processing tools and spell-check are available to you.

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 E.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 an informational UDE because they introduce the text by offering key plot information (a girl standing in line among only boys).

Exhibit E.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 E.4, from a NAEP Grade 4 block, illustrates two informational UDEs. The first informational 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 informational UDE appears in the form of a vocabulary pop-up box defining the Turkish word "akche."

Exhibit E.4. Example of Two Informational UDEs From NAEP's "Five Boiled Eggs" Block

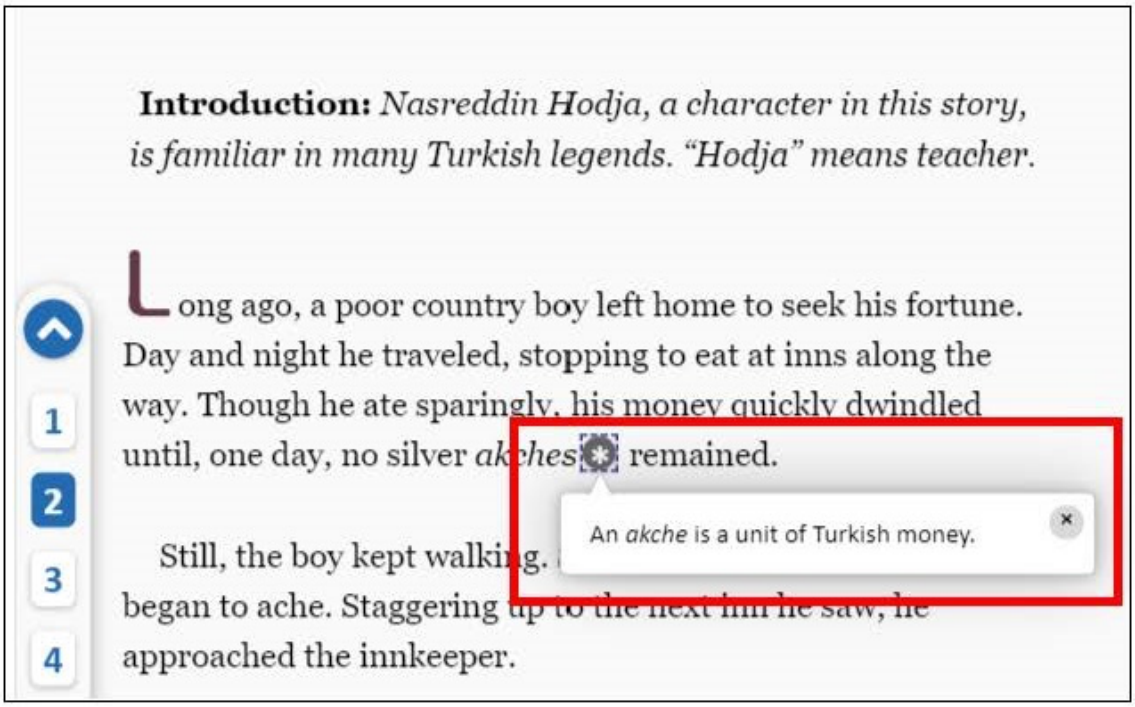
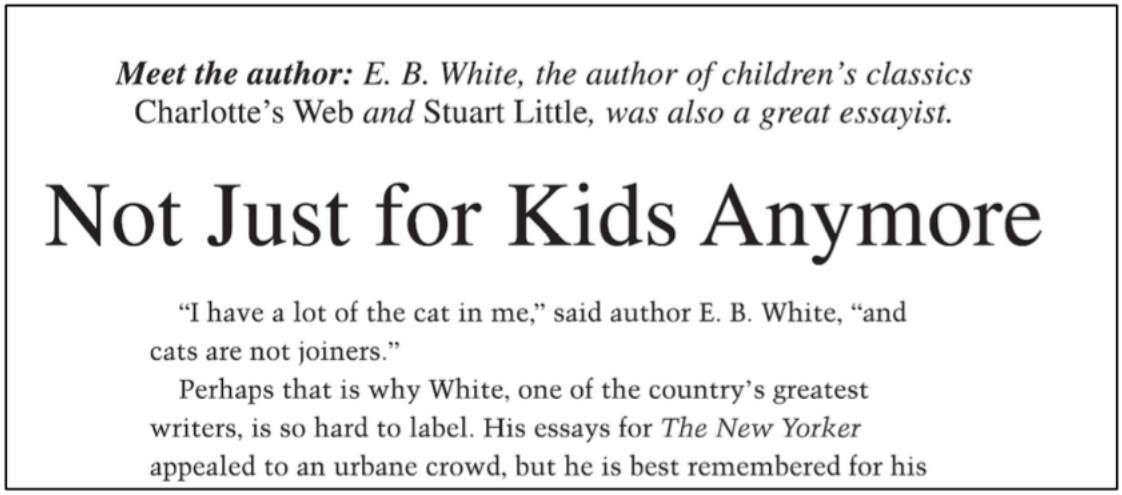


Exhibit E.5 illustrates two different written introductions, one for each of two texts. In Example 1, an informational UDE appears in the form of an introduction to an article about the writer E. B. White. In Example 2, an informational UDE appears in the form of an introduction to an essay by E. B. White, which explains that the author of the essay is also a children's author.

Exhibit E.5. Two Examples of Informational UDEs in the Form of Passage Introductions From a Released NAEP 2019 Block on E. B. White

Example 1



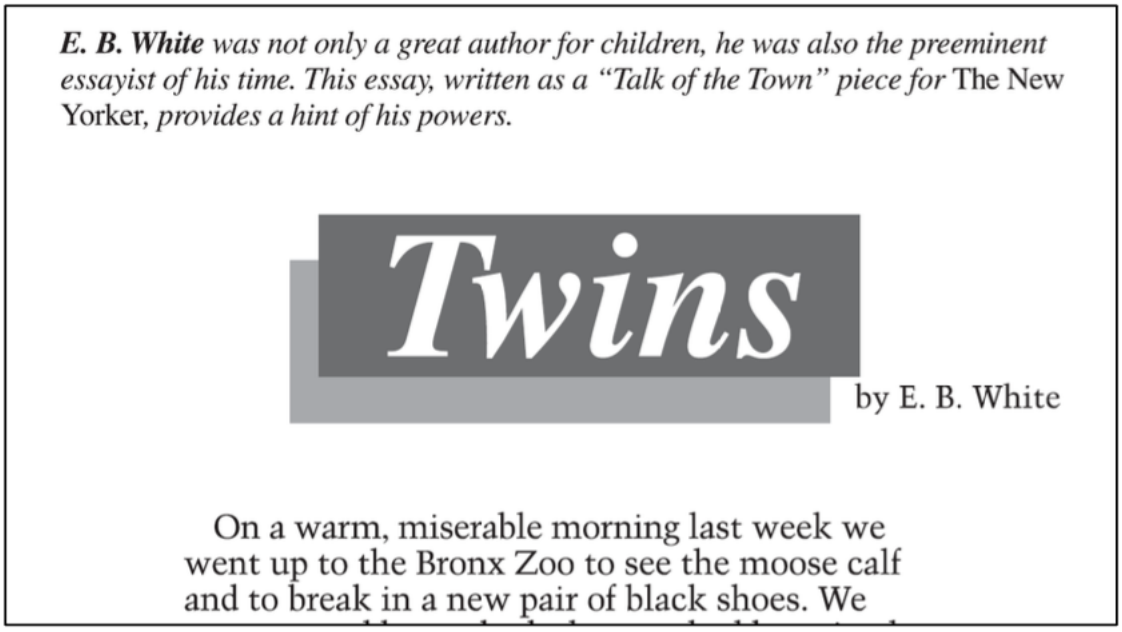
Meet the author: *E. B. White, the author of children's classics* Charlotte's Web *and* Stuart Little*, was also a great essayist.*

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 2



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

Exhibit E.6, from Michigan's reading assessment for Grade 4 students, illustrates three informational 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.

Exhibit E.6. Example of Three Informational 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.

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). Routledge.

Alexander, P. A. (2012). The Disciplined Reading and Learning Research Laboratory. Reading into the future: Competence for the 21st century. *Educational Psychologist*, *47*(4), 259–280.

Alexander, P. A., & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Vol. III* (pp. 285–310). Lawrence Erlbaum Associates.

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

American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Author.

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). 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. 1* (pp. 255–291). Routledge.

Appleman, D. (2017). Literary theory in the secondary school. In *English Language Arts Research and Teaching* (pp. 172–188). Routledge.

Armbruster, B. B., & Anderson, T. H. (1985). Producing 'considerate' expository text: or easy reading is damned hard writing. *Journal of Curriculum Studies, 17*(3), 247–274. DOI: 10.1080/0022027850170303

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). Lawrence Erlbaum Associates.

Barron, N. (2015). *Words onscreen: The fate of reading in a digital world.* Oxford University Press.

Beach, R., & Castek, J. (2016). 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). IGI Global.

Bergner, Y., & von Davier, A. A. (2019). 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. 5* (pp. 79–98). Routledge.

Bråten, I., Stadtler, M., & Salmerón, L. (2018). The role of sourcing in discourse comprehension. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *Handbook of discourse processes*. Routledge.

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.

Buehl, D. (2017). *Developing readers in the academic disciplines*. Stenhouse Publishers.

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.). Heinemann.

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. 5* (pp. 237–260). Routledge.

Charap, L. G. (2015). Assessing historical thinking in the redesigned advanced placement United States history course and exam. In K. Ercikan & P. Seixas (Eds.), *New directions in assessing historical thinking* (pp. 159–170). Routledge.

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). Guilford.

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

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.

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *102*(3), 687–700. https://doi.org/10.1037/a0019452

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). Harper Collins.

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). Lawrence Erlbaum Associates.

de Jong, T. (2006). Scaffolds for scientific discovery learning. In J. Elen & D. Clark (Eds.), *Handling complexity in learning environments: research and theory*, (pp. 107–128). Elsevier Science.

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

Durán, R. P. (2006, December). *State implementation of NCLB policies and interpretation of the NAEP performance of English Language Learners*. American Institutes for Research.

Durlak, J., Domitrovich, C., Weissberg, R., & Gullotta, T. (Eds.). (2015). *Handbook of social and emotional learning: Research and practice*. Guilford Press.

Eccles, J., O'Neill, S., & Wigfield, A. (2005). Ability self-perceptions and subjective task values in adolescents and children. DOI: 10.1007/0-387-23823-9_15.

Educational Testing Service. (2019, January 22). *NAEP Reading special study: Scenario-based task (SBT) and discrete task (DT) report. 2018 special study grades 4, 8, 12. Deliverable in response to ID Task 9.2.1.* Author.

Every Student Succeeds Act (ESSA), 20 U.S.C. § 6301 (2015). https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf

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., & Schleppegrell, M. J. (2010). Disciplinary literacies across content areas: Supporting secondary reading through functional language analysis. *Journal of Adolescent & Adult Literacy, 53*, 587–597.

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

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

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

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

Giroux, C. S., & Moje, E. B. (2017). Learning from the professions: Examining how, why, and when engineers read and write. *Theory Into Practice*, *56*(4), 300–307.

Goldman, S. R., & Bisanz, G. L. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 19–50). Lawrence Erlbaum.

Goldman, S. R. (2018). Discourse of learning and the learning of discourse. *Discourse Processes*, *55*(5–6), 434–453.

Goldman, S., Britt, M. A., Brown, W., Cribb, G., George, M., Greenleaf, C., Lee, C. D., Shanahan, C., & Project READI. (2016). Disciplinary literacies and learning to read for understanding: A conceptual framework for disciplinary literacy. *Educational Psychologist, 51*(2), 219–246. http://dx.doi.org/10.1080/00461520.2016.1168741.

Goldman, S. R., Greenleaf, C., Yukhymenko-Lescroart, M., Brown, W., Ko, M.-L. M., Emig, J. M., George, M., Wallace, P., Blaum, D., & Britt, M. A. (2019). Explanatory modeling in science through text-based investigation: Testing the efficacy of the Project READI intervention approach. *American Educational Research Journal*, *56*(4), 1148–1216.

Goldman, S.R., & Pellegrino, J.W. (2015). Research on learning and instruction: Implications for curriculum, instruction, and assessment. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 33–41. doi:10.1177/2372732215601866.

Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal*, *115*(2), 210–229.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395.
Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*(1), 5–26.
Guthrie, J. T., & Humenick, N. M. (2004). Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 329–354). Paul H Brookes Publishing Co.
Guthrie, J. T., & Klauda, S. L. (2015). Engagement and motivational processes in reading. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 41–53). Routledge.
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., 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). Springer Science.
Hartman, D. K., Morsink, P. M., & Zheng, J. (2010). From print to pixels: The evolution of cognitive conceptions of reading comprehension. In D. J. Leu (ed.), *The new literacies: Multiple perspectives on research and practice* (p. 131–164). Guilford Press.
Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge University Press.
Heath, S. B. (2012). *Words at work and play: Three decades in family and community life*. Cambridge University Press.
Ho, A. (2017). *Advancing educational research and student privacy in the "Big Data" era*. Washington, DC: National Academy of Education.
Hopkins, M., Thompson, K. D., Linquanti, R., Hakuta, K., & August, D. (2013). Fully accounting for English learner performance: A key issue in ESSA reauthorization. *Educational Researcher*, *42*(2), 101–108.
International Test Commission. (2019). ITC guidelines for the large-scale assessment of linguistically and culturally diverse populations. *International Journal of Testing*, *19*(4), 301–336. DOI: 10.1080/15305058.2019.1631024
International Reading Association (IRA) & National Council of Teachers of English (NCTE) Joint Task Force on Assessment. (2010). *Standards for the assessment of reading and writing*. Authors.
Jaeger, R. M. (2003). *Reporting the results of the National Assessment of Educational Progress. NAEP Validity Study. Working Paper Series*. Retrieved from https://nces.ed.gov/pubs2003/200311.pdf.
Johnstone, C. (2003). *Improving validity of large-scale tests: Universal design and student performance. Technical Report 37.* National Center on Educational Outcomes, University of Minnesota.
Johnstone, C., Altman, J., & Thurlow, M. (2006). *A state guide to the development of universally designed assessments*. National Center on Educational Outcomes, University of Minnesota.

Johnstone, C., Altman, J., Thurlow, M., & Thompson, S. (2006). *A summary of research on the effects of test accommodations: 2002 through 2004. Technical Report 45.* National Center on Educational Outcomes, University of Minnesota.

Katz, M. L., Brynelson, N., & Edlund, J. R. (2019). "Enacting rhetorical literacies: The expository reading and writing curriculum in theory and practice." In D. Alvermann, N. Unrau, & M. Sailors (Eds.), *Theoretical models and processes of literacy* 7th ed., pp. 533–562). Routledge.

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). 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,* 391–398. doi: 10.3102/0013189X18777740

Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*. Naval Technical Training Command Millington, TN, Research Branch.

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

Kinzer, C. K., & Leander, K. M. (2003). Reconsidering the technology/language arts divide: Electronic and print-based environments. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook of research on teaching the English language arts* (pp. 546–565). Lawrence Erlbaum Associates.

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

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

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*. International Reading Association.

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

Lee, C. D. (2011). Education and the study of literature. *Scientific Study of Literature*, *1*(1), 49-58.

Lee, C. D. (2016a). 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. (2016b). 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* (pp. 286–304). Routledge.

Lee, C. D. (2017). An ecological framework for enacting culturally sustaining pedagogy. In D. Paris & H. S. Alim (Eds.), *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world* (pp. 261–273). Teachers College Press.

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: Vol. 5* (pp. 37–56). Routledge.

Lee, C. D., Goldman, S. R., Levine, S., & Magliano, J. P. (2016). Epistemic cognition in literary reasoning. In J. Green, W. Sandoval, & I. Braten (Eds.), *Handbook of epistemic cognition* (pp. 165–183). Routledge.

Lee, C. D., & Spratley, A. (2010). *Reading in the disciplines and the challenges of adolescent literacy.* Carnegie Foundation of New York.

Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science* (pp. 87–113). Routledge.

Linn, R. L., & Dunbar, S. B. (1992). Issues in the design and reporting of the National Assessment of Educational Progress. *Journal of Educational Measurement*, *29*(2), 177-194.

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). Literacy Research Association.

Marzano, R. J. (1988). *Dimensions of thinking: A framework for curriculum and instruction*. The Association for Supervision and Curriculum Development, 125 N. West St., Alexandria, VA 22314-2798.

McCarthy, K. S., & McNamara, D. S. (2021). The multidimensional knowledge in text comprehension framework. *Educational Psychologist*, *56*(3), 1–19.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes*, *22*(3), 247-288.

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

Meola, M. (2004). Chucking the checklist: A contextual approach to teaching undergraduates Web-site evaluation. *portal: Libraries and the Academy*, *4*(3), 331–344.

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

Mislevy, R. J. (2019). Advances in measurement and cognition. *The ANNALS of the American Academy of Political and Social Science*, *683*(1), 164–182.

Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent & Adult Literacy*, *52*(2), 96–107.

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

Moje, E. B., & Luke, A. (2009). Literacy and identity: Examining the metaphors in history and contemporary research. *Reading Research Quarterly, 44,* 415–437.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2019). *PIRLS 2021 assessment frameworks.* 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., & Hooper, M. (2017). *ePIRLS 2016 international results in online informational reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/pirls2016/international-results/

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. (NASEM, 2017). *Evaluation of the achievement levels for mathematics and reading on the National Assessment of Educational Progress.* The National Academies Press. https://doi.org/10.17226/23409.

National Academies of Sciences, Engineering, and Medicine. (NASEM, 2018). *How people learn II: Learners, contexts, and cultures*. The National Academies Press. https://doi.org/10.17226/24783.

National Academies of Sciences, Engineering, and Medicine. (NASEM, 2019). *Monitoring educational equity*. The National Academies Press. https://doi.org/10.17226/25389.

National Assessment of Educational Progress Authorization Act, Pub. L. 107-279, 3 U.S.C. §§ 301-305. (2002). https://www.govinfo.gov/content/pkg/PLAW-107publ279/pdf/PLAW-107publ279.pdf

National Assessment Governing Board. (1992). *Reading framework for the 1992 National Assessment of Educational Progress* (Report No. 92-5002). NAEP Reading Consensus Project. National Assessment Governing Board.

National Assessment Governing Board. (2009). *Reading framework for the 2009 National Assessment of Educational Progress.* National Assessment Governing Board and U.S. Department of Education.

National Assessment Governing Board. (2014). *NAEP testing and reporting on students with disabilities and English language learners* [Policy statement]. https://www.nagb.gov/content/dam/nagb/en/documents/policies/naep_testandreport_studentswithdisabilities.pdf.

National Assessment Governing Board. (2017). *Reading framework for the 2017 National Assessment of Educational Progress.* 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. (2018a). *Framework development policy statement*. Author, adopted March 3, 2018.

National Assessment Governing Board. (2018b). Developing student achievement levels for the National Assessment of Educational Progress: Policy statement. Author, adopted November 17, 2018. https://www.nagb.gov/content/dam/nagb/en/documents/policies/ALS-revised-policy-statement-11-17-18.pdf

National Assessment Governing Board. (2019a). *The National Assessment Governing Board charge to the Visioning and Development panels for the 2025 National Assessment of Educational Progress (NAEP) Reading Framework*. Author. https://www.nagb.gov/content/dam/nagb/en/documents/what-we-do/quarterly-board-meeting-materials/2019-02/09-naep-reading-framework-update.pdf

National Assessment Governing Board. (2019b). *Reading framework for the 2019 National Assessment of Educational Progress*. National Assessment Governing Board and U.S. Department of Education.
National Center for Education Statistics. (2018). Going digital: NAEP assessments for the future. *National Assessment of Educational Progress.* 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. (2019a). NAEP Accommodations Increase Inclusiveness. U.S. Department of Education, Institute of Education Sciences. Retrieved from https://nces.ed.gov/nationsreportcard/about/accom_table.aspx.
National Center for Education Statistics. (2019b). NAEP Report Card: 2019 Reading Assessment. https://www.nationsreportcard.gov/highlights/reading/2019/.
National Center for Education Statistics. (2020). Digitally based assessments. *National Assessment of Educational Progress.* U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/nationsreportcard/dba/.
National Center for Education Statistics. (2021). Overview (April 23, 2021) of the NAEP Reading Assessment and Projections. Extracted from the Material of the May 13–15 Quarterly Meeting of the National Assessment Governing Board.
National Council for the Social Studies. (2013). The College, Career, and Civic Life (C3) Framework for Social Studies State Standards: Guidance for Enhancing the Rigor of K-12 Civics, Economics, Geography, and History. Silver Spring, MD: Author.
National Governors Association Center for Best Practices and Council of Chief State School Officers. (NGA-CCSSO, 2010). *Common core state standards for English language arts and literacy in history/social studies & science*. http://www.corestandards.org/Standards/K12/.
O'Hallaron, C. L., Palincsar, A. S., & Schleppegrell, M. J. (2015). Reading science: Using systemic functional linguistics to support critical language awareness. *Linguistics and Education*, *32*, 55–67.
O'Reilly, T., Sabatini, J., & Wang, J. (2018). Using scenario-based assessments to measure deep learning. In K. Millis, D. Long, J.Magliano, & K. Weimer (Eds.), *Deep comprehension: Multi-disciplinary approaches to understanding, enhancing, and measuring comprehension* (pp. 197–208). Routledge.
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*, *30*(9), 1344–1351.
Organisation for Economic Cooperation and Development. (OECD, 2018). *PISA 2018 released field trial new reading items*. https://www.oecd.org/pisa/test/PISA-2018-Released-New-REA-Items.pdf
Organisation for Economic Cooperation and Development. (OECD, 2019). *PISA 2018 assessment and analytical framework.* OECD Publishing. https://doi.org/10.1787/b25efab8-en
Ostenson, J. (2014). Reconsidering the checklist in teaching internet source evaluation. *portal: Libraries and the Academy*, *14*(1), 33–50.
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* (pp. 135–165). 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.
PARCC/New Meridian. (2019). Released Items. Retrieved from https://resources.newmeridiancorp.org/released-items/
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). Lawrence Erlbaum Associates.
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). Teachers College Press.
Pearson, P. D., Hansen, J., & Gordon, C. (1979). The effect of background knowledge on young children's comprehension of explicit and implicit information. *Journal of Reading Behavior*, *11*, 201–210.
Pearson, P. D., & Hiebert, E. (2014). The state of the field. *The Elementary School Journal, 115*(2), 161–183.
Pearson, P. D., Palincsar, A. S., Biancarosa, G., & Berman, A. (2020). *Reaping the rewards of the Reading for Understanding Initiative.* National Academy of Education.
Pellegrino, J. W. (2013). Proficiency in science: Assessment challenges and opportunities. *Science* (American Association for the Advancement of Science), *340*(6130), 320–323. https://doi.org/10.1126/science.1232065
Pellegrino, J. W., Chudowsky, N., Glaser, R., & National Research Council. Committee on the Foundations of Assessment. (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press.
Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown & P. Hagoort (eds.), *The neurocognition of language* (pp. 167–208). Oxford University Press.
Phillips Galloway, E., Brown, J., & Uccelli, P. (2020). Broadening the lens on the science of reading: A multifaceted perspective on the role of academic language in text understanding. *Reading Research Quarterly*, *55*(1), 331–345. doi.org/10.1002/rrq.359
Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. Schunk & J. Meece (Eds.), *Student perceptions in the classroom: Causes and consequences* (pp. 149–183). Lawrence Erlbaum Associates.
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.
Rabinowitz, P. J. (1987). *Before reading: Narrative conventions and the politics of interpretation*. Cornell University Press.
RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension.* RAND. https://www.rand.org/pubs/monograph_reports/MR1465.html
Rogers, C. M., Lazarus, S. S., & Thurlow, M. L. (2016). *A summary of the research on the effects of test accommodations: 2013-2014*. NCEO Report 402. National Center on Educational Outcomes.

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. DOI:10.1080/15305058.2018.1551224

Scalise, K., & Gifford, B. (2006). Computer-based assessment in e-learning: A framework for constructing "intermediate constraint" questions and tasks for technology platforms. *Journal of Technology, Learning, and Assessment, 4*(6), 3–44.

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.

Schreiner, T. (2014). Using historical knowledge to reason about contemporary political issues: An expert novice study. *Cognition and Instruction, 32,* 314–352.

Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Harvard University Press.

Seixas, P. (2010). A modest proposal for change in Canadian history education. *International Review of History Education, 6,* 11–26.

Seixas, P., Gibson, L., & Ercikan, K. (2015). A design process for assessing historical thinking: The case of a one-hour test. In K. Ercikan & P. Seixas (Eds.), *New directions in assessing historical thinking* (pp. 102–116). Routledge.

Serafini, F., & Gee, E. (Eds.). (2017). *Remixing multiliteracies: Theory and practice from New London to new times*. Teachers College Press.

Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *The Elementary School Journal*, *115*(2), 184-209.

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.

Skerrett, A. (2020). Social and cultural differences in reading development: Instructional approaches, learning gains, and challenges. In E. B. Moje, P. Afflerbach, P. Enciso, & N. Lesaux (Eds.), *Handbook of reading research: Vol. 5* (pp. 328–344). Routledge.

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.

Smarter Balanced Assessment Consortium (SBAC). (2020). Sample items. Retrieved from http://sampleitems.smarterbalanced.org

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, *328*(5977), 450–452.

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). 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. (2015). Cognitively-based assessment of research and inquiry skills: Defining a key practice in the English language arts. *ETS Research Report Series, 2015,* 1–55.

Stenner, A. J. (1996). Measuring reading comprehension with the Lexile framework. Fourth North American Conference on Adolescent/Adult Literacy. Washington, D.C.

Street, B. V. (1984). *Literacy in theory and practice.* 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.

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

Thompson, S., Thurlow, M., & Malouf, D. B. (2004). Creating better tests for everyone through universally designed assessments. *Journal of Applied Testing Technology*, *6*(1), 1–15.

Uegaki, C., & Leng, Q. (2014). *Hana Hashimotosixth violin.* Kids Can Press.

Valencia, S. W., Wixson, K. K., Kitmitto, S., & Blankenship, C. (July 2019). *A comparison of NAEP Reading and NAEP Writing Assessments with current generation state assessments in English language arts: Expert judgment study.* American Institutes for Research.

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. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, *328*, 453–456.

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* (pp. 123–139). IAP Information Age Publishing.

van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*(5-6), 360–372.

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. Graesser (Eds.), *Models of understanding text*, (pp. 165–187). Lawrence Erlbaum Associates.

van Drie, J., & van Boxtel, C. (2008). Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review*, *20*(2), 87-110.

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.

Vorstius, C., Radach, R., Mayer, M. B., & Lonigan, C. J. (2013). Monitoring local comprehension monitoring in sentence reading. *School Psychology Review*, *42*(2), 191–206.

Wigfield, A., & Wentzel, K. R. (2007). Introduction to motivation at school: Interventions that work. *Educational Psychologist, 42*(4), 191–196.

Willingham, D. T. (2006). How knowledge helps: It speeds and strengthens reading comprehension, learning–and thinking. *American Educator*, *30*(1), 30–35.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of educational Psychology*, *83*(1), 73.

Wineburg, S., & McGrew, S. (2017). Lateral reading: Reading less and learning more when evaluating digital information. Stanford Histoy Education Group Working Paper No. 2017-A1. Retrieved from https://ssrn.com/abstract=3048994.

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

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* (p. 3–12). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199958450.013.0001.

Zhang, M., & Quintana, C. (2012). Scaffolding strategies for supporting middle school students' online inquiry processes. Computers & Education, 58(1), 181–196.