Quarterly Meeting of August 5-6, 2021

Via Zoom **and** Tysons Corner Ritz-Carlton 1700 Tysons Blvd McLean, VA 22102 (844) 631-0595



AGENDA

Thursday, August 5: 10:30 am – 5:30 pm (EDT)

10:30 am – 12:00 pm	Executive Committee Meeting (see separate agenda) Haley Barbour, Chair		
12:00 – 1:15 pm	Lunch Break		
1:15 – 1:30 pm	Welcome		
	Approval of August 2021 Agenda		
	Approval of May 2021 Minutes		
	Haley Barbour, Chair		
1:30 – 1:45 pm	Executive Director's Update		
	Lesley Muldoon, Executive Director		
1:45 – 2:15 pm	National Center for Education Statistics Update		
	Peggy Carr, Acting Commissioner, National Center for Education Statistics		
2:15 – 2:45 pm	Committee Updates		
	Dana Boyd, Assessment Development Committee		

	Gregory Cizek, Committee on Standards, Design and Methodology Jim Geringer, Nominations Committee Tonya Matthews, Reporting and Dissemination Committee
	ACTION: Release Plans for NAEP Long-Term Trend and High School Transcript Study
	Alice Peisch, Executive Committee
	 ACTION: Nomination for Board Vice Chair for the Term October 1, 2021 – September 30, 2022
	ACTION: Change to Long-Term Trend Administration 2022
2:45 – 3:00 pm	Break
3:00 – 4:00 pm	ACTION: NAEP Reading Assessment Framework Haley Barbour
3:00 – 4:00 pm	ACTION: NAEP Reading Assessment Framework Haley Barbour Alice Peisch

Friday, August 6: 10:00 am – 4:00 pm (EDT)

10:00 – 11:30 am	NAEP Budget and Assessment Schedule (CLOSED) Peggy Carr, Acting Commissioner, National Center for Education Statistics Lesley Muldoon	
11:30 – 11:45 am	Break	
11:45 am – 1:15 pm	WORKING LUNCH: Briefing on Upcoming NAEP Releases (CLOSED) Linda Hamilton, National Center for Education Statistics Grady Wilburn, National Center for Education Statistics	
1:15 – 2:00 pm	Records Management Briefing and Discussion (CLOSED)	

	Jason M. Lautenbacher, Chief, Information Management Branch, U.S. Department of Education
2:00 – 2:15 pm	Break
2:15 – 3:30 pm	Across the Board: Understanding Recent NAEP Results Ebony Walton, National Center for Education Statistics
3:30 – 4:00 pm	Farewell Remarks Gregory Cizek Governor Jim Geringer

2021 - 2023 QUARTERLY BOARD MEETING DATES AND LOCATIONS

November 18 - 20, 2021	TBD
March 3 - 5, 2022	TBD
May 12 - 14, 2022	TBD
August 4 - 6, 2022	TBD
November 17 - 19 , 2022	TBD
March 3 - 4, 2023	TBD



FIFTH FLOOR BALLROOM LEVEL





Governing Board Members 2020 - 2021

Honorable Haley Barbour, Chair

BGR Group, Founding Partner Yazoo City, Mississippi

Representative 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

Representative 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

National Assessment Governing Board

Committee Structure (2020-2021)

Assessment Development Committee

Chair Vice Chair Dana Boyd Mark Miller Christine Cunningham Frank Edelblut Patrick Kelly Reginald McGregor Nardi Routten *Michelle Blair (Staff)*

Committee on Standards, Design and Methodology

Chair Vice Chair

Greg Cizek Carey Wright Jim Geringer Rick Hanushek Suzanne Lane Alice Peisch Julia Rafal-Baer Russ Whitehurst *Sharyn Rosenberg (Staff)*

Reporting and Dissemination Committee

Chair Tonya Matthews Vice Chair Marty West Alberto Carvalho Tyler Cramer Paul Gasparini Beverly Perdue Ron Reynolds Mark White Laura LoGerfo (Staff)

Nominations Committee

Chair Jim Geringer Dana Boyd Tyler Cramer Paul Gasparini Tonya Matthews Reginald McGregor Mark Miller Alice Peisch *Tessa Regis (Staff) Lisa Stooksberry (Staff)*

Executive Committee

Chair Haley Barbour Vice Chair Alice Peisch Dana Boyd Greg Cizek Jim Geringer Tonya Matthews Mark Miller Beverly Perdue Marty West Carey Wright *Matt Stern (Staff)*

National Assessment Governing Board

Meeting of May 13-14, 2021 Virtual

OFFICIAL SUMMARY OF GOVERNING BOARD PROCEEDINGS

Complete Transcript Available

National Assessment Governing Board Members Present

Haley Barbour, Chair Alice Peisch, Vice Chair Dana Boyd Alberto Carvalho Gregory Cizek Tyler Cramer Christine Cunningham Frank Edelblut Paul Gasparini Jim Geringer Eric Hanushek Patrick Kelly Suzanne Lane **Tonya Matthews Reginald McGregor** Mark Miller **Beverly Perdue** Julia Rafal-Baer **Ron Reynolds** Nardi Routten Martin West Mark White Grover Whitehurst Carey Wright Mark Schneider (ex-officio)

National Assessment Governing Board Staff

Lesley Muldoon, Executive Director Lisa Stooksberry, Deputy Executive Director Michelle Blair Stephaan Harris Donnetta Kennedy Laura LoGerfo Munira Mwalimu Tessa Regis Sharyn Rosenberg Angela Scott Matt Stern Anthony White

National Center for Education Statistics (NCES)

Lynn Woodworth, Commissioner Peggy Carr, Associate Commissioner Tammie Adams Gina Broxterman Samantha Burg Jing Chen Brian Cramer James Deaton Alison Deigan Enis Dogan Veda Edwards Patricia Etienne Eunice Greer Shawn Kline Tina Love Daniel McGrath Nadia McLaughlin Taslima Rahman **Emmanuel Sikali** Tom Smith Holly Spurlock William Tirre Ebony Walton William Ward Grady Wilburn Angela Woodard Roberta Woods

American Institutes for Research (AIR)

Markus Broer Jack Buckley Danielle Ferguson Kim Gattis Cadelle Hemphill Martin Hooper Young Yee Kim Sami Kitmitto Gabrielle Merken Jasmine Park Yan Wang Xiaying Zheng

Council of Chief State School Officers, CCSSO

Fen Chou Scott Norton

CRP, Inc.

Shamai Carter Subin Hona Anthony Velez Edward Wofford

Department of Education

James Forester Marco Sanchez Craig Stanton

Educational Testing Service (ETS)

Siva Angappan Marc Berger Jonas Bertling Kelly Bruce Jay Campbell Peter Ciemins Gloria Dion Hugo Dos Santos Amy Dresher Kadriye Ercikan Gary Feng Janel Gill Yue Jia Cassandra Malcom Daniel McCaffrey Tenaha O'Reilly Hilary Persky Emilie Pooler Sarah Rodgers Kathleen Scalise

Adrienne Sgammato Lisa Ward Nancy Waters Karen Wixson

Hager Sharp

James Elias David Hoff Joanne Lim Debra Silimeo

The Hatcher Group

Jenny Beard Sami Ghani Robert Johnston Zoey Lichtenheld David Loewenberg Alexandra Sanfuentes Devin Simpson Nandini Singh Jenna Tomasello

Management Strategies

Brandon Dart Harrison Moore

Optimal Solutions Group

Imer Arnautovic Sonika Mishra

Pearson

Scott Becker Cindy Flockhart Abigail Keller-Dombrock Eric Moyer Paula Rios Pat Stearns Cathy White Llana Williams

Westat

Chris Averett Greg Binzer Lauren Byrne Zully Hilton Jason Nicholas Lisa Rodriguez Rick Rogers Keith Rust

<u>WestEd</u>

Matthew Gaertner Georgia Garcia Cynthia Greenleaf Mira-Lisa Katz Mark Loveland Sonya Powers Steve Schneider Megan Schneider Sarah Warner

Other Attendees/Speakers

Sarah Aguirre, Reading Development Panel Member Vickie Baker, West Virginia Department of Education Angela Battaglia, Utah State Board of Education Amanda Beaumont, U.S. Senate Committee on Health, Education, Labor and Pensions (HELP) Derek Briggs, University of Colorado Boulder Nancy Brynelson, California State University Kymyona Burk, Foundation for Excellence in Education Michael Casserly, Council of the Great City Schools Gina Cervetti, University of Michigan Allegra Chilstrom, Neal Gross Julie Coiro, University of Rhode Island Theresa Deeney, University of Rhode Island Kathilia Delp, University of the Cumberlands Danielle Dennis, University of Rhode Island Donna Dubey, New Hampshire Department of Education Jeremy Ellis, Missouri Department of Education Andrea Faulkner, North Carolina Department of Public Instruction Jennifer Fletcher, California State University Elena Forzani, Boston University Rachael Gabriel, University of Connecticut John Guthrie, University of Maryland College Park

Kathleen Hinchman, Syracuse University Andrew Ho, Harvard University Christy Hovanetz, ExcelinEd Linda Jacobson, The 74 Million Carol Jago, University of California, Los Angeles Laura Jimenez, Boston University Michael Kolen, University of Iowa Andrew Kolstad, P20 Strategies LLC Beth LaDuca, Oregon Department of Education Regina Lewis, Maine Department of Education Tamara Lewis, Maryland State Department of Education Brian Lloyd, Michigan Department of Education Rebecca Logan, Oklahoma State Department of Education Susan Loomis, Independent Consultant Scott Marion, National Center for the Improvement of Educational Assessment Michael McShane, EdChoice Raina Moulian, Alaska Department of Education & Early Development Jim Patterson, College Board P. David Pearson, University of California, Berkeley Marianne Perie, Measurement in Practice, LLC Renee Savoie, Connecticut State Department of Education Mandy Schaumburg, U.S. House Committee on Education and Labor Wanda Sims, WMS Associates Emily Slack, U.S. Senate Committee on Appropriations Michael Slattery, NAEP WTDOM Mark Stephenson, Kansas Department of Education Christy Talbot, American Educational Research Association Paola Uccelli, Harvard Graduate School of Education Zuowei Wang, University of Memphis Joyce Zurkowski, Colorado Department of Education Jeanna, ACS AI Media

Opening Remarks

Haley Barbour, Chair, called the session to order at 12:30 p.m. and welcomed attendees to the May 13, 2021, National Assessment Governing Board (Board or Governing Board) meeting.

Approval of May 2021 Agenda

Barbour requested a motion for approval of the May 2021 agenda. Jim Geringer moved to accept the revised agenda circulated to the Board members on May 12, 2021. Reginald McGregor seconded the motion which passed unanimously.

Approval of March 2021 Board Meeting Minutes

Barbour requested a motion for approval of the March 2021 meeting minutes. Alice Peisch moved to approve the minutes, and Suzanne Lane seconded the motion which passed unanimously.

Executive Director's Update

Governing Board Executive Director Lesley Muldoon provided a quarterly update.

Muldoon noted that plans are underway to convene in-person for the August 2021 Board meeting based on the current vaccination rollout in the United States. The meeting approval plans are dependent on prevailing CDC guidelines and requisite approvals from the Department of Education. Muldoon stated that the Board would be updated on developments regarding that decision.

Muldoon announced that the Board's Assistant Director of Assessment Development, Michelle Blair, would depart the Governing Board staff at the end of May to embark on a doctoral program in economics. Muldoon thanked Blair for her 14 years of service to the Board and called for the Board and staff to congratulate her.

Muldoon then reflected on challenges due to the COVID-19 pandemic as well as difficult decisions facing the Board. Virtual meetings, necessitated by COVID-19 restrictions, limited members' full experience of their service, specifically the collaboration, collective responsibility, and camaraderie which are hallmarks of the Board.

Muldoon highlighted the Board's foundational values and each member's responsibility to the National Assessment of Educational Progress (NAEP). The NAEP program has evolved over time yet continues to reflect the importance of education assessment, which it has exemplified since its inception. In the mid-1980s, then-Secretary of Education William Bennett commissioned a report about on how to make the Nation's Report Card useful to states and localities, responsive to state and local concerns but free from partisan politics. In that report, the commission recommended a new governance structure for NAEP which would accomplish three goals:

- 1. Be broadly representative of education stakeholders, in order to provide wisdom, stability, and continuity;
- 2. Mesh the assessment needs of states and localities with those of the nation; and
- 3. Be buffered from manipulation by any individual, level of government, or special interest group in the field of education.

These recommendations later served as a template for what became the National Assessment Governing Board in 1988. Muldoon said that all staff and Governing Board members bear a responsibility to uphold the founding principles of the Board. The Board, through sound policy and tradition, helps members find common ground to fulfill Board duties.

Muldoon believes the Board has reached a point where it has departed from institutional norms and traditions, which threatens the Board's functioning and its ability to fulfill its

responsibilities. Muldoon underscored that her comments were not directed at any individual or specific action. Muldoon reiterated the importance of the Board as a nonpartisan, deliberative body in which each member's perspective is valued and given equal weight. Governing Board meetings provide time and space for the deliberative process required to hear diverse perspectives, gather facts, and arrive at a consensus. She noted that nonpartisan agreement requires each Board member remain honest and open in expressing viewpoints, while committing to listen to other viewpoints. Muldoon concluded her statements by stating that the Board's primary goal with respect to the Reading Framework discussion should be to identify areas of additional consensus regarding the revised Reading Framework and allow the Board to proceed with taking action in August.

Tonya Matthews thanked Muldoon for her comments, insight, and oversight. Matthews expressed appreciation for Muldoon's recognition of the challenges brought on by the pandemic. She questioned the notion that virtual meetings due to the pandemic were the central reason that Board members were not forming collaborative, respectful relationships. Matthews described the explicit and implicit tutoring and support she received in her first years on the Board.

Matthews acknowledged the palpable tension among the Board which had gone unaddressed. She and her fellow Board members are stewards of a long history and legacy and must support and respect each other in this work. Matthews recalled her nomination to be Vice Chair by former Board member Ronnie Musgrove and remembered feeling terrified upon meeting him for the first time because he had an accent Matthews had been taught to fear. Matthews recalled that Musgrove corrected her misconception quickly through warm handshakes and Southern humor. Matthews reminded the Board that the genius of the Board derives from its bipartisan composition, non-partisan structure, and shared moments of revelry.

Barbour thanked Matthews for her remarks and highlighted changes to the meeting agenda. He noted that staff removed a session from the program to give members sufficient time to review the revised Reading Framework. Barbour announced that time had been allocated, within the day's agenda, for everyone to study the revisions which had emerged during discussions the prior week. The Chair stated the Board would reconvene at 2:30 p.m. to discuss member reactions and allow each Board member time to provide comments limited to two minutes. The floor then would open for discussion. Before breaking for the independent review time, Barbour invited members to express viewpoints.

Julia Rafal-Baer thanked Matthews for her statement, citing it as both brave and poignant. She requested additional background on the Reading Framework milestones, including the public comments received during the summer of 2020 and how it was addressed in subsequent drafts. Muldoon referenced existing summaries of the public comment and stated that she would re-send them to all Board members. Mark Miller, Vice Chair of the Assessment Development Committee (ADC), noted that Rafal-Baer could review the 2018 Board meeting minutes, available online, for additional context.

Barbour thanked Miller and announced the beginning of the independent reading period.

Independent Review Time

The May 13, 2021, session of the Board meeting adjourned at 12:57 p.m. to allow members to individually review the NAEP Reading Framework. The meeting was reconvened at 2:30 p.m.

NAEP Reading Framework Policy Discussion

Barbour opened the discussion by reminding members that the Governing Board's Executive Committee decided in April that the action on the NAEP Reading framework update would occur in August 2021 instead of during this Board meeting.

Barbour summarized activities that have taken place in discussions on the Reading Framework:

- Board members had an opportunity to read the most recent revision of the framework, which is dated April 21, 2021.
- Members participated in an informational webinar on April 30, 2021.
- Office hours were held by Governing Board staff to address member questions.

Barbour acknowledged that the Board members had the previous hour to review Board member Russ Whitehurst's edits to the draft framework, development panel feedback to those edits, and additional materials that Muldoon sent the evening of May 12, 2021, and the morning of May 13, 2021.

Before initiating the discussion, Barbour invited Dana Boyd, ADC Chair, to provide a brief summary of the May 7, 2021, ADC meeting.

Boyd noted that when she took the oath of office as a Board member, she expressed her commitment to students, parents, and teachers as the elementary school principal representative and noted that she would continue to advocate for all students. She stated that parents are entrusting the Board to make the best decisions for their children's future. As ADC Chair, her leadership style as a principal focuses on empowering others and being empowered by team members. She believes that every Board member should have an equal voice and respect one another even when they disagree. Boyd thanked the Development Panel for their work and expertise and also thanked Board staff member Michelle Blair for her role throughout the process.

Boyd then shared a chronology of recent activities:

- 1. On April 23, all Board members received the April 21 version of the draft reading framework.
- 2. On April 30, the Board hosted an informational webinar to review secure item information, which included briefings from the National Center for Education Statistics (NCES) and the Technical Advisory Committee (TAC). During that informational webinar, Board member Whitehurst asked to submit a revised version of the framework, and Chair Barbour requested that Whitehurst send the draft to him the following Monday.
- 3. On May 3, Whitehurst's revisions were sent to the Board Chair who forwarded it to the ADC as the committee of jurisdiction.

4. On May 7, ADC reviewed the revisions and forwarded them to the Development Panel for their expert opinion.

Boyd invited Christine Cunningham to summarize the discussions at the May 7, 2021, ADC meeting and Miller to conclude the briefing.

Cunningham noted that detailed minutes of the ADC meeting were shared with all Board members the previous evening. She reported that ADC members reviewed the proposed edits that Whitehurst made to the April draft and responded with respect to four areas: policy, process, intent, and content.

Addressing each area, Cunningham stated that some proposed edits were not aligned with Board policies. For example, while framework development and update processes should reflect professional standards, key reports, and research, some proposed edits deleted justifications for how the framework is conceptualized. Some of these justifications responded to external comments requesting clarification and are needed for documentation and transparency.

Second, Cunningham explained that the Committee had concerns about process. Board policy requires a process that convenes visioning and development panelists with diverse expertise. There are established processes for review, revisions, and edits by the panelists. ADC members were concerned that these processes were not being followed.

Third, Cunningham explained that the perceived intent of the edits was discussed. According to Cunningham, the proposed changes address framing and wording but do not change the assessment itself.

Fourth, Cunningham stated that the proposed content changes were discussed. She said that ADC agreed that this framework should express a commitment to equity and fairness, particularly given how equity is positioned in the Strategic Vision, in the previous Board-adopted framework, in National Academies' consensus reports, and in professional testing standards. Cunningham explained that, while the ADC objected to the deletion of related references in the edited document, it sent the edited draft to the Development Panel for their feedback. The Development Panel's feedback overlapped with several concerns about content raised by the ADC and included two areas of potential revision for the final framework.

Concluding her recap of the ADC meeting, Cunningham asked NCES Commissioner Lynn Woodworth to share NCES' position on measuring specific beliefs and attitudes in the subjectspecific contextual variables. Woodworth stated that NCES had no time to review these documents thoroughly so he declined to comment at this time. Cunningham invited Miller to provide remarks.

Citing the substantial differences between the June 2020 draft and the April 2021 draft, Miller applauded the panel for its work. He added that the ADC agrees that the April 2021 draft is a strong product that reflects the spirit and the letter of the Board's framework development policy and indicated that he looked forward to additional Board input.

Barbour thanked Boyd, Cunningham, and Miller and also expressed appreciation to Board member Whitehurst for his proposed edits to the framework. Barbour clarified that the Board had agreed to vote on the Reading Framework in August 2021. He then asked each member to respond to two questions:

- 1. Are members ready to support the framework draft presented by ADC, as proposed by the reading panel?
- 2. If so, why? If not, why not?

The Chair then called on each individual member to provide remarks with a time allotment of two minutes each.

Boyd began by commending the ADC's hard work and commitment to treating one another with respect. Boyd urged Board members to make sure the final framework is grounded in research and best practices. Boyd said that the framework development process has become divisive, including some critiques that have been interpreted as personal attacks. The presentations made by the panel and other experts provided important context, background, and details to help members and the public better understand the updated framework. She said that staff member Michelle Blair supported the ADC's review in accordance with ADC requests. Boyd was appreciative that the ADC has the opportunity to listen to each Board member. Expressing hope for a majority consensus, Boyd stated her support for the framework that the panel has proposed.

Next, Barbour called on Alberto Carvalho, who had stepped away to address an urgent matter. Later, however, Carvalho responded that he was inclined to vote for the framework. He reported that he had productive discussions with staff and all his questions were answered to his satisfaction. He had also read the letters and stakeholder input which further solidified his position.

Gregory Cizek stated he had three disappointments and one hope. He said that NAEP is the gold standard of trend monitoring, not trend setting. Cizek expressed his belief that the proposed new definition of reading in the framework changes the construct measured by NAEP and would break trend. He added that while the Governing Board has an honorable tradition of struggling to reach consensus and speaking with a singular voice, he worried that the framework development process has been degraded by factionalism. Cizek expressed concern that NAEP would lose credibility if the Board moves forward without fixing this draft.

In closing, Cizek shared an image of three children of different heights standing on boxes of different heights so that each could look over a fence at the same level. Cizek described the image, stating kids are standing on boxes so they can see over a fence labeled "Proficient," representing the desire, held by the Board, for all students to attain that level of achievement. Cizek said the boxes represent the supports for those students, which, if in place for assessments, lead to the wrong conclusion that the students are equal in their reading comprehension. Cizek explained why this is a problem. First, because the students are not equal in their reading comprehension. Second, when the test is over and they face the literacy demands of college and career, the tall kid is well-prepared but the short kid whom NAEP deemed Proficient is left to struggle. Cizek emphasized that Governing Board members were appointed to ensure the collection and reporting

of accurate data and emphasized to Board members that it was not too late to pause, reset, and refocus on what is best for students.

Tyler Cramer described his background in education advocacy and how E.D. Hirsch, Checker Finn, Russ Whitehurst, Mark Schneider, and Rick Hanushek have influenced his work. Cramer stated that what gets measured gets done and that the Reading Framework will influence how reading is taught and learned. He noted that frameworks and assessments change just like everything else. He added that he often needs help seeing such changes more clearly. He supported colleagues in wanting valid, reliable, unbiased, and in-depth assessments that illuminate inequities and inequalities – one of NAEP's most important functions.

Cramer emphasized the importance of contextual questions for illuminating inequities and making education more effective. Cramer stated that Chapter Four of this draft focuses on reporting and contextual questions, but he worried that some of the contextual questions may be challenging to implement. Cramer thanked David Pearson for leading the Development Panel and gave great weight to his recommendations, stating that Pearson is very well-respected. Cramer stated that he believed the same was true for all who served on the panels.

Cramer explained that he reached out to former Board members Carol Jago and Linda Rosen for their feedback on the framework, noting that while these experts had questions about the April 21 draft of the Reading Framework, they recommended its adoption. Cramer stated that he was leaning towards voting for the framework because he believed trend would be maintained and necessary contextual questions would be added.

Cunningham commended the range of expertise involved in the framework process and the diverse backgrounds of Board members helping to inform the process. Acknowledging that reading was not her area of study, she expressed gratitude for the framework development process and the group of highly respected reading experts who contributed to its development. She noted that many people and organizations have helped revise the framework. Although not every viewpoint could be adopted, the revisions have strengthened the framework. Concluding that NAEP is charged to look to the future while being grounded in recent research and practice, Cunningham affirmed her support for adopting the latest draft of the Reading Framework.

Frank Edelblut said that Board members share the same goal—to help all U.S. students achieve the best possible outcomes. He urged members to sustain their passion for creating bright futures for students across the nation. As a member of ADC, he was at first skeptical of the framework recommendations and met with David Pearson to understand what was being done and where the efforts were going. Like other Board members, he wanted to make sure that student outcomes would be measured in a valid and reliable way.

Understanding that the Board could keep trend on the NAEP Reading Assessment, he would support the framework. He added that he hopes the Panel will further revise the draft in response to Whitehurst's feedback and noted Whitehurst's bravery in sharing his feedback with the Board, as well as ADC's patience with the process and willingness to ensure that all voices were heard.

Paul Gasparini stated that he supported the proposed framework because he believes the framework development process had been faithfully adhered to.

Geringer stated that as a former governor, former teacher, and parent, he had encountered his share of contentious policy issues. He urged Board members to return to the work that needs to be done, in the most respectful and gracious way, and refrain from building coalitions outside of the structure of the institution. Geringer expressed his support for the ADC's work and recommended adopting the Reading Framework.

Hanushek said that the Board has at times conflated the idea of assessing reading with improving reading, stating that the latter was not in its charge. Hanushek asked: If we change views about the theory of learning, should we change the way we assess learning? Secondly, is the ADC draft evolutionary, or is it revolutionary? Hanushek said he gets imprecise and off-topic responses to the first question about how testing should relate to learning theory. Hanushek referenced the book, *How People Learn II*, and the importance of tests developed based on learning progressions rather than the science of learning.

Hanushek moved to address question two, deciphering whether Board efforts were seen as evolutionary or revolutionary. Hanushek stated that NCES adopts an evolutionary view, which he supports. Addressing the ADC draft, Hanushek asserted that the Board was trying to be revolutionary and diving into how to change reading rather than just assessing reading. Hanushek does not believe that the Board should ratify a learning theory, even if all Board members believe in that theory; the Governing Board's primary role is to assess performance.

Additionally, he asserted that the Board does not do a good job of setting the performance standards and cited the high percentage of Black eighth graders performing below the NAEP Basic achievement level as an example of an inequitable system.

Hanushek noted that he did not believe that adding a couple of contextual questions would allow the Board to get at the causality of things that affect policy. Hanushek concluded by expressing hope to return to what the Governing Board specializes in, focusing on assessment and not on learning theories and learning policies.

Patrick Kelly thanked Board members for their work on contributing to the updated framework. He expressed support for the Reading Framework produced by the Development Panel. Kelly referenced the color-coded chart on page 103 of the meeting materials, titled TABLE 1: Similarities and Differences Between the 2009–2019 and 2026 NAEP Reading Frameworks. Kelly stated that he believed the framework should be amended based on the Development Panel feedback that came as a result of Whitehurst's notes.

Kelly stated his support for the framework because the process has been responsive to feedback and concerns about sociocultural context and the importance of reading as a cognitive process. Kelly stated that he does not think creating broad definitions of reading is needed at this point as the issue was addressed in an earlier version of the framework. Kelly added that a majority of the public comments the Board received indicated that most people supported the original framework, dated [insert month year]. Kelly cited the expertise of the Development Panel and indicated that he trusted their recommendations. Kelly reiterated that he could support the April 21 version of the framework, because he supports the changes and because so many experts agreed on the high likelihood of maintaining trend. Kelly added that it was necessary for the Board to look at how to measure socioeconomic status beyond free- and reduced-price eligibility.

Lane referred to the 2014 version of Standards for Educational and Psychological Testing. She stated that fairness is a fundamental validity issue that requires attention throughout all stages of test development and use. She also shared that many test developers have incorporated socio-cognitive theories of learning into test development processes. She noted that while she was initially skeptical about the framework because of the role of UDEs in assessing reading comprehension, she has become more comfortable having read more of the background documents. Lane expressed support for the assessment framework.

Matthews stated her support for the framework and pride in the Board's work. She acknowledged that this decision was part of a bigger conversation about how the Board and the public think about assessments and the bias that may or may not be embedded in them. Matthews said she trusted her colleagues and the process that they have followed.

McGregor stated that he had the privilege of hearing from industry colleagues as well as education colleagues throughout the development process. He stated his support for the April 21 version of the framework. McGregor concluded that he supports the framework because it responds to comments from the field, is focused on students, and should allow the maintenance of trend.

Miller expressed gratitude for the opportunity to speak as an eighth-grade teacher and said he believes the April 21 NAEP Reading Framework would serve students well. He reasoned that the April 21 draft framework aligns well with the Governing Board's Strategic Vision to serve as a catalyst for action to improve student achievement and will continue to tell the story of achievement over time and in context.

Miller addressed the importance of student participation in NAEP. As a teacher, Miller said he has heard students ask: "What are we taking this test for? What happens if I don't do well? Is this going to affect my grade?" He said that the April 21 draft provides a framework for an assessment which engages students and includes motivational UDEs so as to allow all students to show what they know and can do. Miller expressed support for the April 21 version of the NAEP Reading Framework.

Beverly Perdue stated that NAEP is the North Star in assessment and that the Nation's Report Card reflects something more important than just math and reading. It reflects the ability of Americans to come together with different opinions, reach consensus, and move forward.

Perdue expressed hope that there will be greater opportunity to reflect on why 50 percent of African American children may not be achieving as they should. Having read most of the supporting documents, Perdue believed that the framework, with the revisions, was well-thought

out. Perdue expressed support for the April 21 version of the framework. Perdue concluded by thanking Whitehurst for his feedback.

Julia Rafal-Baer expressed concern that some elements of the framework and/or framework development had been conflated and recent critiques may instead reflect earlier framework drafts. In particular, letters about the framework suggested a belief that the framework is grounded in a sociocultural model, even though the revision plan states that is not the case.

Rafal-Baer worried that if the Board did not distribute the latest revision of the framework for another round of public comment, and clarify what had been removed, the Governing Board may appear partisan and risk losing support from all sides. Rafal-Baer stated that families want to know if NAEP is maintaining the trend line and developing questionnaires to better understand students' socioeconomic status. She urged additional public comment.

Ron Reynolds stated that regardless of what a person is reading, certain mental operations remain fundamental to reading. If that were untrue, then the concept of trend would have little or no meaning. Reynolds stated that the decisive factor in his evaluation of the draft Reading Framework was maintaining trend. Maintaining trend assures that NAEP will continue to assess the cognitive elements that have remained and will remain fundamental to the construct of reading. Stating that this should be NAEP's primary goal, he added that he believed trend could and will be maintained. Reynolds indicated support for the April 21 version of the framework.

Nardi Routten reflected on changes in reading expectations during her 24 years of teaching. She explained that many states, including Florida—where she spent most of her teaching career—requires fourth-grade students to read up to three texts on the same topic, synthesize information, and respond to a prompt by providing evidence and citing their source. Still, teachers, including Routten, often help students draw on personal experiences to make connections when they read text.

She added that although reading is a cognitive process, students naturally bring to bear their experiences and sociocultural resources. She asked, shouldn't the Board strive to create fair and equitable assessments? She stated that she believed a good start is providing UDEs and recognizing and understanding how students convey their experiences to reading. Routten emphasized the importance of recognizing students' differences, cultural identity, and voice. Routten concluded that if recent and updated research were dismissed in the framework, maintaining the gold star or the North Star standard would be difficult. She concluded by expressing support for the April 21 revision of the framework.

Martin West noted that the Board has reached consensus on many elements of the current framework. The Board seemed to agree that maintaining trend is a high priority. The Board also seemed to agree on a new comprehension target and disciplinary target for reading and associated text types and agreed to include all forms of UDEs currently on the operational NAEP as a strategy to reduce construct irrelevant variance. West noted that the extent of agreement is underscored by the fact that Whitehurst's proposed edits would not require changes to the assessment.

West said that his main concern about the framework was that efforts to minimize the role of background knowledge in the assessment could obscure group differences in achievement and distract educators from building background knowledge as a strategy for improving literacy.

West said Board members have been told by technical advisory committee members that none of the UDEs included in the framework update should assist a particular group of students. West did not think that this message was communicated effectively in the framework and echoed Rafal-Baer's concern about the nature of some of the misunderstandings in the feedback the Board received from stakeholders. West said that misunderstandings highlight the need for additional clarification and editing the framework. West concluded that he was optimistic about the Board's ability to get the April 21 draft approved by improving it so that it would be clearly understood by the field.

Mark White stated that he was not ready to support the framework because he wants to better understand the framework, the goal of the update, and the development process. He added that he wants to make certain that the Board was not setting up a pendulum effect for the NAEP assessment going forward where it swings one way and then the other. White concluded that as a new Board member he is trying to understand the background resources and was unable to support the framework at this time.

Whitehurst stated that he could not support the current draft framework but that he could support it if there were willingness and openness to edits. He stated that additional edits should not be onerous or upsetting but that one issue on which he seeks more clarity is language around equity and fairness. Whitehurst referenced the same chapter Lane had mentioned previously, noting that the term "fairness" has no single meaning. He stated that, while it is possible that individuals endorse fairness and testing as a desirable social goal, they can reach quite different conclusions about the fairness of a given testing program. The key element missing in the current draft, he said, is an explanation of what is considered fair or not fair.

Whitehurst then cited Mississippi's recent improvement in NAEP scores and Massachusetts's decline in scores. He asked if the implication was that these results revealed something was wrong or unfair about the assessment. Whitehurst agreed that more work would be required to avoid political pushback on the framework and urged colleagues not to hold on to something that was not absolute or necessary to move NAEP forward.

Carey Wright urged the Board to consider all input around the framework in revising the framework. She said that both the cognitive and sociocultural orientations deserve consideration and should not be in competition with each other. Wright emphasized that all Board members want an assessment system that will measure what children know and can do in a way that informs teachers, districts, and state departments of education. Wright found the revised framework and summary material helpful but asked for another review to determine if the framework was honoring both researchers who believe in cognitive and sociocultural orientations. Wright concluded by stating support for more public comment.

Peisch stated that she was comfortable taking action on the framework at the Board's next meeting. She found the current draft responsive to the concerns that were raised with the initial

draft. Peisch added that the question about maintaining trend had been answered: If Andrew Ho believed NAEP would maintain trend under the new framework, she was comfortable that it would. She said the time had come to make a decision and to move on and that once action was taken, the Board must clearly communicate what had been done and clarify any lingering confusion.

Barbour then opened the floor for questions and discussion. Geringer asked for clarification about next steps for the Board and a timeline for receiving the final framework document. Barbour clarified that the Board would not adopt the framework at this meeting; some edits would need to be made together with additional input and guidance from the Board. Muldoon stated that ADC would present an updated version of the framework during the August meeting.

Wright asked if Board members would have time to review the updated version of the framework prior to the August meeting. Barbour stated they would, although outstanding issues must be addressed and clarifications provided. He said there appeared to be a two to one sentiment in favor of adopting the framework. He did not think there was time to put out the framework for another round of public comment. Barbour clarified that the vote on the framework would take place in August. He also suggested convening a webinar in the interim and hoped the Board would come to a consensus by the time members reconvened.

Rafal-Baer asked the Board to take time to build bipartisan support for the framework and clarify confusion by seeking public input. She recommended creating a summary that could be shared publicly. She said that agreeing upon a revised framework need not require a long process, but it is necessary to remain intentional about reaching out to groups on both sides of the political spectrum and to document the input. Rafal-Baer also urged her colleagues to ensure that all state leaders have the opportunity to add their input directly, as well as district leaders who participate in the Trial Urban District Assessment (TUDA). Ultimately, she said, more input and conversations would help make the framework and the process reflect everyone's needs and concerns.

Kelly echoed Rafal-Baer's concerns about public comment and asked if it were possible to invite feedback from organizations that responded to the initial version of the framework. Doing so would give the Board the ability to measure shifts in opinion and shield the Board from some of the partisanship which emerged during the process. Kelly added that he agreed that fairness is a loaded term, referencing Cizek's presentation with the diagram of students on boxes and noting that the framework lacked clear definitions of equity and fairness.

West urged seeking public comment in a targeted way to get valuable input from the field and to encourage people to review the Board's work more closely. He believed that this could prove valuable to the reception of the eventual framework and would not be particularly onerous.

Looking toward the next Board meeting, Barbour cautioned against extending the timeline for voting on the framework because it would delay implementation of the new framework until the 2028 NAEP Reading Assessment. Barbour asked for clarity from Board members on the level of additional input they wanted and how long they wanted to wait to receive it.

Cramer stated that he agreed that a lot of the differences of opinion had been narrowed in the revised framework but was not sure that others were aware of the changes. Cramer believed the Board could reach out to the public to clarify how feedback had been addressed. He also endorsed formulating a definition of equity and fairness for NAEP.

Matthews suggested that if additional input is sought, the Board reach out to those who already commented so as not to restart the process from the beginning. Matthews concluded that one of the Board's strengths is that it is not bipartisan, but non-partisan.

Boyd stated that it was in ADC's best interest to listen to all points of view and asked for clarification from Rafal-Baer if it would be possible to hold a public webinar to provide clarity of where they were, as a Board, but not necessarily put the framework out for public comment. Boyd stated that everybody's ideas were noted and the ADC would reconvene soon to decide what to integrate into the next iteration of the framework. Boyd reiterated that a public webinar could be held to address concerns and clarifications with the framework.

Whitehurst stated that the revision process, so far, would likely not generate changes that would prove sufficient for him. He was unsure if he could vote for the framework, in part, because he said the Visioning Panel seemed to have just one view and that such an advisory group should reflect differences of opinion and that ideally those differences would overlap with concerns that were held by members of the Board. He stated that because no one on the Visioning Panel appeared to share the concerns voiced by some Board members, he encouraged further review and comments from Board members before finalizing the framework. He hoped that such a process would get the Board to a unanimous vote, which he believed would be beneficial for the Board and for NAEP.

Miller reassured the Board that the ADC would meet after this meeting to create a new version that addresses the latest comments and suggestions.

Barbour said ADC would try as quickly as possible, no later than next week, to draft a plan and timeline for revisions. Barbour reiterated the need to take action on the framework during the August quarterly Board meeting. He concluded that the Board owed it to the public and to NAEP to produce the best framework possible.

National Center for Education Statistics Update

Barbour noted that NCES Commissioner Lynn Woodworth's tenure was quickly coming to a close. Barbour expressed appreciation for Woodworth's service to the country and invited him to address the Board and provide the NCES update.

Woodworth shared his appreciation for returning to public service at NCES. He noted he had spent 6 years in the United States Marine Corps, 11 years as a public-school teacher, 5 years as a researcher, and a number of years as a student at various levels, all in public institutions.

Woodworth thanked Peggy Carr and commended Carr on her extensive experience in all areas of NAEP's work. He recognized Carr as an outstanding individual who has made invaluable contributions to NAEP. He thanked the rest of the NCES staff and expressed appreciation to Enis Dogan, Daniel McGrath, Patrick Keaton, Eunice Greer, Gina Broxterman, Grady Wilburn, and Ebony Walton.

Woodworth also thanked the Governing Board staff, noting that the work had been tough and that many long and frank conversations took place, but they were all valued for working in the best interest of the country and its students. He noted the Governing Board's commitment to NAEP. Woodworth referenced the Reading Framework discussions, stating that the conversation is emblematic of what should be focused on: the country and its students. Woodworth expressed appreciation for the hard work of NAEP contractors, particularly during the COVID-19 pandemic.

Woodworth acknowledged the unprecedented nature of the pandemic, mentioning that never, in the history of the United States, had nearly every public school in the country been shut down. The adjustments made, by the NCES staff, the Governing Board staff, and the contractors challenged everyone to work harder. Woodworth mentioned that the NAEP program will return to the field in 2022 to gain an accurate measure of achievement and learn where students are academically, based on the impact of COVID, which will prove vitally important. That assessment, by Woodworth's standards, will be the most important NAEP assessment done in the history of the United States due to the unprecedented events that have unfolded over the past year.

Woodworth stated that the current NAEP began in 1990 and it was time for certain updates to NAEP. The methodology to develop the NAEP assessments in 1980 brilliantly solved the issue about how to assess the existing content and materials without spending an excess amount of time doing it. The people who designed NAEP determined that a matrix design could be used, which would cover a large scope, meaning that no student had to spend five days taking the assessment. This development met the needs that would allow estimates for large sub-populations at the national and state levels. However, Woodworth stated that the assessment did not meet the needs of today.

Woodworth added that there are a number of advances in computing power which have led to different uses of NAEP data. Woodworth cited that one of today's more common uses is regression analysis to examine correlations across the data set and across student characteristics. Woodworth claimed that the current NAEP does not work well for this use because those data analyses require reliable and replicable scores for each student. Woodworth encouraged the Board to consider moving NAEP to an item-level computer adaptive assessment. These assessments would provide accurate estimates at all levels of student achievement—at the very top, the middle, and at the bottom. The assessments would produce comparable direct scores for each student that could be used to run these analyses. Not only could group outcomes be evaluated, but also individual-level data.

Woodworth recommended that NAEP move toward online and device-agnostic administrations, stating that the expenses of administering NAEP in person would not be sustainable moving

forward. Woodworth clarified that collecting data for NAEP assessments was the most expensive part of NAEP. The COVID-19 crisis highlighted flaws in the administration model, because people needed to enter schools to set up, check, and operate equipment. Ultimately, the pandemic prevented the assessment from taking place. An online and device-agnostic NAEP could have occurred, despite COVID-19.

Woodworth also recommended that NAEP move toward two-subject design. No other major assessment still uses a one-student, one-subject design. Woodworth stated that there were advances to run correlations between subjects if a student took both reading and math. Currently, there is a set of students assessed for reading and a separate set of students assessed for math, so those types of analyses and comparisons are impossible.

Woodworth challenged the Governing Board to persuade NCES to certify the achievement levels as being reasonable and informative to the public. He claimed that the previous six commissioners had agreed with him that the NAEP achievement levels do not work. He pointed out that half the students in major sub-populations score below NAEP Basic based on the current standards. The last time the 12th grade math assessment was administered, 75 percent of 12th grade math students scored below *NAEP Proficient* and 40 percent scored below *NAEP Basic*. Woodworth declared that these results lead to one of two conclusions: either the United States' education system is a disaster, failing huge swaths of kids, or the achievement levels are in the wrong place.

Woodworth acknowledged the Governing Board's interest in scenario-based tasks (SBT) for NAEP but noted their immense expense. The Board wants more frequent assessments and more state level assessments, but the current budget cannot support those goals. SBTs cost twice more than a similar block of multiple-choice items, even though they do not provide twice the information power. Woodworth requested that the Board evaluate the SBTs from a scientific standpoint and determine if they merit the investment.

Woodworth said that NAEP is struggling, along with every other type of data collection, to recruit students and schools to participate. He recommended that Congress decide what subjects and grades are most important to assess and provide funds necessary to complete those assessments. He added that Congress should mandate all states participate in those assessments.

Woodworth stressed the importance of NAEP, reasoning that if it was important enough to spend millions of dollars to administer and for kids to miss time from school, then the assessments are important enough for Congress to say that states must participate. Woodworth added that NAEP is one of the Department of Education's most important activities and plays an important role in measuring education performance and educational equity across the nation.

Woodworth concluded that, without NAEP, there would be no standard to measure education system outcomes, no way to compare performance from one state to the next, and no way to understand the gaps that keep children from achieving their potential. Woodward reiterated that NAEP remains a critical tool in the fight for the American education system and he invited everyone to move toward a better NAEP.

<u>Resolution in Honor of Michael Casserly, Executive Director, Council of the Great City</u> <u>Schools</u>

Barbour announced that Michael Casserly, Executive Director of the Council of Great City Schools was retiring after 30 years of service. He thanked Casserly for his collaboration, invaluable insights, and his advocacy and contribution to NAEP.

Barbour called attention to a resolution, drafted in Casserly's honor to recognize Casserly for his support, contributions, accomplishments, and partnership. Barbour noted that Casserly is an ardent advocate for NAEP and his work has bolstered the strength and power of NAEP in harnessing NAEP data. He has focused on improving achievement in large cities.

The Chair expressed gratitude for Casserly's insightful leadership and his founding of the Trial Urban District Assessment (TUDA) program – a subset of NAEP which represents samples of students in select urban districts across the United States.

Barbour then read the resolution to honor Casserly (see Appendix X for the full text of the resolution).

Barbour requested a motion for approval of the resolution of Michael Casserly. Cramer moved to approve the resolution; the motion was seconded by Martin West and passed unanimously. Barbour then invited Casserly to share remarks.

Casserly thanked the Chair for his words of affirmation and his leadership. Casserly also thanked Mark Schneider of the Institute of Education Sciences; Lynn Woodworth and Peggy Carr of NCES; Executive Director of the Governing Board, Lesley Muldoon; her predecessors; and the Governing Board staff for their hard work and dedication.

Casserly noted that the autumn of 2000 was a fraught period in the history of NAEP. Casserly approached the Board and proposed the development of what became the Trial Urban District Assessment (TUDA) for three reasons:

- 1. Districts wanted a way to compare themselves to other big city school districts facing similar challenges;
- 2. Districts wanted to see if their efforts to improve were actually reaping benefits, if they were improving as they thought they would; and,
- 3. Districts wanted to test whether or not the reforms they were pursuing actually worked and which did not.

Casserly added that NAEP has helped TUDAs improve instruction for students in urban school districts and has served as an invaluable tool. He expressed gratitude for the Board's forward thinking in agreeing to the TUDA proposal over 20 years ago. While more must be done, the strides taken so far have immensely improve urban education throughout the decades. Casserly concluded by thanking the Board.

Concluding Remarks

Barbour announced that the next day's sessions (Friday, May 14) were closed to the public as members would be receiving embargoed NAEP Science results and engage in budget discussions. Barbour thanked the Board, its staff, NCES, and all who helped facilitate the meeting.

The meeting adjourned for the day at 5:28 p.m.

Results from the 2019 NAEP Science Assessment (CLOSED)

Under the provisions of exemption 9(B) of §552b(c) of Title 5 U.S.C., the National Assessment Governing Board (Governing Board) met in closed session on Friday, May 14, 2021, from 12:39 p.m. to 1:47 p.m. to receive a briefing from Grady Wilburn of the National Center for Education Statistics (NCES) on results from the 2019 National Assessment of Educational Progress (NAEP) in Science.

Chair Barbour called the meeting to order. He explained that the session was closed to the public because results of the science assessment have not been released to the public. He noted that the initial release event will occur on May 25, 2021. Board member Christine Cunningham and former Board member Cary Sneider will participate in the release along with NCES staff.

Wilburn provided a briefing on the NAEP 2019 science results at grades 4, 8, and 12. He noted that the science assessment was administered between January and March 2019 to a national sample of 4th, 8th and 12th graders. This administration marked the first time the science assessment was given via tablet, but the assessment was administered through both paper-and-pencil and digital platforms to preserve trend lines of student performance over time. Wilburn reported that the NAEP Science assessment includes three content areas—physical science, life science, and earth and space sciences. The assessment measures four science practices—(1) identifying science principles; (2) using science principles; (3) using scientific inquiry; and (4) using technological design.

Wilburn highlighted sample scenario-based tasks in the science assessment at each grade. He then provided an overview of student performance at grades 4, 8, and 12 and compared the science results with the NAEP 2015 and 2009 science results by:

- Average scale score
- Percentile score
- Content area scores
- NAEP achievement levels

He then highlighted the results by grade and average scale scores. The results were compared with the 2015 and 2009 assessment results. Student performance was described by race/ethnicity, participation in the National School Lunch Program, and gender. Wilburn also shared insights from the contextual questionnaires, including how often students engaged in scientific inquiry-related classroom activities and their enrollment in biology, chemistry, and physics classes. Board members asked questions during the presentation, which Wilburn addressed.

Recess

The May 14, 2021, Governing Board meeting recessed at 1:47 p.m. and reconvened at 2:05 p.m.

NAEP Budget and Assessment Schedule (CLOSED)

Under the provisions of exemption 9(B) of §552b(c) of Title 5 U.S.C., the Governing Board met in closed session on Friday, May 14, 2021, from 2:05 p.m. to 3:46 p.m. to receive a briefing from Peggy Carr, Associate Commissioner, NCES, on the NAEP budget and Assessment Schedule. Chair Barbour announced that the session was closed to the public and that only Governing Board members could ask questions. Congressional staff attended the session for information gathering purposes only.

Carr provided a briefing on the NAEP budget and its impact on the NAEP Assessment Schedule. The briefing covered three areas:

- COVID implications in 2022
- Current budget cash flow for FY 2020 through FY 2024
- Marginal costs of assessments

Carr concluded the briefing by highlighting the President's budget proposal timeline for Fiscal Year 2023:

- July-August 2021: Offices submit requests and justifications to Department Budget Service
- September 2021: Department submits to OMB
- November/December 2021: OMB pass back/Department appeals
- February 2022: President submits FY23 budget proposal to Congress

Board members asked questions during and after the presentation, which Carr addressed.

Next Generation NAEP (CLOSED)

Under the provisions of exemption 9(B) of §552b(c) of Title 5 U.S.C., the Governing Board met in closed session on Friday, May 14, 2021, from 3:47 p.m. to 4:35 p.m. to receive a briefing from Peggy Carr on the Next Generation NAEP.

The briefing covered three areas:

- Costs relative to changes in NAEP's scope of work over time
- Reviewing major cost drivers
- Investment needed to modernize NAEP

Carr addressed questions from Board members throughout her presentation.

The meeting adjourned at 4:35 p.m.

I certify the accuracy of the minutes.

Chair Chair

<u>July 22, 2021</u> Date



Resolution in Honor of Michael Casserly Executive Director of the Council of the Great City Schools

Approved May 13, 2021

Whereas, Michael Casserly has served as Executive Director of the Council of the Great City Schools for 29 years and is stepping down from his leadership position and assuming a role of consultant and advisor;

Whereas, through his insightful leadership of the Council, Michael Casserly inspired and initiated the Trial Urban District Assessment Program (TUDA) of the National Assessment of Educational Progress (NAEP), which assesses representative samples of students in urban districts across the United States through which invaluable data are provided for these districts to understand and to improve the educational experiences and outcomes of their cities' students;

Whereas, Michael Casserly strengthened and sustained the TUDA program for more than two decades, expanding the program from six districts in its first year to 27 districts that currently participate by eloquently explaining the unique value and immense power of TUDA data to urban school district policy and function;

Whereas, the TUDA data facilitate urban school districts' efforts to improve student performance and close achievement gaps by allowing districts to conduct comparative analyses with districts similar in demographic profile, to learn lessons from peers' experiences and successes, and to discern and implement effective education practices, with large cities making significant score gains and their performance on NAEP Reading and Mathematics improving faster than the nation;

Whereas, Michael Casserly's leadership of the TUDA Task Force with the Governing Board affords vitally important feedback to the Governing Board to inform and improve policy, research, and communications related to NAEP and to the TUDA program, leading to significant improvements;

Whereas, Michael Casserly's ardent advocacy for NAEP neither began nor concluded with the TUDA program in that he consistently and persuasively championed for support for the Nation's Report Card and worked tirelessly to advance the NAEP program through his savvy, sage, expert advice, his keen insights on strategic implementation of assessment programs, and his thoughtful, collegial collaboration with both the National Center for Education Statistics and the Governing Board members and staff;

Therefore, be it resolved that the National Assessment Governing Board expresses its profound appreciation and gratitude for Michael Casserly's unwavering support and monumental contributions to NAEP and student achievement in our nation;

Be it further resolved that a copy of this resolution be entered permanently into the minutes of the National Assessment Governing Board meeting of May 2021.

Signed on this Thirteenth Day of May, Two-Thousand and Twenty-One

aly anton

Haley Barbour, Chair National Assessment Governing Board

National Assessment Governing Board Executive Committee Meeting Report of May 5, 2021

CLOSED SESSION

Executive Committee Members: Haley Barbour (Chair), Alice Peisch (Vice Chair), Dana Boyd, Gregory Cizek, Jim Geringer, Tonya Matthews, Mark Miller, Bev Perdue, Martin West, Carey Wright.

Executive Committee Members Absent: None.

National Assessment Governing Board Members: Ron Reynolds, Mark Schneider (exofficio).

National Assessment Governing Board Staff: Michelle Blair, Stephaan Harris, Donnetta Kennedy, Laura LoGerfo, Lesley Muldoon, Munira Mwalimu, Tessa Regis, Sharyn Rosenberg, Angela Scott, Matthew Stern, Lisa Stooksberry, Anthony White.

National Center for Education Statistics Staff: Peggy Carr, Pat Etienne, Dan McGrath, Holly Spurlock, James Lynn Woodworth, Brian Cramer, Ebony Walton, Enis Dogan, Eunice Greer, Jing Chen, Shawn Kline, William Tirre.

The Executive Committee met in closed session from 3:00 p.m. to 5:00 p.m. to discuss the NAEP budget and assessment schedule, in addition to other Governing Board priorities.

The closed session was called to order by Chair Haley Barbour at 3:00 p.m.

These discussions were conducted in closed session because the disclosure of cost data would significantly impede implementation of contract awards. Therefore, this discussion is protected by exemption 9(B) of section 552b(C) of Title 5 U.S.C.

Barbour reminded members of the confidential nature of the discussions. Barbour informed other Executive Committee Members that he would have to leave the meeting early and that Vice Chair Peisch would assume the responsibilities of the Chair, in his absence.

Barbour provided a brief update on the NAEP Reading framework timeline and process and the Board's recordkeeping policies based on guidance from the U.S. Department of Education. Barbour reminded Executive Committee Members that more information on the recordkeeping requirements would be shared at a later date.

Barbour then introduced Peggy Carr, Associate Commissioner, National Center for Education Statistics (NCES). Carr led a presentation on the NAEP Budget. Carr provided information about projected costs for the program, including the impact of COVID and school closures, updating the reading framework, and forthcoming research and development. Carr described projected budget implications for the NAEP Assessment Schedule.

Lesley Muldoon, Executive Director, then facilitated a discussion on potential implications for the NAEP Assessment Schedule in the short- and long-term. Muldoon facilitated members' discussion on a transition to the next generation of digitally based assessments and how NAEP modernization could lead to long-term cost savings.

At 4:52 p.m. Vice Chair Peisch adjourned the meeting.

I certify the accuracy of these minutes.

Haley Barbour, Chair

July 1, 2021 Date

National Assessment Governing Board

Assessment Development Committee

Report of May 7, 2021

Assessment Development Committee (ADC) Members: Dana Boyd (Chair), Mark Miller (Vice Chair), Christine Cunningham, Patrick Kelly, Reginald McGregor, Nardi Routten.

Other Governing Board Members: Tyler Cramer, Ron Reynolds, Mark Schneider (ex officio). **Governing Board Staff:** Lesley Muldoon, Lisa Stooksberry, Michelle Blair, Stephaan Harris, Laura LoGerfo, Munira Mwalimu, Sharyn Rosenberg, Angela Scott, Matthew Stern,.

Other Attendees: Doretha Allen, Imer Arnautovic, Sadaf Asrar, Scott Becker, Rebecca Bennett, Jonas Bertling, Derek Briggs, Markus Broer, Nancy Brynelson, Jack Buckley, Lauren Byrne, Gina Broxterman, Jinghong Cai, Peggy Carr (Associate Commissioner, NCES), Shamai Carter, Jay Campbell, Gina Cervetti, Jing Chen, Julie Coiro, Brian Cramer, Kathilia Delp, Gloria Dion, Patricia Donahue, Jeremy Ellis, Kadriye Ercikan, Patricia Etienne, Gary Feng, Chester Finn, Elena Forzani, Matt Gaertner, Georgia Garcia, Kim Gattis, Eunice Greer, Laura Goadrich, Cynthia Greenleaf, Kristin Hamilton, Kathleen Hinchman, David Hoff, Martin Hooper, Maria Hyler, Linda Jacobson, Carol Jago, Robert Johnston, Beth LaDuca, Carol Lee, Joanne Lim, Mira-Lisa Katz, Young Yee Kim, Mark Loveland, Scott Marion, Daniel McGrath, Nadia McLaughlin, Gabrielle Merken, Jim Patterson, P. David Pearson, Hilary Persky, Susan Pimentel, Emilie Pooler, Stanley Rabinowitz, Sarah Rodgers, Rick Rogers, Alicia Ross, Megan Schneider, Steve Schneider, Nandini Singh, Paola Uccelli, Sheila Valencia, Ebony Walton, Sarah Warner, Karen Wixson, Edward Wofford.

NAEP Reading Framework Update

Vice Chair Mark Miller welcomed Committee members and other Board members in attendance. Mark reviewed three goals for today's Assessment Development Committee (ADC) session on the NAEP Reading Framework: (1) to review guidance from the Technical Advisory Committee (TAC) for the NAEP Reading Framework Update, as issued to the Framework Development Panel; (2) to review suggested edits recently submitted by Board member Russ Whitehurst; and (3) to review suggested edits submitted by the Framework Development Panel in April 2021.

Miller then introduced Matt Gaertner, WestEd Measurement Lead, to provide a report on the most recent meeting of the TAC for the NAEP Reading Framework Update. Gaertner noted that the TAC has provided ongoing advice for the Framework Development Panel as they have been drafting NAEP Reading Framework Update recommendations. He then introduced two members of the eight-person TAC who were in attendance: (1) Derek Briggs, Professor of the Research
and Evaluation Methodology at the University of Colorado, Boulder and director of the Center for Assessment Design Research and Evaluation (CADRE); and (2) Scott Marion, Executive Director of The National Center for the Improvement of Educational Assessment.

At its recent meeting, the TAC discussion focused primarily on issues related to topic knowledge and Universal Design Elements (UDEs), with a special emphasis on knowledge-based UDEs recommended by the Framework Development Panel. Gaertner began by reviewing the TAC's discussion of UDEs. The TAC was asked to provide their advice about the types of evidentiary standards that should be used to understand whether UDEs are functioning in an assessment as intended. Gaertner noted that for assessment accommodations, appropriate supporting evidence would be data showing that the accommodation in question was assisting only the students who needed it. In contrast, Gaertner indicated that UDEs are intended for all students; that is, none of the UDEs proposed in the NAEP Reading Framework update are intended to assist a particular group of students, e.g., by gender, socioeconomic status, or race/ethnicity. Because of this, he noted that the appropriate supporting evidence would be data showing that a given UDE assists all students; evidence that a UDE assists a particular group of students could, therefore, be a cause for concern.

In the TAC discussion about topic knowledge, Gaertner summarized that the assessment construct in the proposed NAEP Reading Framework Update specifically excludes textindependent domain knowledge, as does the current NAEP Reading Framework. The framework update specifies that although topic knowledge is important, it is not part of the intended assessment construct. Gaertner noted that the two knowledge-based UDEs that are proposed in the NAEP Reading Framework update are: (1) short introductions, of one or two sentences, to potentially unfamiliar topics; and (2) providing definitions for words that are likely to be unfamiliar and are not part of the construct being assessed. Gaertner stated that the knowledge-based UDEs in the framework update support the intended assessment construct. Further, they have precedent and are deemed non-controversial in assessment practice, i.e., they appear in various state-wide summative assessments across the U.S. For example, in a recent CCSSO survey, most states indicated that they use similar UDEs on their assessments.

Gaertner asked Briggs and Marion to offer any additional comments about the TAC advice to the Framework Development Panel. Marion clarified that the introduction of topic knowledge could be a threat to the validity of reading comprehension assessment. Briggs asserted that while the term 'topic knowledge' has been used extensively in discussions about the intended construct for the NAEP Reading Assessment, it would be more appropriate to think of the issue as one of 'topical familiarity.' The knowledge-based UDEs are more like "primes" to allow students to be cued in to what to look for, getting students ready to engage with the text; they are not intended to be an instructional intervention.

Patrick Kelly asked whether UDEs interplay with NAEP's time constraint in a way that improves the ability of the assessment to measure student achievement. Marion noted that the knowledgebased UDEs are intended to activate the students' prior knowledge. The amount of extra reading relating to the knowledge-based UDEs is miniscule. Further, he stated that the NAEP exam is not intended to be speeded. Briggs noted that the provision of UDEs increases student engagement and increases the student's ability to take agency over their assessment experience. Hence, there are likely to be some efficiency gains as students progress through the assessment and potentially spend less time on each assessment task. Marion agreed and noted that the TAC does not believe these UDEs will add to assessment time.

Chair Dana Boyd reported that a few days ago Board member Russ Whitehurst shared a suggested rewrite of the NAEP Reading Framework update. Per the Governing Board Framework Development Policy and NAEP law, active participation of a wide array of stakeholders is required for each framework process. Accordingly, Boyd observed that the Board typically provides feedback to Framework panels in a more collaborative manner and at an earlier stage in the framework development process. She asserted, however, that the Board wants the best framework possible. Hence, it is important to carefully consider the critiques and suggestions. She asked the Committee to share thoughts about whether and how these edits should be shared with the Development Panel as guidance from the Governing Board.

Christine Cunningham said that she thoroughly reviewed the suggested edits and had four areas of concern relating to: (1) policy, (2) process, (3) intent, and (4) content. In the issues of policy, the edits are misaligned with many of the policies that are supposed to be informing frameworks and framework development. Cunningham cited the fourth principle of the Governing Board Framework Development Policy which says, "The NAEP framework development and update processes shall be informed by a broad, balanced, and inclusive set of factors. The framework shall reflect current curricula and instruction, research regarding cognitive development and instruction, and the nation's future needs and desirable levels of achievement. This delicate balance between 'what is' and 'what should be' is at the core of the NAEP framework development process." This Board policy also instructs framework panels to consider a wide variety of resources "...including but not limited to curriculum guides and assessments developed by states and local districts, widely accepted professional standards, scientific research, other types of research studies in the literature, key reports having significant national and international interest, international standards and assessments, other assessment instruments in the content area, and prior NAEP frameworks, if available." To summarize the policy, Cunningham stated that the Board is supposed to prioritize what is currently occurring in research, professional standards, and key reports, and this all highlights the forward-thinking "what should be" nature of the Board's work.

Cunningham then noted that the Board's policy also mandates that each framework must set forth the content to be measured, the format for the assessment, and the levels of achievement. Therefore, removing achievement level descriptions from any framework would be in violation of Board policy.

Regarding the issue of process, Cunningham noted that Board policy requires the participation of many stakeholders, which is why the Visioning and Development panels convened for a framework update must be comprised of nationally renowned experts. Cunningham asserted that we cannot and should not invite experts to work over many months in an iterative drafting effort, and then allow one person to edit their carefully written document. Cunningham concluded her remarks regarding process by noting that allowing any Board member to rewrite a framework breaches the norms of transparency and open deliberation that are central to NAEP framework processes.

Cunningham then described her concerns about the intent of the edits proposed by Whitehurst. Cunningham noted that in the preamble, where Whitehurst summarized his edits, he stated that his proposed edits do not require any changes in the assessment plans recommended in the Framework Development Panel's April 2021 draft. Cunningham observed that this implies that the edits put forward do not relate to the concerns articulated in the Board's policy deliberations over the last nine months – namely, maintaining trend, Universal Design Elements, or what is best for students. Cunningham indicated that many of the words struck in the edited document are widely accepted and espoused by current research. Cunningham reported that she could not agree with these deletions, and she suspected that there may be other Committee members, members of the Framework Development Panel, classroom teachers, reading researchers, and members of the public across the country that might also find many of these edits objectionable.

Cunningham conducted general classifications of the edits suggested by Whitehurst. In her final area of concern, Cunningham summarized these content issues by asking a series of questions to understand the sense of the Committee:

- 1. Do we as a Committee have a commitment to equity?
- 2. Do we as a Committee have a commitment to fairness?
- 3. Do we aspire to create assessments that are equitable, non-biased, valid, fair, rigorous, precise, and accurate? (She clarified that all of these words with the exception of one appearance of the word "valid" were removed in the proposed revisions of the framework.)
- 4. Should we as a Committee and as a Board minimize test bias?
- 5. Should the framework reflect advances in current research about learning in reading?
- 6. Should the framework draw on work and recommendations of non-partisan National Academies committees?
- 7. Do we as a Committee believe that social and cultural experiences shape learning?

- 8. Do contexts, readers, texts, and activities all affect students' readings?
- 9. Do we see as consistent with our mission designing assessments that call attention to malleable factors that are most likely to lead to improved policies and practices that can shift students' outcomes? (She clarified that this last question largely relates to edits proposed for Chapter 4, which deals with reporting.)

All Committee members in attendance affirmatively answered each of these nine questions by raising their hands. Cunningham then explained that each of her questions related to sections of the framework that were either deleted or rephrased. She inferred that the Committee's affirmative answers to her questions meant that many of these edits are problematic.

Cunningham noted that the latest Strategic Vision, adopted by the Governing Board in September 2020, observes that equity is a prevalent goal of educational practitioners. Language about equity also appears in the NAEP Technology and Engineering Literacy Assessment – as a content area for the assessment and also as an ideal for the assessment itself. Given the commitment to equity listed in the current Governing Board Strategic Vision, current NAEP frameworks, and the Committee's aforementioned consensus, Cunningham concluded that it is appropriate for the word "equity" to appear in a NAEP framework. The April 2021 draft of the framework from the Framework Development Panel defines equity, notes it as an important outcome, and references several non-partisan national reports that have extensively addressed equity issues.

Miller thanked Cunningham for her analysis. He commented that her effort helps to ensure that the Committee carefully considers the decisions ahead. Nardi Routten expressed appreciation for Cunningham's remarks. Routten reported that she also noted the repeated deletions of certain words, and she found these deletions offensive as an educator.

Kelly appreciated the work that Whitehurst devoted to his revisions within a short period of time. Kelly commented that Whitehurst's passion is remarkable and commendable. Kelly said he was concerned that the Board is not effectively communicating about the framework because some Board members are still considering the public comment draft that was released in June 2020. Since then, there have been several revisions to account for and incorporate public comment and Board feedback. The April 2021 draft before the Committee today is dramatically different from what was put forward in June 2020; and Kelly claimed that this difference is a testimony to the quality and strength of the Board's framework development process. As a Board member, he agrees that adhering to research evidence is an important principle for assessment. The framework's citation of Universal Design of Assessment (UDA) principles is well-sourced and reflects what good assessment design ought to encompass. He observed that Whitehurst did not edit these principles, which likely implies that he also views UDA principles as important for best practice. Looking at the UDA principle for inclusive assessment populations, Kelly stated that equity is essential to this principle, and designing assessments to ensure equitable access is an appropriate goal for a framework to put forward. Kelly also noted that he wanted to hear what the Framework Development Panel thought about Whitehurst's edits, because they are the experts empaneled to provide guidance on the assessment construct and the assessment design.

Cunningham noted that the theoretical basis and research evidence for selected contextual variables are required in the Governing Board's policy on NAEP contextual questionnaires. Cunningham said that another content aspect of the suggested edits was the deletion of theoretical rationales for the contextual variables recommended for subject-specific questionnaires under the NAEP Reading Framework update. However, many of these deletions counter Board policy. Cunningham clarified that theoretical rationale is not a philosophical or instructional tool; all assessments are grounded in a theory of how people learn - stated or unstated. Yet, theoretical rationales in the NAEP Reading Framework update have been largely deleted in the proposed edits. By removing these rationales, the remaining research that is cited is dated in nature and does not reflect the widely accepted view of research in education, what most educators know about teaching, or the consensus identified in different National Academies reports. The role of cultures, societies, social interactions, and contexts in learning and in learning reading have been well-established. To relabel these as environmental factors is not consistent with current research. By deleting references to context, for example, the framework becomes incomplete, relative to current research. Cunningham reiterated that the Board is mandated to ensure that frameworks reflect current research. Routten agreed that it is important to reflect recent and updated research in assessment. Boyd said that as a Committee, the ADC has several practitioners. She agreed that from a practitioner perspective, it is inappropriate to exclude consideration of recent research.

Kelly observed that the public comment draft of the framework may have referred to sociocultural theory in ways that sparked several critiques, and these critiques were then used to refine the framework update into the improved draft that the Panel put forward in April 2021. Kelly also observed that some of the ongoing critiques assert that the framework is attempting to delegitimize the role of content and background knowledge in reading, but this seems inaccurate. The framework even cites Daniel Willingham, whom we know is an expert who believes that background knowledge matters in reading. Kelly noted that in the preamble where Whitehurst summarized his edits, Whitehurst argued that the sociocultural perspective is "elevated to conceptual preeminence" in the April 2021 draft. Kelly suggested, however, that this was not accurate, and it could be that Whitehurst was referring to the public comment draft (released in June 2020). In the April 2021 draft, the NAEP Definition of Reading Comprehension emphasizes reading as a complex cognitive process. In the original feedback from the public, only one person accused the public comment draft of activism, and in response, the references to sociocultural theory were substantially scaled back. It is, therefore, difficult to understand how

sociocultural theory could be framed in "politically charged and divisive" terms (as noted in the Whitehurst preamble summary) in this latest draft from the Framework Development Panel (as submitted in April 2021).

Cunningham added that many of the suggested edits removed rationales or details that were requested in the public comment period. These rationales are also important for the various validity research activities undertaken by NCES. Kelly noted that it is possible that Whitehurst removed these rationales to improve the accessibility of the framework as a document, e.g., by deleting excessive research citations. Still, Kelly noted that it is more important for the framework to be well-documented and supported than it is for the framework to be highly readable. Miller noted that abridged frameworks are developed after a framework is adopted, and that these shorter documents tend to be more user-friendly and accessible to the general public.

Regarding the contextual variables, Cunningham asked NCES to confirm whether the Panel's framework recommendations are consistent with past NAEP practice. Dan McGrath reported that NCES only measures things that are closely related to achievement, which includes attitudes, confidence, and perseverance related to reading, for example. Kelly noted that the NAEP survey question that Whitehurst cited related to self-efficacy does appear to be asking students to share a personal belief. Kelly said that having this type of NAEP survey question should be carefully reviewed and revisited by the Governing Board, but the measurement of self-efficacy occurs in other NAEP subject areas as well – so this is not specific to the NAEP Reading Framework Update.

Reginald McGregor applauded the dedication that Whitehurst showed in helping the Board make the right decisions on the NAEP Reading Framework. He also praised the extensive efforts that have been made to collect feedback from various stakeholders in the development of the framework update. McGregor noted the importance of having documents that reflect inclusiveness, whether the document is an employee handbook or an assessment framework. He also said that current research must be cited in whatever NAEP does. He remarked that he also hoped that negative connotations of the word "equity" could be set aside.

McGregor then reminded the Committee that Governing Board members have largely expressed two major areas of concern with the framework update under discussion: Universal Design Elements and topic knowledge. He encouraged the ADC to stay focused on these two issues and praised Cunningham for her thoughtful analysis.

Cunningham said that with the proposed removal of current research citations, it is not clear what would be new in the framework update; if these edits are accepted, it is not clear that a framework update is even needed. She added that misalignment with current research would not be aligned with the Strategic Vision, which calls for innovation.

In closing the discussion of these suggested edits, the Committee agreed to forward the edits to the Framework Development Panel to determine if any should be considered for inclusion in the framework, based on their expertise. Miller also noted that it is important for the Panel to also receive a summary of the current ADC discussion, as context for this request.

Miller then opened discussion of the April 2021 draft of the framework, which incorporates feedback received at the March 2021 Quarterly Board meeting. For example, multimedia knowledge-based Universal Design Elements are removed, and there are other clarifications – such as removing redundancy and improving graphics. He asked whether the Governing Board Framework Development Policy was implemented faithfully, particularly the responses to public comment and Board feedback. Kelly stated that this draft is responsive to the Board's charge, public comment, and Board feedback over multiple iterations. It aligns NAEP with what is appropriate in the field based on current research and understandings, without being revolutionary and unnecessarily introducing instability to what the NAEP Reading Assessment measures. Kelly noted that page 53 has an example item with a UDE that occurs as a video. He asked if this was an error given that knowledge-based UDEs that involve video are supposed to be relegated to a special study. David Pearson responded that this video component might have been a part of the original text, but he needed to confer with the lead authors of Chapter 3 to confirm. Framework Development Panel member Julie Coiro was in attendance. As a lead author for Chapter 3, she clarified that the video was a motivational UDE, to get students intrigued with the scenario. Coiro shared that her understanding was that knowledge-based UDEs could not be in a video format, but that motivational UDEs allowed for the possibility of video.

Boyd thanked the Panel for their ongoing and thorough efforts.

NAEP Mathematics and Reading Framework Processes

Miller stated that it is important for the Committee to review how framework processes have been implemented, as the ADC deliberates on potential process improvements for future framework projects. To provide a report on recent implementations, Miller introduced Mark Loveland, the WestEd Project Co-Director for both the Mathematics and Reading Framework updates.

Loveland reviewed the Governing Board policy principles that drive how WestEd has been implementing framework processes. To launch the work, each Panel receives various resources including copies of the current framework, specifications, NAEP survey questionnaires, and access to items from the current assessment. So, both the NAEP Mathematics Framework Development Panel and the NAEP Reading Framework Development Panel anchored their work in the current frameworks, rather than writing completely new framework documents. He noted that the COVID-19 pandemic prompted a transition to virtual meetings and timeline adjustments.

Loveland summarized the ongoing involvement of the group of psychometricians that comprise the Technical Advisory Committee (TAC) for each framework panel – attending Panel meetings, holding their own TAC meetings, reviewing draft documents, and responding to Panel questions.

Then, Loveland discussed the demographic panelist selection criteria, such as teaching experience, geographic region, locale (i.e., urban, suburban, rural), gender, and grade band (i.e., elementary, middle, high school). He described the areas of expertise that were needed for each Panel; these areas were identified in the Governing Board Framework Review, which was conducted before each framework update project was launched. Finally, Loveland summarized the types of stakeholders that were represented on each Panel, as well as the organizations that were represented. For example, he noted that the Council of the Great City Schools, the National School Boards Association, the National Association of Elementary School Principals, and the National Association of Secondary School Principals were each represented on both mathematics and reading update panels.

Loveland reported on how panelists rated their experiences in each framework process. Most importantly, mathematics panelists reported that the opportunity to contribute their thinking was either very good or exceptional. Similarly, mathematics panelists reported that the opportunity for others to contribute their thinking was also very good or exceptional. Some mathematics panelists were surprised by how "big influencers" impacted the framework process. Reading panelists reported that opportunities to contribute their thinking and for others to contribute their thinking was satisfactory, very good, or exceptional. Some reading panelists expressed frustration that they were working "to make things more 'palatable'... for Board members."

Loveland reported on the public comment period and the related outreach that WestEd organized. Despite the COVID-19 pandemic, the number of public comments submitted for reading were about 50 percent higher than in the mathematics public comment period. Loveland also reported on the ways in which the Board engaged with the Panel. This involvement begins with the Board-adopted charge to each panel. Loveland observed that the initial charge for the NAEP Mathematics update prioritized maintaining trend, while the charge for NAEP Reading did not. Still, Loveland noted that both panels navigated tensions of supporting NAEP's innovations and relevance while also presenting changes that allow for stable reporting of student achievement trends. He stated that both panels scaled back their original vision in their public comment drafts based on Board feedback.

Loveland described common themes in both the mathematics and reading framework update projects. He noted that both framework update panels expressed strong interest in improving

NAEP reporting, which relates to many of each panel's contextual questionnaire recommendations. Loveland concluded by sharing reflections on panel size and structure.

Boyd thanked Loveland for his clear presentation. Miller asked about the selection procedures for identifying panel candidates and eventually selecting finalists to serve on each panel. Loveland noted that there is a fairly iterative process of review that involves Governing Board staff and Governing Board members. Miller asked if larger panels might be helpful to the framework process. Although each Panel has current and former classroom teachers, Loveland responded that if the Board wants to increase the number of panelists who are currently serving as classroom teachers, then it might be helpful to have a larger panel.

Kelly noted that it was important for the Board to improve the charges that are given to panels. For example, it would have prevented some frustration if reading framework panelists knew earlier that the Board highly prioritized trend. Kelly also asked for the Board to consider how Board feedback is delivered to the Panel. He noted that having the panel hear the Board's feedback in real time could be harshly received, counterproductive, and damaging to the Board's ability to recruit panelists for future framework projects.

Other Framework Processes: Initial Discussion

Boyd called attention to a white paper about assessment framework processes outside of NAEP (Attachment C in the Committee's advanced materials). She noted that this paper is also useful in thinking about potential process improvements suggested by the larger assessment landscape.

Boyd announced that the ADC is planning to convene with COSDAM in June for a joint committee session about framework processes. These discussions will help the ADC in detailing a procedures manual, which is one Strategic Vision activity.

NAEP Science Framework

Miller announced that the Board is gearing up for a review of the NAEP Science Framework. The ADC will need to develop a recommendation about whether this framework needs to be updated.

At recent ADC meetings, the Committee discussed the importance of gathering public comment early to inform deliberations. The Committee agreed that this is a good approach, and a draft public comment request is in the ADC's advanced materials (Attachment D).

Cunningham asked about the appropriate time for the Board to discuss merging the NAEP Technology and Engineering Literacy Framework with the NAEP Science Framework. Michelle Blair noted that the issue may surface in the initial public comment on the NAEP Science Framework, which would make it a formal part of the Governing Board's Framework Review and related Board deliberations.

Boyd adjourned the meeting at 7:28 p.m. E.T.

I certify the accuracy of these minutes.

Bard Dana Boyd, Chair

May 12, 2021 Date

National Assessment Governing Board Committee on Standards, Design and Methodology Report of May 3, 2021

COSDAM Members: Gregory Cizek (Chair), Carey Wright (Vice Chair), Jim Geringer, Eric Hanushek, Suzanne Lane, Alice Peisch, Julia Rafal-Baer, and Russ Whitehurst.

Other Governing Board Members: Dana Boyd, Christine Cunningham, and Patrick Kelly.

Governing Board Staff: Executive Director Lesley Muldoon, Deputy Executive Director Lisa Stooksberry, Michelle Blair, Laura LoGerfo, Munira Mwalimu, Sharyn Rosenberg, Angela Scott, and Matt Stern.

NCES Staff: Commissioner James (Lynn) Woodworth, Associate Commissioner Peggy Carr, Gina Broxterman, Samantha Burg, Jing Chen, Brian Cramer, Enis Dogan, Pat Etienne, Eunice Greer, Daniel McGrath, Nadia McLaughlin, Taslima Rahman, Eddie Rivers, Holly Spurlock, Bill Tirre, and Ebony Walton.

Other Attendees: American Institutes for Research: George Bohrnstedt, Markus Broer, Kim Gattis, Martin Hooper, Young Yee Kim, and Xiaying Zheng. CRP: Shamai Carter, Subin Hona, and Edward Wooford. Educational Testing Service: Jay Campbell, Gloria Dion, Amy Dresher, Helena Jia, Hilary Persky, Sarah Rodgers, Luis Saldivia, Karen Wixson, and Meng Wu. Hager Sharp: James Elias, Cailin Jason, and Joanne Lim. The Hatcher Group: Jenny Beard, Nandini Singh, and Jenna Tomasello. Pearson: Jennifer Galindo, Eric Moyer, and Cathy White. Westat: Lauren Bryne and Keith Rust. WestEd: Matthew Gaertner, Cynthia Greenleaf, Mira-Lisa Katz, Mark Loveland, Sonya Powers, and Steve Schneider. Other: Rebecca Bennett (Massachusetts Department of Education), Derek Briggs (University of Colorado – Boulder), Kathilia Delp, Donna Dubey (New Hampshire Department of Education), Laura Goadrich (Arkansas Department of Education), Michael Kolen (University of Iowa), Beth LaDuca (Oregon Department of Education), Raina Moulian (Alaska Department of Education), David Pearson (University of California – Berkeley), Linda Rosen, and Wayne Sheffield (North Carolina Department of Public Instruction).

Welcome and Overview of Agenda

Chair Gregory Cizek called the meeting to order at 12:02 p.m. ET and noted that the agenda included four topics: a brief update on the math and reading achievement level descriptions, a brief update on framework development processes generally, a discussion and question and answer session with NCES on the proposed NAEP Reading Framework, and an NCES presentation and discussion on NAEP Long-Term Trend for 17-year-olds.

Review and Revision of Mathematics and Reading Achievement Level Descriptions

Cizek began by noting that this topic has been on the COSDAM agenda several times and provided a brief explanation of achievement level descriptions (ALDs). NAEP has three achievement level policy definitions: *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*. ALDs translate these general policy definitions into specific expectations for a given subject and grade assessed by NAEP that are more informative about what students at each achievement level should know and be able to do. He explained that the ALDs provide important validity evidence for the NAEP achievement levels, and that the study to be conducted by Pearson will show us whether students within a given achievement level can actually do the things that the ALDs claim they should be able to do.

Cizek noted that a written update on the project and minor changes to the Design Document were provided in the advance materials. He invited Pearson project director Eric Moyer to briefly highlight the changes to the project since the March COSDAM meeting.

Moyer stated that the major project update was an extension of the project schedule to account for planning the panel meetings in person rather than virtually. With the additional time, the technical advisory committee (TAC) for the project spent more time reviewing the Design Document and procedures. This led to two additional recommendations: 1) removing the discrimination criteria for determining the anchor sets, and 2) conducting the item review by passage for reading.

Suzanne Lane complimented Moyer on the written materials and presentation. Cizek agreed and asked about the TAC's rationale for removing the discrimination criteria. Moyer noted that the criteria used in previous studies was norm-referenced, which meant that it would vary across the different subjects and achievement level categories. Pearson examined other alternatives and discovered that few items would be removed from the anchor sets regardless of the discrimination criteria employed, but that the excluded items might address important content that would not otherwise be represented because some of the anchor sets are fairly small. Cizek reiterated that the study would include items to illustrate the range of content that students know and can do at each achievement level without worrying about how well those items differentiate between students at different achievement levels.

Rick Hanushek stated that the large percentages of students—especially in some subgroups-scoring below the *NAEP Basic* level suggests to him that NAEP is not providing very useful information about those students. Cizek responded that he understands the concern about the need for better measurement and description of lower-performing students but that the concern is distinguishable from the purpose of the ALD project, which is based on the current achievement levels policy. Cizek noted that he would work with Board staff to plan a session for the August COSDAM meeting to begin discussing various options for improving measurement and reporting at the low end of the performance distribution.

There were no other questions or concerns about the project update or Design Document.

Framework Development Processes

Cizek explained that he has been thinking about the Board's framework development process and is interested in seeing how it can be improved; this is not meant to impact the Reading Framework update that is nearing completion, but rather to see how lessons learned from that process can inform the Science Framework. He began by differentiating between content standards and performance standards. Content standards define what is to be measured (knowledge, skills, and abilities), while performance standards indicate how high performance must be to be classified as *NAEP Basic, NAEP Proficient*, or *NAEP Advanced*. Framework development is focused on the first type of standards.

Cizek noted that the current Board policy on framework development was updated by the Board a few years prior and that it is important to be familiar with that policy. He highlighted three points in the framework process that he believes warrant further discussion to minimize tumult: 1) balancing the role of NAEP frameworks as a "mirror" versus "crystal ball" (i.e., the extent to which frameworks should reflect the current state of education versus trying to predict future practices); 2) reflecting contemporary practice (for example, by reconsidering the proportion of currently practicing teachers in the subject area); and 3) frequency of review (i.e., more frequent review could result in smaller, more incremental changes that could pose less threat to trend). Finally, he noted that the Board materials include two papers that will not be discussed at the current meeting due to time constraints but will provide background for an upcoming joint meeting of COSDAM and the Assessment Development Committee.

Russ Whitehurst responded to the crystal ball analogy, noting that NAEP has a special status because it can be self-fulfilling when NAEP tries to predict the future; that is, states may adjust their own assessment and instruction to maximize their performance on NAEP, even though NAEP is prohibited from influencing curriculum. Lane added that certification programs also grapple with trying to identify what knowledge and skills are emerging in the next 5-10 years; this challenge is not unique to NAEP or to educational assessment. Cizek clarified that part of the challenge is to figure out how to ease into emerging practices rather than trying to lead or influence what other people are doing.

Carey Wright observed that the Visioning Panel plays a large role in this process, and that she would appreciate knowing more about how they are selected and how they relate to the Development Panel. Cizek responded that the Development Panel is a subset of the Visioning Panel, and that he believes that the Development Panel has the most powerful role because they do the actual writing of the framework document.

Hanushek stated that the frameworks should be predictive of skills that are relevant to society and that increasing the proportion of educators to reflect current practice should be considered against increasing the input of employers and other users of NAEP data and the power of NAEP scores for predicting post-schooling outcomes. Cizek responded that the Visioning Panel does include several types of users of NAEP data but that there may be an opportunity to broaden the groups who are included.

Discussion of NAEP Reading Assessment

Cizek began by noting that the informational webinar on the Reading Framework held on April 30 was helpful; this COSDAM discussion is an additional opportunity to ask specific questions of NCES related to operationalizing the assessment. He reminded members that specific information about budget or secure items or passages could not be discussed during an open meeting.

Hanushek suggested that improved survey questions could result in better understanding the causal structures of reading and better policies to improve reading. He believes that these claims go beyond what NAEP can do, and that the framework should stick to defining what is being measured rather than how the data should be analyzed and used. He also raised concerns regarding the use of National School Lunch Program (NSLP) eligibility as the current measure of socioeconomic status because of the characteristics of those data.

Other Board members also expressed concerns about how NAEP currently measures socioeconomic status and suggested that alternatives be explored; they noted that existing problems with NSLP eligibility have been exacerbated by the pandemic because many states are now providing free lunch to all students. States are also struggling with how to measure socioeconomic status. Governing Board Executive Director Lesley Muldoon noted that the Reporting and Dissemination Committee has been discussing the measurement and reporting of socioeconomic status on NAEP, and that this topic is likely to be on a full Board agenda later this year. NCES Commissioner of Education Statistics Lynn Woodworth added that NCES is aware of the problems with the current measure and has been working with the Census Bureau on developing an alternative to NSLP eligibility using a kriging modeling methodology that estimates the socioeconomic status of a given point on a map. This work has been time consuming and has not yet been vetted strongly enough to implement in operational NAEP.

Cizek referred members to the documents in the advance materials prepared by NCES to address open questions from the Board. He stated that he has an interest in understanding whether the design elements intended to provide topic knowledge are having their intended effect; that is, do they differentially impact students who lack that topic knowledge? Woodworth responded that previous studies have not attempted to capture students' topic knowledge; additional research to address this question might be able to be conducted if resources allow, but he was not willing to commit to such a study at this time.

In response to two questions about the size and representativeness of the student samples for cognitive interviews and pretesting, Eunice Greer of NCES noted that typically cognitive interviews are conducted individually with 9-12 students, small scale tryouts include 50-100 students, and pilot studies use regular NAEP samples of 2500-3000 students. NCES works with their contractors and provides sampling guidelines to ensure that a diverse student population is included in these activities.

Several Board members raised questions about the NCES written response stating that the cost of implementing a framework can only be determined after Board approval of that framework. Some members expressed concerns about adopting a framework without understanding the costs upfront. Associate Commissioner of NCES Peggy Carr, clarified that the costs largely

depend on how much new item development and field testing is needed, along with bridge studies. She indicated that she was unable to provide specific numbers during this open discussion but that she would prepare some information about the cost of implementing the current draft framework for an upcoming closed budget briefing. Holly Spurlock added that NCES submits a formal response memo to all frameworks during the public comment period and focuses on any concerns related to framework implementation, including cost. For example, during the development of the math framework, NCES indicated that the large number of scenario-based tasks called for in the public comment draft would be very costly; the final math framework included a reduction of scenario-based tasks.

In response to questions about whether the existing universal design elements (UDEs) increase testing time, Spurlock indicated that there is a wide range of UDEs and that NCES does not consider most of them to be accommodations. The current tasks are developed as 30-minute segments and incorporate elements that are necessary for good measurement practice; UDEs are part of this development rather than being conceptualized as an add-on. Woodworth noted that process data could be analyzed, if time and resources allow, to examine how much time students spend on some of the existing UDEs.

Cizek noted that it is considered best practice to screen reading passages to ensure that everything students need to answer the assessment questions is included in the passage. He asked whether NCES has a procedure to ensure that students do not need to rely on topic knowledge to answer questions on the reading assessment. Greer responded that NCES does attend to this concern in the development and review process. Spurlock added that the Assessment Development Committee reviews all reading passages before NCES even begins writing assessment items.

Technical Advisory Committee (TAC) member Derek Briggs stated that many of the concerns raised by COSDAM were also the subject of discussions between the TAC and the Visioning Panel; throughout the process, part of the TAC's role was to remind the panel that NAEP operates within certain constraints. Some of the panel's original vision and aspirations were scaled back from the initial draft. In the current draft of the framework, the knowledge-based UDEs are not intended to serve as mini tutorials to teach students about a topic; rather they are very brief attempts (e.g., two sentence introductions or a short pop-up definition) to prime students and provide basic familiarity with a context.

TAC member Scott Marion agreed with Briggs and stated that he views NAEP as the most technically sound assessment with the most extensive development process in the country. He believes that the typical NAEP development process will provide data to address the questions raised by COSDAM members.

Cizek thanked NCES for the helpful information that they provided.

Discussion of 2022 NAEP Long-Term Trend

Cizek noted that the final agenda topic was related to the Long-Term Trend (LTT) assessment; he introduced Enis Dogan of NCES.

Dogan began by noting that the LTT assessment is age-based rather than grade-based; the assessments for 9-year-olds, 13-year-olds, and 17-year-olds are administered during different times of the school year. The assessment was scheduled to be conducted during the 2019-2020 school year; data collection for 9-year-olds and 13-year-olds was completed as scheduled but the 17-year-old administration was scheduled to begin in March 2020 right when schools shut down due to COVID-19. The Governing Board subsequently moved the LTT assessment of 17-year-olds from 2020 to 2022 on the NAEP Assessment Schedule. Dogan stated that the purpose of this session is to ensure that COSDAM members are informed about technical considerations related to the planned administration of LTT for 17-year-olds in 2022.

Dogan noted that although the data from the three age groups were scaled together during the initial 1971 assessment, the age groups have subsequently been scaled separately. From the perspective of scaling, there is no concern with conducting the 17-year-old data collection at a different point in time than the other age groups.

In terms of item functioning, Dogan reported that there were some challenges with the 2020 data collection, particularly for 13-year-olds. Several items had to be split during data analysis; that is, they were treated as if they were different items in 2012 and 2020 even though the items themselves did not change. Dogan hypothesized that even more items could require splitting for 17-year-olds given the larger gap from the prior administration and potential effects from the pandemic. He concluded that this challenge is not insurmountable given the data analysis tools available.

Dogan described concerns related to interpretation of the results for 17-year-olds. Depending on the status of COVID-19 in 2022, changes to administration conditions could be a confounding variable impacting the results. In addition, if there has been a lot of learning loss over the last couple of years, 17-year-old results could be closer to the typical 13-year-old results, which could also pose challenges to their interpretation.

Finally, Dogan noted that there may be limited utility to collecting data on 17-year-olds in 2022. One of the primary purposes for the 2020 LTT data collection was to establish one final data point before transitioning the LTT assessment to a digital platform in 2025. The field test for the 2025 LTT administration will take place in January 2023, and there is not enough time to analyze the 2022 LTT data to allow for changes prior to the field test. Dogan concluded his presentation by stating that there were several technical issues to consider but that the challenges were not insurmountable.

Hanushek requested clarification on the concern related to testing 9-year-olds and 13-year-olds prior to the pandemic and 17-year-olds after the pandemic. Dogan explained the need for some degree of match between what the assessment is measuring and what students know; the concern is partly that precision could be reduced if some items are not providing much information about what students know. In addition, there are concerns about interpreting the results given the potential for some users to compare performance across the three age groups when the pandemic affected only the measurement of 17-year-olds.

Lane asked about plans for examining item misfit and whether there is an analysis of the content of such items. Dogan responded that data to model fit is examined routinely, along with

differential item functioning analyses. He explained that the content of problematic items is examined as well. For example, if most problematic items were in a single content area, that could indicate a systematic issue that could affect the representativeness of the construct.

Woodworth added that the unique circumstances of the pandemic could cause some items to function differentially in 2022 but return to functioning normally in 2025. It would be difficult to know how to interpret whether item performance fluctuations are due to COVID or due to diminishing relevance or other factors until there is an additional data point available. This is one reason why Woodworth is hesitant to proceed with the 2022 LTT administration.

Julia Rafal-Baer asked whether it is likely that COVID will continue to have an impact on the data collected, given that many students and families are still reluctant to return to school in person. Woodworth responded that he is aware of concerns that some students will never fully recover from the educational impacts of the pandemic, but that he believes that the deficit for 17-year-olds will be smaller in 2025 than it would be in 2022.

Hanushek countered that he is interested in what individual students can do and argued that the group of 17-year-olds in 2022 is a different cohort than the students who will be 17-year-olds in 2025. He does not believe there is a benefit to dropping the 2022 data point.

Cizek complemented the NCES staff on the presentation and discussions throughout the meeting and noted that Board members will keep this information in mind when making a decision about whether to proceed with the 2022 LTT assessment for 17-year-olds.

Cizek adjourned the meeting at 2:02 pm ET.

I certify the accuracy of these minutes.

Gregory Cizek, Chair

<u>July 15, 2021</u> Date

Reporting and Dissemination Committee Meeting

May 10, 2021 10:00 am - 12:00 pm

Reporting and Dissemination Committee Members: Chair Tonya Matthews, Vice Chair Marty West, Alberto Carvalho, Tyler Cramer, Paul Gasparini, Governor Bev Perdue, Ron Reynolds, Mark White

Governing Board Staff: Stephaan Harris, Donnetta Kennedy, Laura LoGerfo, Lesley Muldoon, Munira Mwalimu, Sharyn Rosenberg, Angela Scott, Matt Stern

National Center for Education Statistics Staff: Gina Broxterman, Samantha Burg, Peggy Carr, Jing Chen, Brian Cramer, Jamie Deaton, Pat Etienne, Doug Geverdt, Eunice Greer, Daniel McGrath, Taslima Rahman, Holly Spurlock, William Tirre, Ebony Walton, William Ward, Grady Wilburn, Lynn Woodworth

Contractors: <u>AIR</u>: George Bohrnstedt, Markus Broer, Cadelle Hemphill, Young Kim, Xiaying Zheng; <u>CRP</u>: Shamai Carter, Subin Hona, Emilie Pooler, Anthony Velez, Edward Wofford; <u>ETS</u>: Jonas Bertling, Hugo Dos Santos, Amy Dresher, Gloria Dion, Robert Finnegan, Laura Jerry, Sami Kitmitto, Courtney Sibley, Lisa Ward, Ryan Whorton, Karen Wixson; <u>Hager Sharp</u>: James Elias, David Hoff, Joanne Lim; <u>The Hatcher Group</u>: Sami Ghani, Zoey Lichtenheld, David Loewenberg, Nandini Singh, Jenna Tomasello; <u>Optimal Solutions Group</u>: Imer Arnautovic, Sadaf Asrar, Charlotte Notaras; <u>Pearson</u>: Eric Moyer, Stanley Rabinowitz, Pat Stearns; <u>Silimeo Group</u>: Debra Silimeo; <u>Westat</u>: Chris Averett, Kavemuii Murangi, Lisa Rodriguez

Other: Rebecca Bennett (Massachusetts Department of Education), Kathilia Delp, Donna Dubey (New Hampshire Department of Education), Jeremy Ellis (Missouri Department of Education), Jasmine Fletcher-For, Laura Goadrich (Arkansas Department of Education), Beth LaDuca (Oregon Department of Education), Regina Lewis (Maine Department of Education), Rebecca Logan (Oklahoma Department of Education)

Chair Tonya Matthews called the Reporting and Dissemination Committee meeting to order at 10:00 am on Monday, May 10, 2021. Matthews provided an overview of the agenda and the goals for the meeting.

The meeting began with a focus on the Strategic Vision. Matthews briefly described both the general pillars and the specific goals for which the Reporting and Dissemination (R&D)

Committee serves as primary agent. Much of R&D's work resides within all of the pillars, with a few particularly salient to R&D's strengths. For example, the Board's Strategic Vision includescircumscribing the purpose and use of NAEP, specifically what NAEP can inform and what NAEP cannot. To that end, R&D members have urged the Board to develop a body of evidence to improve interpretations of NAEP.

Matthews noted that more than one committee tackles the work for each pillar. Even a goal about setting the assessment schedule falls to R&D, because the Board needs to communicate the expected schedule (and any changes) and to elicit feedback from stakeholders on whether the Board's communications efforts succeed.

Laura LoGerfo, Governing Board assistant director for reporting and analysis, then outlined how the Strategic Vision's pillars and goals drive the Board staff's work plans. The communications contractor, The Hatcher Group, then executes the plans developed and approved by R&D through activities included in the recently updated communications and outreach plan.

LoGerfo provided an example of how this work develops in ways that leverage the resources of the small staff. One of the Strategic Vision's pillars is inform, which leads to meetings with stakeholders to determine their interests in NAEP data. These interests become incorporated into release events and post-release activities and materials, all of which are presented in the communications plan. For a specific example, representatives of school districts in the Trial Urban District Assessment (TUDA) program convene as part of the Board's TUDA Task Force, led by the Council of the Great City Schools. The Board considers these TUDA districts as priority stakeholders, seeks their feedback and insights, and develops strategies based on this input.

Vice Chair Marty West acknowledged how these examples validate R&D's contributions to the Strategic Vision and affirmed the prominence of the inform and engage pillars as uniquely suited to R&D. He explained that where the Board cannot align neatly with stakeholder needs, when the Board or NAEP cannot directly answer questions, those gaps can be addressed with innovation, the third pillar in the Strategic Vision.

Communications and Outreach: Accomplishments and Aspirations

With this context established, Stephaan Harris, the Governing Board's assistant director of communications, and Robert Johnston, lead for the Board's communications contractor, shared achievements from the last communications plan and outlined plans for the next two years.

Johnston presented first, crediting the daily collaboration between Board staff and the Hatcher team, for recent successes. As West observed moments earlier, the two pillars of inform and engage guide activities, which can be organized as (1) release events and post-release activities; (2) social media management; (3) artifacts such as videos and graphics featuring NAEP data; (4) community outreach events, e.g., outreach dinners in off-site Board meeting locations like El

Paso and Montgomery; (5) Governing Board member recruitment campaigns; and (6) a monthly newsletter sent to 5,600 recipients.

In reviewing the Board's social media strategy on Twitter, Facebook, and LinkedIn, the Board increased its followers across all platforms, and permission to use paid digital ads on Facebook boosted impressions of the Board's posts about release events. Evidence indicates that this small investment reaps significant rewards. In 2018, the Board seldom engaged with LinkedIn, but current practice focuses on this platform as a critical component of a tripartite social media strategy.

Attention then shifted to Stephaan Harris who explained the Governing Board's new communications and outreach plan to the committee. The Strategic Vision adopted by the Board in 2020 conveniently coincided with the contractually obligated update to the Board's communications plan. Thus, the plan neatly aligns its content with the Strategic Vision.

At the start, the communications plan declares the priority audiences for the Board's outreach work as education administrators, researchers, education advocacy organizations, and policymakers. This declaration derives from the Strategic Vision itself and from several meetings with Board staff. Introducing NAEP and the Board's work to new audiences represents Inform, after which sustained collaboration can reflect the Engage pillar. The Board needs a systemic approach to establishing regular avenues of communication with current collaborators and with new stakeholders, such as social media, media outreach, conferences, and emails. Such collaboration compels reciprocation; heeding stakeholders' needs while also soliciting their assistance in disseminating the Board's messaging.

These messages center most squarely on the unique and valuable data NAEP provides about how the nation's students are learning. Tantamount in importance, NAEP, the gold standard in student assessment, serves as a catalyst for action to improve student achievement. Equitable education policy and outcomes rely on understanding what students know and can do. Vice Chair West noted that many recommendations in the plan call for either deepening relationships with a few stakeholders or broadening the Board's reach to include new additional voices. He suggested that the most effective, expedient impact may emerge first from delving more deeply into NAEP with a select few stakeholders, then leveraging that initial work to broaden the audience and enhance the Board's ability to attract new stakeholders.

In the near future, Hatcher is interested in elevating the profiles of Board members and their diverse areas of expertise as a feature in journalists' stories not only in education-related media but also in more mainstream outlets that publish on broader issues to wider audiences. This fall, the Board can capitalize on the postponement of NAEP Day (the biannual initial release of results in NAEP Reading and Mathematics) to host an event about related topics like the divergent trend lines or to showcase the transcript data, which will be released sometime in Autumn 2021.

Matthews then invited questions and discussion, recommending R&D members consult Attachment B which lists upcoming releases and reports, including the upcoming 2019 NAEP Science release event on May 25. Matthews cautioned that the Board must be sensitive in releasing data at this moment and acknowledge the lag between the collection and release of data. Namely, these data capture performance prior to COVID. Matthews urged the committee members to consider additional communications challenges which they anticipate may affect imminent releases.

NCES Commissioner Woodworth asked Matthews if he could share a few updates with the committee. NCES is currently building an equity dashboard tool for the U.S. Department of Education, which should include information from the Governing Board, or even lead to a companion website for the Governing Board. He encouraged the Board to collaborate with NCES on this front. Marty West expressed appreciation for this news and emphasized that the Board should deem NCES and the Department itself as critical stakeholders for communications efforts.

West asked when the Governing Board expects to return to assuming that events should be held in person, with virtual as only a back-up option. Or, should the Board eliminate the in-person experience, because the Board gains more and more widespread viewers through remote attendance? He also wondered why the monthly school COVID-19 survey, administered through the NAEP infrastructure, is not branded as a NAEP product? Commissioner Woodworth explained that the school survey data do not undergo the same rigorous quality control checks as the NAEP data, thus NCES distinguishes these survey results from NAEP so as not to raise doubts about NAEP data.

Ron Reynolds praised the outreach plan as comprehensive and well-conceived and observed how the plan strikes the appropriate balance between broadcasting and narrowcasting. Broadcasting NAEP means informing those in the education field unfamiliar with NAEP, but who, once cognizant of NAEP's value, could inform others. But Reynolds could perceive the expedient benefits of engaging key stakeholders, or narrow-casting. He then asked for examples of how the Board should engage the general public. Harris described the general public as those who do not normally seek NAEP data or who are unfamiliar with NAEP, e.g., advocacy groups which do not specialize in education and media who cover issues other than education. Also, in response, Matthews coined a new term of "in-between casting," i.e., casting to specific task force members, congressional leaders, and superintendents invested in NAEP.

Matthews replied that the Board grapples with this issue perennially. Because NAEP does not capture school-level or student-level data, proving NAEP's value to the general public becomes challenging, but not impossible. The Board aims to engage parents and parent advocates, which is currently accomplished through recruiting them as nominees for general public representative Board seats.

The presentation on the outreach plan inspired several questions from general public representative, Tyler Cramer. Cramer asked who on the Board staff identifies current and prospective NAEP users, because the Board needs to research who these audiences are (Hatcher and Board staff jointly undertake this task). He urged staff to include state legislators and the legal community among key stakeholders to approach. Equity is typically decided in the courts, so the Board should teach lawyers how to use NAEP. Cramer encouraged NAEP to spearhead the effort to decide what equity means for educational assessment. Cramer also expressed concerns about the NAEP Reading assessment framework, which calls for new reporting on new contextual data. How can the R&D Committee, whose jurisdiction includes responsibility for core contextual variables, address these recommendations?

This session concluded, and Matthews transitioned to the next agenda item by affirming Cramer's point that more attention should be paid to equity in assessment. Updating the assessment frameworks provokes intense questions of equity, which may require engaging with new audiences.

Measuring SES: Recap & Discussion

Equity, or a lack thereof, can be found through understanding differences in performance. A typical measure of equity in school systems is the extent to which family background generally--and SES specifically--factors into student outcomes. NAEP should be a tool to measure the extent to which student achievement depends on SES and how that relationship changes over time. However, that is a challenging duty, given the relatively imprecise data NAEP collects on student SES.

At the March meeting of the R&D Committee, Markus Broer, a researcher at the American Institutes for Research, Eric Hanushek, a current Governing Board member and scholar at Stanford University, and Tom Kane, a professor at Harvard University, presented their suggestions on how to improve NAEP's measure of SES. Their thought-provoking presentations left the committee with little time then to discuss the proposed approaches. Beyond allotting ample time for member discussion at this meeting, the committee also invited NCES staff to share their reactions as to the technical feasibility of each proposal.

West began by reviewing the proposed approaches and the assumptions which 'constrain' the feasibility of those approaches, such as using extant variables, so that any new iteration of SES can be applied to previous data to chart trends across time. However, fixating on that proviso precludes new information and prevents a shift to a more precise, improved measure. West encouraged R&D members and NCES to think more flexibly about that principle as an absolute constraint. West ceded the intractability of a few challenges, e.g., the meaning and interpretation of contextual questions may change over time and missing or erroneous data, because the data come from student self-reports on their family circumstances.

To jog committee members' memories, West summarized the three proposals from the March meeting:

- Broer measures SES in an additive index comprising (1) number of books at home; (2) student eligibility for the National School Lunch Program (NSLP); (3) school percent of students eligible for NSLP; (4) parents' highest level of education. This index offers stronger explanatory power than eligibility for NSLP alone and more than the SES index in PISA data. Broer examines how the relationship between SES and student achievement has evolved over time, but finds little evidence of change, just modest narrowing of gaps.
- Hanushek's proposal aligns closely with Broer's approach, with its emphasis on household possessions. Unlike Broer, Hanushek employs principal component analysis to construct his SES index.
- Kane eschews all variables but household income, which is one component of SES but does not reflect the entire construct. The NAEP program does not collect these data, but Kane links NAEP schools in the restricted-use datafile to neighborhood income data from the Census Bureau. He and his team aggregate the data and analyze measures of variance of achievement within schools and the variance of income in the schools' surrounding neighborhoods to draw inferences about the relationship between income and achievement
 - This approach assumes that schools in NAEP generally serve their immediate surrounding neighborhoods, but that assumption is flawed to a degree, given the variable prevalence of charter schools, other schools of choice, and the nature of gentrification.
 - Kane and his colleagues find a narrowing of differences over time, consistent with others' findings, even though income inequality has increased over time.

West concluded his review by inviting NCES staff to respond, and Dan McGrath, director of reporting for NAEP, addressed the committee first. He agreed that a measure of SES on NAEP should not be constrained only to variables previously administered. Such a constraint would benefit those invested in maintaining trend, but that should not be a mandatory requirement. Indices like what Broer and Hanushek use allow for changes in the meaning of individual variables over time while conserving the overall meaning of the construct for analyses of trend. In short, indices depend on the importance of the collective, not the individual items which comprise the index. For example, an index can include specific items on what technology students own or use regularly, but the specific variable can transform from asking about Palm Pilots to smart phones while retaining the actual meaning of "technology use in the home."

Bill Ward of NCES chimed in next and deemed none of the three proposals were too outlandish. In fact, this might be the appropriate time to explore how to enact some of the ideas. Ward also suggested that supplemental reporting about new indices may prove illuminating and address some concerns with the SES data.

Ebony Walton of NCES shared a caution about using the NSLP eligibility indicator in NAEP. She warned that even before the Community Eligibility Provision (CEP) for NSLP (above a certain percentage of students in a given school eligible for free- and reduced-price lunch, the entire school is classified as eligible), the NSLP indicator did not measure individual need. Walton drew an analogy to Title I. The NAEP program already asks the school administrator whether the school provides Title I for individual students or for the entire school. Perhaps NAEP could edit the wording of the NSLP item on NAEP to mark that same distinction. Ebony shared that the relationship between NSLP and performance is growing weaker, because changes in policy have changed how schools participate in NSLP, all of which disrupts the interpretation of trend. Lynn Woodworth added that COVID caused further change, because the U.S. Department of Agriculture extended free lunches to every student across the country during the pandemic. This gives further proof that education researchers should not rely on this variable as a measure of student need.

To that point, Matthews suggested capturing indicators of environmental equity, which refers to resources available to students in their communities, such as nearby libraries, which relate to student learning. She also urged NCES not to overlook mental wellness, which may transcend traditional notions of socioeconomic status but certainly affect student achievement. Both Woodworth and West supported the inclusion of measures of environmental equity or linking to datasets with such information.

Broer requested to reply to queries and comments raised by R&D Committee members. He discussed his analyses of NSLP eligibility data at both the state and national levels through 2019 which show continued relevance for this objective measure. Admittedly, schoolwide eligibility quadrupled since the implementation of the Community Eligibility Provision (CEP), but schools can determine individual eligibility, which has remained stable. Prior to CEP implementation, which began in 2011, schools with 75-100% of their student population eligible for NSLP already basically had schoolwide eligibility, so the policy change made no impact for them. Indeed, the estimates before and after 2011 support this claim. Since 2011, the NSLP eligibility indicator has proven an even stronger predictor of performance in some states and some urban school districts in the Trial Urban District Assessment program.

In Broer's analyses, variables about household possessions do not add precision or reliability to the index and actually reduce internal consistency. To which West inquired why Rick Hanushek persists in using them if such items do not add much unique value to the SES index.

Cramer shifted the focus from the federal lunch program to the private sector and wondered if NAEP could partner with Amazon, Costco, and other companies to procure these companies' data on personal consumption in neighborhoods. Or simply ask students whether they own their own cellular telephone. Broer explained that NAEP at one point did inquire about cell phones at home and currently asks about smartphones. These data showed that students who had high SES through other measures had a steep drop in cell phone usage, students in mid-range SES had no change, and low-SES students increased their cell phone ownership and use. Not to overinterpret the data, but these findings may be linked to higher-SES parents now more likely to regulate their children's time with technology and lower-SES students using cell phones as the only means to access the internet. West suggested that these results would make a fascinating presentation (duly noted by staff for future meetings!).

Cramer again raised a familiar question about the interoperability of NAEP data with other federal datasets such as those available from the U.S. Census and the U.S. Department of Labor. West reassured Cramer that analysts with restricted use data licenses can link individual students on NAEP to any measure of local geography, which is Tom Kane's approach. However, depending on neighborhood SES data to impute or infer school SES can pose a challenge, because in gentrifying urban areas, students enrolled in the school do not necessarily reside in the surrounding neighborhood.

West asked about the legality of Kane's proposal to gather information about students' addresses only temporarily until a match to aggregated income data is made, when all personally identifiable information (PII) is discarded. Woodworth clarified that this approach lies outside the law. However, McGrath shared that NCES is already pursuing an approach akin to Kane's. This work is led by Doug Geverdt who has invested more than two years into geocoding addresses of student participants in NAEP and building a poverty index from the Common Core of Data for all schools in the nation.

Geverdt and his SIDE project team at NCES work with schools where the NAEP program lists sampled students' names, and the schools append data relevant to students' addresses and pass that relevant data to NCES. This follows the same process as collecting data on NSLP eligibility. States and districts hold the student-level information, not the federal government, which side steps any worries about PII.

The unique value of SIDE estimates dwells in their specificity; Geverdt can produce address-specific estimates of income-to-poverty ratios, which uses a continuous measure from 0 to 999, with 100 equal to the poverty threshold, 130 and 185 the income-to-poverty ratios that determine NSLP eligibility, and higher numbers corresponding to further distance from poverty. The 0-999 continuum allows researchers to peer past the current 185 cap on the measure and examine middle income and affluent schools as well, which facilitates new analytic possibilities.

Geverdt essentially attains a measure strong in reliability like tract-level estimates (based on more data, so high reliability) but with the precision of the block level, which captures the best of both approaches. If NCES can convince states about the benefits to this strategy, then NAEP can request SIDE estimates, just like the typical request for NSLP data. Through NCES' Statewide Longitudinal Data Systems grant program, 16 states already provide information to check SIDE estimates as a feasibility test.

West inquired whether researchers can link or append those poverty indices to schools in the restricted-use NAEP datafile. Geverdt demurred that this effort should be classified as a pilot test at this point, with the geospatial infrastructure required for its operation still nascent. But there is hope for the relatively near future. Maryland is developing alternative measures to capture SES beyond NSLP. The Council of the Great City Schools and the Atlanta school district use zip codes to capture corresponding Census data on SES.

Reynolds asked a process question. Adding contextual variables to NAEP seems like both a policy question, under the Governing Board's domain, and a technical question, NCES' bailiwick. Whose responsibility is it? West and LoGerfo explained that the Board decides on the constructs of importance to include, and NCES gives feedback on technical feasibility.

Walton underscored the importance of including contextual variables on NAEP itself. Data linkages work well for secondary research, but obtaining additional data from external sources to predict student outcomes may produce biased results. The NAEP conditioning model, which generates plausible values for the assessment, uses data from the questionnaires. Thus the most accurate statistics about students' backgrounds come from the NAEP dataset.

In the final moments of the meeting, McGrath summarized next steps. He and the NCES team would explore items pertinent to environmental equity, which can be obtained through geocoding, as well as to mental health and wellness, which may be tricky to operationalize within the constraints on intrusiveness. LoGerfo noted later that the TUDA Task Force requested communication from the Governing Board that acknowledges the issues with relying only on NSLP eligibility to measure SES and updating them about progress underway to improve the measure.

At 12:01 pm, the meeting concluded. Matthews made a motion to adjourn, which West seconded.

I certify the accuracy of these minutes.

M.h. Mathies

Tonya Matthews, Chain

July 20, 2021

National Assessment Governing Board

Nominations Committee

May 11, 2021

Open Session

Nominations Committee Members: Governor Jim Geringer (Chair), Dana Boyd, Tyler Cramer, Paul Gasparini, Tonya Matthews, Reginald McGregor, Mark Miller, Alice Peisch.

Board Member: Suzanne Lane.

Board Staff: Stephaan Harris, Donnetta Kennedy, Lesley Muldoon, Munira Mwalimu, Tessa Regis, Lisa Stooksberry.

Governor Geringer welcomed members and thanked Board member Suzanne Lane for attending the meeting. He reviewed the agenda topics for discussion.

Outreach Strategy for 2022 Nominations Campaign

Stephaan Harris, Assistant Director of Communications, briefed the committee on the outreach strategy for the 2022 nominations campaign and the role of the communications contractor, the Hatcher Group, in this effort. In the 2021 campaign staff and contractors pursued typical strategies such as calls with key stakeholders and extensive social media outreach. However, new activities included a well-received webinar with the Governing Board alumni, who were asked to tap into their professional networks to get the word out on the campaign. Harris reported that the 2021 strategy yielded success with a large, diverse pool of candidates.

The 2022 campaign will be launched via a splash page in the summer of 2021 before the site is open to applicants. Vacant positions in 2022 include 4th Grade Teacher, 8th Grade Teacher, Secondary School Principal, and General Public Representative (Parent Leader). A tool kit will be developed, and a webinar will be convened to attract candidates for all open categories with a focus on the parent leader's category. In addition, outreach will include a range of organizations with networks of potential candidates for the open positions (e.g., the National Parent Teacher Association (PTA), among many others). Harris suggested that Tonya Matthews write a blog about her experiences serving on the Board as a General Public Representative.

A survey will be sent to Board members this summer to solicit outreach ideas.

Procedures Manual

Geringer asked members if they had any suggested edits to the Nominations Procedures manual. He referenced a previous request for clarification regarding qualifications in the Testing and Measurement category. He reported that the Office of General Counsel (OGC) had provided guidance on candidate eligibility in this category, noting that candidates who work at a testing company are not automatically disqualified. Circumstances for candidates are reviewed and determined on an individual basis by OGC.

There were no changes suggested to the Procedures Manual. Tessa Regis reported that she planned to survey members to evaluate the 2021 review and rating processes and to also request feedback on the Procedures Manual.

Closed Session

Nominations Committee Members: Governor Jim Geringer (Chair), Dana Boyd, Tyler Cramer, Paul Gasparini, Tonya Matthews, Reginald McGregor, Mark Miller, Alice Peisch.

Board Member: Suzanne Lane.

Board Staff: Stephaan Harris, Donnetta Kennedy, Lesley Muldoon, Munira Mwalimu, Tessa Regis, Lisa Stooksberry.

Under the provisions of exemptions 2 and 6 of § 552b (c) of Title 5 U.S.C., the Nominations Committee met in a closed session on Tuesday, May 11, 2021 from 6:10 p.m. to 6:30 p.m. to receive a briefing on the slate of 2021 finalists for submission to the Secretary for consideration and appointment.

Lisa Stooksberry reported on the status of 2021 candidates and referenced internal preparations for a forthcoming meeting with the Secretary's office. Geringer then provided a briefing on a technical issue that has since been resolved. The committee discussed this matter with an eye toward preventing such issues in the future.

Geringer thanked the Nominations Committee members for their efforts and staff for their support.

The meeting adjourned at 6:30 p.m.

I certify the accuracy of these minutes.

Jim Geringer Jm Geringer, Chair

July 9, 2021 Date

Executive Committee

August 5, 2021 10:30 am – 12:00 pm ET Salon I & II, 5th level and via Zoom

AGENDA





Change to Long-Term Trend Assessment Administration in 2022

During the May 2021 Board meeting, there was discussion about amending the NAEP Assessment Schedule to replace the age-17 LTT with age-9 LTT assessment in 2022.

Since age-9 LTT was the last NAEP assessment administered before COVID, the rationale behind the proposed change is that repeating age-9 would provide valuable data of "learning loss" to researchers that want to analyze a comparison of 2020 to 2022 results.

NCES has confirmed that they would be able to operationalize this change and is already preparing to do so, in anticipating of potential action by the Board at this August Board meeting.

In addition, NCES has reported no additional impacts to the NAEP budget as a result of making this change to the assessment schedule.

Please see the amended Assessment Schedule that reflects this change.



National Assessment of Educational Progress Schedule of Assessments Approved March 5, 2021

The *National Assessment of Educational Progress (NAEP) Authorization Act* established the National Assessment Governing Board to set policy for NAEP, including determining the schedule of assessments. (P.L. 107-279)

		National	State	TUDA
Year	Subject	Levels	Grades	Grades
	, i i i i i i i i i i i i i i i i i i i	Assessed	Assessed	Assessed
2020	Long-term Trend*	9-year-olds		
		13-year-olds		
2021				
2022	Reading	4, 8	4, 8	4, 8
	Mathematics	4, 8	4, 8	4, 8
	Civics	8		
	U.S. History	8		
	Long-term Trend*	9-year-olds		
2023				
2024	Reading	4, 8, 12	4, 8	4, 8
	Mathematics	4, 8, 12	4, 8	4, 8
	Science	8		
	Technology and Engineering Literacy	8		
	Transcript Studies			
2025	Long-term Trend	~		
2026	READING	4, 8	4, 8	4, 8
	MATHEMATICS	4, 8	4, 8	4, 8
	Civics	8		
	U.S. History	8		
2027				
2028	Reading	4, 8, 12	4, 8, 12	4, 8
	Mathematics	4, 8, 12	4, 8, 12	4, 8
	SCIENCE	4, 8	4, 8	4, 8
	Technology and Engineering Literacy	8	8	
	Transcript Studies			
2029	Long-term Trend	~		
2030	Reading	4, 8	4, 8	4, 8
	Mathematics	4,8	4, 8	4, 8
	CIVICS	4, 8, 12	8	
	U.S. HISTORY	4, 8, 12	4 0 10	1.0
	WRITING	4, 8, 12	4, 8, 12	4, 8

NOTES:

* Long-term Trend (LTT) assessment not administered by computer until 2024. All other assessments will be digitally based.

~ LTT assessments sample students at ages 9, 13, and 17 and are conducted in reading and mathematics.

BOLD ALL CAPS subjects indicate the assessment year in which a new or updated framework is implemented, if needed.

The National Academies of SCIENCES • ENGINEERING • MEDICINE

OPPORTUNITIES FOR NAEP IN AN AGE OF AI AND PERVASIVE COMPUTATION: A PRAGMATIC VISION

NASEM Project Discussion with the NAGB Executive Committee: August 5, 2021

NASEM discussants

• Karen J. Mitchell, Chair, NASEM Panel

Karen recently retired as senior director of the Medical College Admission Test (MCAT) at the Association of American Medical Colleges (AAMC). At AAMC, Karen oversaw the work of her colleagues on test development and scoring, test administration and reporting, test preparation services, testing research, and outreach and communication. She directed the redesign and 2015 launch of the current version of the MCAT exam and directed its continued administration during COVID-19.

• Stuart W. Elliott, Study Director, NASEM Panel

Stuart is a scholar at NASEM where he has directed numerous studies related to testing and assessment. For 10 years he served as the director of NASEM's Board on Testing and Assessment. He also spent three years at the OECD working on the PIAAC assessment of adults.

Discussion questions

- NAGB's 2025 strategic vision includes an aim to innovate to achieve your policy priorities of utility, frequency, and efficiency. How should the panel think about the relative importance of cost cutting, as emphasized in our Statement of Task (see below), from the perspective of NAGB's priorities?
- The panel hopes to makes its recommendations clear and actionable, in terms of decision-making authority, cooperation between organizations, and realistic expectations about costs, cost savings, and timelines. Do you have any suggestions about the things we should think about in making recommendations that will be clear and actionable?

NASEM Project Statement of Task

The National Academies of Sciences, Engineering, and Medicine will appoint an ad hoc panel to consider several innovations that could substantially reduce the cost structure of NAEP while maintaining its technical quality and value in informing the public about education progress. The panel will review the major cost components of NAEP and related assessment programs and consider the following possible changes to the NAEP program: 1) automatic item generation; 2) remote test administration; 3) computer adaptive testing; and 4) consolidation and elimination of substantive overlaps between NAEP assessments and between NAEP and other assessments, such as PISA, TIMSS, and PIRLS. The panel will also solicit and consider suggestions of other major changes that reflect modern methods of assessment and that could substantially reduce NAEP costs while largely preserving its technical quality and informative value. The panel will review relevant research and industry practice to draw conclusions about the likely effects of these potential changes on the cost, technical quality, and informative value of NAEP.

The panel will produce a short and broadly accessible report that summarizes its findings and conclusions about these potential changes to NAEP and recommends potential assessment or programmatic changes and research needed for NAEP to explore innovations while balancing the competing objectives of cost reduction, technical quality and informative value.

Panel Members

- Karen J. Mitchell, Chair, Association of American Medical Colleges (retired)
- Isaac I. Bejar, Educational Testing Service (retired)
- Sean P. (Jack) Buckley, Roblox
- Brian Gong, Center for Assessment
- Andrew D. Ho, Harvard Graduate School of Education
- Stephen Lazer, Questar Assessment
- Susan M. Lottridge, Cambium Assessment, Inc.
- Richard M. Luecht, University of North Carolina at Greensboro
- Rochelle S. Michel, Curriculum Associates
- Scott Norton, Council of Chief State School Officers
- John Whitmer, Federation of American Scientists

Strategic Vision 2025 Update August 2021

Since October 2020, the National Assessment Governing Board has been engaged in designing and implementing its legislatively-mandated body of work under the auspices of <u>Strategic Vision 2025</u>. Managed at the staff level, implementation of the vision is overseen by the Executive Committee. Staff provide quarterly updates in March, May, and August and will produce a comprehensive annual progress report every November.¹

Strategic Vision 2025 is organized by three pillars: *Inform, Innovate, and Engage*. Housed under the three pillars are eight strategic priorities. One of the underlying functions of the strategic vision is to bridge work across committees. Staff continue down the path of creating committee-level work plans that identify collaborative opportunities. Current staff-developed work plans focus on activities through September 30, 2022.

Since the May 2021 Quarterly Board Meeting, a number of activities have taken place that address strategic priorities led by the Executive Committee; Assessment Development Committee (ADC); the Committee on Standards, Design and Methodology (COSDAM); and the Reporting and Dissemination Committee (R&D). Below is a brief summary of those activities by committee.

Executive Committee

*INNOVATE: Monitor and make decisions about the NAEP assessment schedule based on the Board's policy priorities of utility, frequency, and efficiency to ensure NAEP results are policy relevant. (SV 5)*²

The Executive Committee is responsible for oversight of the Strategic Vision, on behalf of the Governing Board, and leads activities associated with the NAEP Assessment Schedule (SV 5).

During the March and May 2021 Board meetings, NCES provided funding flow projections to the Governing Board. As follow up, Executive Committee leadership sent a letter to Secretary Cardona, cc'ing Members of Congress, in support of administering the full assessment schedule which aligns with the Board's strategic vision goal of utility by advocating that the assessment schedule, as adopted, is crucial to understanding what America's students know and can do in the various subjects. In addition, this effort supported the Board's strategic priority of frequency by advocating that assessments should not be cut from the schedule and should be administered as frequently as the budget allows.

More recently, Board staff held meetings with U.S. Department of Education staff in the Secretary's office and Office of Planning, Evaluation, and Policy Development, to express support for an increase to the NAEP budget to maintain the assessment schedule. Ultimately, President Biden's FY 2022 Budget

¹ Initially, staff expected to provide committee-level reports in March, May, and August. Upon further consideration, however, separate reports do not reflect the cross-cutting nature of the strategic priorities.

 $^{^{2}}$ To avoid the perception that the priorities are rank ordered, they are not numbered in Strategic Vision 2025. However, for the purposes of working documents, numbers are used for ease and clarity.

Request for the U.S. Department of Education included a \$15 million increase to the NAEP program that would cover most costs associated with the projected budget deficit. In addition, the House Labor-HHS-Education 2022 appropriations bill included a \$40 million increase which is \$25 million more than President Biden's budget request. The additional \$25 million would be reserved to conduct a state-level Civics assessment in 2024.

Assessment Development Committee (ADC)

INNOVATE: Optimize the utility, relevance, and timing of NAEP subject-area frameworks and assessment updates to measure expectations valued by the public. (SV 4)

ADC is charged with leading the priority focused on NAEP subject-area frameworks and assessments (SV 4). Chief among ADC's responsibilities since May has been to shepherd the 2026 NAEP Reading Framework toward full Board action in August 2021. Alongside that effort and due to how the Reading Framework development process has played out over the last 18 months, the ADC has also initiated plans for reviewing and revising the framework development policy and procedures.

The 2026 NAEP Reading Framework activities include:

- Development of a Chair's Draft, led by a Working Group composed of eight Board members including the Chair and Vice Chair;
- Multiple reviews of drafts by ADC;
- Stakeholder outreach on the Chair's Draft, including a call for written feedback and briefings with key stakeholders; and
- Ongoing communication with the Visioning and Development Panels charged with developing the Reading Framework recommendations to the Board.

While in the nascent stage, ADC has begun to explore improvements that can be made to the framework update process. The Board commissioned two white papers to inform the design of future framework processes. ADC and COSDAM will be meeting in September to discuss lessons learned in the two most recent framework updates. Over the coming months, ADC will flesh out recommendations that will be shared with the full Board for their input and feedback. With a potential update to the Science Framework on the horizon, the process improvement conversations are time sensitive and critical.

Committee on Standards, Design, and Methodology (COSDAM)

INFORM: Link NAEP resources with external data sources and disseminate what is learned from the sources so that NAEP can inform policy and practice in understandable and actionable ways. (SV 3)

INNOVATE: Develop a body of evidence to improve the interpretation and communication of NAEP achievement levels to ensure that they are reasonable, valid, and informative to the public. (SV 6)

COSDAM leads two priorities. The first focuses on linking NAEP to external data sources (SV 3) and the other on developing a body of evidence to improve interpretation and communication of NAEP achievement levels (SV 6).

The Board is currently recruiting for an Assistant Director of Psychometrics; Sharyn Rosenberg has transitioned to a new role as Assistant Director for Assessment Development, working with ADC. When this role is filled, the staff will set up a cross-committee Working Group of members from COSDAM and R&D to identify policy-relevant findings from existing linking studies and discuss how this work can be highlighted in ways that are actionable to policymakers.

In September 2020, the Board awarded a contract to Pearson to review and revise the NAEP Reading and Mathematics achievement level descriptions. This project is intended to provide validity evidence to address the most important recommendation from the recent evaluation of NAEP achievement levels. Recent activities for this project include:

- In-person panel meetings for a pilot study and operational meeting are being planned in Atlanta, Georgia, for October 25-28, 2021, and February 22-25, 2022, respectively
- Recruitment for teachers and non-teacher educators to participate in these meetings is currently underway
- Materials and presentations for the panel meetings are in the process of being developed and reviewed by the Technical Advisory Committee

A procurement is planned to conduct additional activities to address the remaining items in the Board's Achievement Levels Work Plan.

Reporting and Dissemination Committee (R&D)

INFORM: Identify the needs of stakeholders and refine resources to promote sustained use of NAEP data, enabling educators, researchers, advocates, and policymakers to understand and improve student achievement. *(SV 1)*

INFORM: Elevate high-quality uses of NAEP resources to demonstrate NAEP's utility and to highlight the unique value of the Nation's Report Card to inform education policy and practice. (SV 2)

ENGAGE: Develop, sustain, and deepen strategic partnerships to ensure that NAEP remains a trusted, relevant, and useful resource. (SV 7)

ENGAGE: Help stakeholders understand how the Governing Board and NAEP can illuminate important skills for postsecondary education pathways. (SV 8)

In May, R&D prioritized its strategic vision activities for the coming months. Since that time, a number of activities have taken place, which include:
- In May, the Governing Board hosted the release of the 2019 NAEP Science results. Nearly 600 attendees joined to hear Acting Commissioner Peggy Carr's presentation of the findings and a policy-focused conversation with Board member Christine Cunningham and Board alumnus, Cary Sneider. The release event featured stakeholders in the science education community, building a network of NAEP-savvy experts within policy and advocacy groups, such as the National Science Teachers Association (NSTA), National Association of Elementary School Principals (NAESP), and the National Association for Research in Science Teaching (NARST).
- In June, the Governing Board partnered with the questioners from the National Science Teachers Association featured at the May release event and Stephen Pruitt of the Southern Regional Education Board (SREB) for a popular Twitter chat. Activities like these solidify strong working partnerships with stakeholders in NAEP Science.
- To accompany the release event and the Twitter chat, the Board produced videos about the subscales measured by the NAEP science assessment and graphics highlighting results. These are circulating on our social media channels.
- Michael Solem of the American Association of Geographers (AAG) tapped the robust collaborative relationship staff developed with him two years ago during conversations about assessment schedule changes. AAG is hosting a symposium on how to analyze NAEP Civics, NAEP U.S. History, and NAEP Geography data. The sessions will also spotlight solid research with these data as exemplars for new researchers to follow. Board staff and NCES staff will present at the symposium and have assisted in its coordination.
- The R&D Committee is reviewing the draft release plans for the Long-Term Trend results and the High School Transcript Study, both of which will be released this autumn.
- At the May meeting of the R&D Committee, members discussed approaches to improve the measure of socioeconomic status in NAEP. This conversation will continue at the August meeting to determine next steps.

Assessment Development Committee

August 5, 2021

9:00 – 10:00 am ET

Salon I and II, 5th level

COLERVING BOR

Zoom: <u>https://www.nagb.gov/governing-board/quarterly-board-meetings/2021/adc-08-2021-pt2.html</u>

AGENDA

9:00 – 9:10 am	Welcome and Review of Agenda				
	Dana Boyd, Chair Mark Miller, Vice Chair				
9:10 – 9:30 am	ACTION: 2026 NAEP Reading Framework	See plenary tab			
	Dana Boyd Mark Miller				
9:30 – 10:00 am	Upcoming ADC Activities and Priorities	Attachment A			
	Dana Boyd Mark Miller Sharyn Rosenberg, Assistant Director for Assessment Development				
Information Items	Item Review Schedule	Attachment B			
	Strategic Vision 2025 Update	See Executive Committee Material			

Upcoming ADC Activities and Priorities

Over the past year, the Assessment Development Committee (ADC) has devoted significant time to discussions of the NAEP Reading Framework update. With Board action on the NAEP Reading Framework planned for the August Board meeting, there is an opportunity for the ADC to focus on other activities and priorities.

This session will provide an opportunity for ADC members to briefly discuss the following activities that are planned for the upcoming year and to reflect on additional priorities that may be undertaken:

- Review of Assessment and Item Specifications for the NAEP Reading Framework
- Review and revision of framework development processes
- Implementation of Strategic Vision activities
- Update of NAEP Science Framework
- Review of NAEP cognitive items and contextual variables (see Attachment B)
- Creation of framework procedures manual
- Review and revision of Board policy on item development and review



550 12th Street SW Washington, DC 20202 Office: 202 123 4567

Assessment Development Committee Item Review Schedule August – December 2021 Updated July 16, 2021

Review Package to Board	Board Comments to NCES	Survey/ Cognitive Review Task		Approx. Number Items	Status
November 2, 2021	November 30, 2021	Survey	Mathematics (4, 8) Existing Item Pool Review	200	
November 2, 2021	November 30, 2021	Survey	urvey Reading (4, 8) Existing Item Pool Review		
November 24, 2021 (Off-cycle)	December 21, 2021 (Off-cycle)	Cognitive	2026 Reading (4, 8) Passage Clearance & Concept Sketch Review	40-60 passages 20-20 concept sketches	

Committee on Standards, Design and Methodology



August 3, 2021 1:00 – 3:00 pm ET (Virtual)

AGENDA

1:00 – 1:40 pm	Item Difficulty and Student Ability Distributions (CLOSED) Gregory Cizek, Chair Enis Dogan, National Center for Education Statistics	Attachment A
1:45 – 2:35 pm	Improving Information about Students Scoring Below the NAEP Basic Achievement Level Gregory Cizek Karla Egan, EdMetric Jing Chen, National Center for Education Statistics Taslima Rahman, National Center for Education Statistics	Attachment B
2:35 – 2:50 pm	Update: Review and Revision of Mathematics and Reading Achievement Level Descriptions <i>Eric Moyer, Pearson</i>	Attachment C
2:50 – 3:00 pm	Next Steps Gregory Cizek	
Information Item	Strategic Vision 2025 Update	See Executive Committee Material



Item Difficulty and Student Ability Distributions in NAEP

Would you include the following item in a fourth-grade assessment?

1+1=...How about this one? Solve for x, where $\log_x 81 = 4$

Obviously, the answer is no in both cases. Setting aside the fact that these items would not be measuring skills in a fourth-grade assessment framework, the items would not provide any "information" about a (typical) fourth-grader's mathematics "ability." There is not a good alignment between the student ability and these two (hypothetical) items; you already know how the student would perform on these items. This example is to illustrate that items should not be too difficult, nor too easy for the students—they need to be … "just right"! In fact, the level of "information" an assessment provides is proportional to the degree of alignment between student ability and item difficulty. The most efficient way to achieve such alignment is through adaptive testing, where items are selected for the student in a way that their difficulty match his/her "ability."

But what to do in a linear test, where items are not selected this way? The solution is less efficient, yet quite straightforward—include items in your assessment that are "just right" for all kinds of students, ranging from poor performing to highest performing. In other words, make sure the item pool varies in difficulty as much as the students vary in "ability" and that it includes items "just right" for every student. If, say, 20% of the students (typically) perform below a certain score, maybe allocate 20% of the item pool to items that are "just right" for this score range.

In this closed session, we will look at the degree of alignment between student "ability" and difficulty of the items in NAEP assessments. Specifically, we will look at what percent of our students perform below the *NAEP Basic* range and compare that to the percentage of items that are "just right" in this score range across a number of grades and subjects. We will do this by examining item/person maps that show student score distributions and the distributions of difficulty of items side-by-side. We will also touch on past and present efforts in improving the alignment between student "ability" and difficulty of the items in NAEP assessments.

Improving Information about Students Scoring Below the NAEP Basic Achievement Level

One of the Governing Board's most important legislated responsibilities is developing the NAEP achievement levels. The Board policy on <u>Developing Student Achievement Levels for NAEP</u> defines three achievement levels: *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*. The policy specifies that, "The remaining region that falls below the *NAEP Basic* cut score shall be identified as 'below *NAEP Basic*' when a descriptor is necessary." The percentage of students in this lowest category is reported but there is no accompanying achievement level description (ALD).

During recent COSDAM meetings, there has been discussion about the importance of better understanding what the lowest performing students know and can do. Some COSDAM members requested additional discussion about potential options for improving measurement and reporting below the *NAEP Basic* achievement level.

Three background materials are included to inform this August COSDAM discussion:

- 1. A literature review conducted by Karla Egan of EdMetric that Board staff commissioned at the request of COSDAM Chair Gregory Cizek, to better understand:
 - the number of state assessments that have Below Basic ALDs;
 - the nature of Below Basic ALDs and how they differ from other categories;
 - the pros and cons of including Below Basic ALDs in state assessments; and
 - additional considerations relevant to NAEP;
- 2. A summary report from a National Center for Education Statistics (NCES) panel convened on December 16-17, 2020, to explore how NAEP can better measure, describe, and report the skills and knowledge of lower-performing students, particularly those below *NAEP Basic;* and
- 3. A highlights report from a recent NCES special study that was conducted to better understand the knowledge and skills possessed by students who performed below the *NAEP Basic* achievement level on the grade 4 reading assessment. Additional information about the study, along with supplementary materials including audio recordings, can be found on the study website at: https://nces.ed.gov/nationsreportcard/studies/orf/.

During the August COSDAM meeting, there will be very brief presentations (approximately 5 minutes each) on the background materials listed above, followed by Committee discussion.



2021 No. 025

Describing the Lowest Achievement Level Final Report

PreparedNational Assessment Governing Boardfor:800 North Capitol Street N.W., Suite 825Washington DC 20002

Authors: Karla Egan, Ph.D. EdMetric LLC Prepared ED-NAG-17-C-0002 under:

Date: March 31, 2021

Headquarters: 66 Canal Center Plaza, Suite 700, Alexendria, VA 22314 | Phone: 703.549.3611 | humrro.org



Describing the Lowest Achievement Level

Table of Contents

Defining Achievement Level Descriptors	1
The Lowest Achievement Level	4
What do Other Testing Entities Do?	4
States: Reporting and Policy ALDs	5
States: Range ALDs	19
TIMSS and PISA	19
Other Research	20
Demographic Composition of the Lowest Achievement Level	20
Discussion	22
References	23

List of Tables

Table 1. Types and Uses of ALDs: Proposed Revised Governing Board Policy and Procedures	
(Governing Board Achievement Levels Procedures Manual, June 2020)	2
Table 2. Grade 8 Mathematics NAEP Achievement Level Descriptors	. 3
Table 3. Numbers of States Reporting at the Lowest Performance Category	. 5
Table 4. Descriptors for Lowest Achievement Level	. 6

List of Figures

Figure 1. Percentage of Students in Below the NAEP Basic achievement level compared	
to NAEP Basic and Above for Black, Hispanic, and White Students, NAEP	
Grade 8 Mathematics	21



Describing the Lowest Achievement Level

The use of achievement level descriptors (ALDs) is an accepted part of K-12 assessment. ALDs are used at standard setting, written for reports, developed for teachers to guide learning and instruction, and used by content experts to write items for large-scale assessments. Even though various aspects of ALDs have been explored in literature, one aspect of ALDs remains unexamined-the description of the lowest achievement level. Testing experts disagree on whether descriptors should be written for the lowest achievement level. This paper examines the utility and appropriateness of writing a descriptor for the lowest achievement level. In 1995, the Governing Board's policy on NAEP achievement levels said subject-matter ALDs are "articulated in terms of what students should know and should be able to do" (1995, p. 8) and "they are not written for content below the *Basic* level" (1995, p. 8). In 2018, the Governing Board reaffirmed this approach when the policy was updated to say, "There shall be no content ALDs developed for performance below the NAEP Basic level" (2018, p. 6). Therefore, since the inception of ALDs in the 1990s, the Governing Board only provides ALDs for Basic, Proficient, and Advanced. This paper first defines ALDs, then examines why ALDs are not written for the lowest level and examines the current state of the field regarding descriptors for the lowest achievement level.

Defining Achievement Level Descriptors

Achievement level descriptors (sometimes called performance level descriptors) are now ubiquitous in K–12 assessment programs. These descriptors define the types of knowledge, skills, and abilities of students at different levels of performance. The specificity of the ALD depends on the use of the ALD. For this reason, the Governing Board adopted *content ALDs*, an umbrella term that encompasses framework ALDs, threshold ALDs, and reporting ALDs. Table 1 illustrates this relationship along with the uses of each type of ALD.

The current paper focuses on policy ALDs and reporting ALDs. In a typical state program, policy ALDs are high-level definitions of the types of performance expected in each achievement level. For NAEP, the policy ALDs include:

- **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.
- Below NAEP Basic. No descriptor.



Table 1. Types and Uses of ALDs: Proposed Revised Governing Board Policy andProcedures (Governing Board Achievement Levels Procedures Manual, June 2020)

Policy definitions: The policy defines three NAEP achievement levels: NAEP Basic, NAEP Proficient, and NAEP Advanced. These policy definitions apply to all main NAEP assessments.

Content ALDs	
ALDs in Framework (for item development and achievement- level setting)	Under the revised policy and procedures for framework development, the framework development panel may develop multiple sets of content ALDs for the purposes of informing item development and for use in the achievement-level setting activities. The framework development panel might also determine that one set of ALDs can serve both of these purposes. These ALDs will continue to be written in terms of what students should know and be able to do. If there is a specific need to revise the content ALDs in advance of an achievement-level setting, then a separate activity will be undertaken to do so, but this is not intended to be necessary in most cases.
Threshold/Borderline ALDs (if applicable)	If descriptions of performance right at the cut scores are needed for setting achievement levels (e.g., if a Bookmark or similar procedure is used), then threshold (or borderline) ALDs will be developed by achievement-level setting panelists. Threshold ALDs are for the panelists' own use and are not reported with the NAEP results. The rationale for having the achievement-level setting panelists create threshold ALDs (rather than providing them at the beginning of the process) is that it is an important task to help panelists fully internalize the ALDs. Because the creation of threshold ALDs is an instrumental activity that occurs as part of the achievement-level setting process, panelists are typically discouraged from spending inordinate amounts of time on their development or focusing on minor edits and wordsmithing.
Reporting ALDs	Reporting ALDs are developed following the first operational administration of an assessment and express the empirical findings as to what students have demonstrated they know or can do at each achievement level. The policy calls for conducting a study to derive the reporting ALDs following the first operational administration of an assessment (and again every 3 administrations or 10 years, whichever comes later).

The reporting ALDs describe the knowledge, skills, and abilities that students in each achievement level (e.g., NAEP Basic, NAEP Proficient, NAEP Advanced) demonstrate. Table 2 provides the current reporting ALDs for Grade 8 mathematics NAEP; however, the ALDs in Table 2 are currently being updated to reflect the new requirement in the NAEP policy that reporting ALDs incorporate empirical data on student performance and describe what students *do* know and *can* do rather on what they should know and be able to do.



Achievement Level	Description
NAEP Advanced (333)	Eighth-grade students performing at the NAEP Advanced level should be able to reach beyond the recognition, identification, and application of mathematical rules in order to generalize and synthesize concepts and principles in the five NAEP content areas. Eighth-graders performing at the NAEP Advanced level should be able to probe examples and counterexamples in order to shape generalizations from which they can develop models. Eighth-graders performing at the NAEP Advanced level should use number sense and geometric awareness to consider the reasonableness of an answer. They are expected to use abstract thinking to create unique problem-solving techniques and explain the reasoning processes underlying their conclusions.
NAEP Proficient (299)	Eighth-grade students performing at the <i>NAEP Proficient</i> level should apply mathematical concepts and procedures consistently to complex problems in the five NAEP content areas.
	Eighth-graders performing at the <i>NAEP Proficient</i> level should be able to conjecture, defend their ideas, and give supporting examples. They should understand the connections between fractions, percents, decimals, and other mathematical topics such as algebra and functions. Students at this level are expected to have a thorough understanding of <i>NAEP Basic</i> level arithmetic operations—an understanding sufficient for problem solving in practical situations.
	Quantity and spatial relationships in problem solving and reasoning should be familiar to them, and they should be able to convey underlying reasoning skills beyond the level of arithmetic. They should be able to compare and contrast mathematical ideas and generate their own examples. These students should make inferences from data and graphs, apply properties of informal geometry, and accurately use the tools of technology. Students at this level should understand the process of gathering and organizing data and be able to calculate, evaluate, and communicate results within the domain of statistics and probability.
NAEP Basic (262)	Eighth-grade students performing at the <i>NAEP Basic</i> level should exhibit evidence of conceptual and procedural understanding in the five NAEP content areas. This level of performance signifies an understanding of arithmetic operations—including estimation—on whole numbers, decimals, fractions, and percents.
	Eighth-graders performing at the <i>NAEP Basic</i> level should complete problems correctly with the help of structural prompts such as diagrams, charts, and graphs. They should be able to solve problems in all NAEP content areas through the appropriate selection and use of strategies and technological tools—including calculators, computers, and geometric shapes. Students at this level also should be able to use fundamental algebraic and informal geometric concepts in problem solving.
	As they approach the <i>NAEP Proficient</i> level, students at the <i>NAEP Basic</i> level should be able to determine which of the available data are necessary and sufficient for correct solutions and use them in problem solving. However, these eighth-graders show limited skill in communicating mathematically.
Below NAEP Basic	Not Described

Table 2. Grade 8 Mathematics NAEP Achievement Level Descriptors



For a typical state summative assessment, ALDs are found on individual score reports to provide meaning to the scale score the student achieved. Individual results, however, are not reported for the NAEP assessments. There is a rigorous psychometric process being used to create NAEP reporting ALDs where scale scores are extrapolated for individual students, those scale scores are assigned to achievement levels, and NAEP items are linked to the extrapolated scale scores. The group of items assigned to each achievement level represents the knowledge, skills, and abilities (KSAs) an "average student" in each level can demonstrate. Even though it is a misnomer to discuss student performance in terms of NAEP, we use this shorthand when describing the KSAs included in the reporting ALDs or the high-level skills found in policy ALDs.

The Lowest Achievement Level

The Governing Board's 1995 policy and updated 2018 policy to *not* describe the lowest level rests on reasonable psychometric principles. First, NAEP assesses the knowledge and skills of students relative to the NAEP frameworks. These frameworks reflect current educational requirements in the United States. Students in the lowest category may demonstrate knowledge of some entry skills or lower-ability skills, but they are not yet able to demonstrate the bulk of the knowledge and skills measured by the framework. NAEP results are not intended to drive instruction; rather, the NAEP provides a snapshot of student performance in the United States on the subject area tests, and the reporting ALDs provide a snapshot of the framework KSAs found in the NAEP Basic, NAEP Proficient, and NAEP Advanced areas of the scale. To capture the KSAs of the lowest achievement category, the NAEP frameworks may need to be expanded to consider pre-requisite skills the students in the lowest category can demonstrate.

Without expanding the item pool, the conditional standard error of measurement is quite large in the area of the test scale below the NAEP Basic range. There are relatively few items that cover this area of the scale in most NAEP assessments, so it is difficult to identify any KSAs with the same precision as there is in the other performance categories. Currently, the Nation's Report Card website provides sample items and item maps to indicate types of KSAs found in the lowest achievement category. Even with an expanded item pool, it will be difficult to encapsulate the KSAs of the diverse student performance found in the lowest achievement category, below the NAEP Basic level. This area of the scale ranges from the lowest obtainable scale score to the scale score just before the NAEP Basic cut score. Student performance at the lowest obtainable scale score by default or because students have performed poorly on the test. We cannot say anything about the KSAs associated with the lowest obtainable scale score; however, much could be summarized regarding the KSAs of the students just below the NAEP Basic cut score.

What do Other Testing Entities Do?

Even though it is difficult, some testing organizations describe student performance in the lowest category. To understand the current state of the field, we gathered information from:

- State websites
- TIMSS and PIRLS
- Academic literature

We examined all 50 states' websites and Washington, DC. Within the state websites, we searched technical reports, score interpretation guides, and standard-setting information for



evidence regarding how each state addressed the lowest achievement level for grade 4 English language arts (ELA) and grade 8 mathematics. This information was not easily located for most states. In some cases, links were broken or outdated. In other cases, information was located for one grade/content area but not another. In Utah, we could only locate the reporting ALDs for grade 6 math. For Alabama, we found the grade 8 math descriptors but not the grade 4 ELA descriptors.

We focused on collecting information related to policy and/or reporting ALDs, as these were the ALDs most often reported to stakeholders. The following states are not included in the list or counts of state ALDs:

- Arkansas, Kentucky, Maryland: Information could not be located
- New Jersey: 2008 descriptors were located and determined to be out of date
- Arizona: an example ALD was located in a guide for score reporting on a graphic of an individual student report but a full list could not be located

For the most part, state departments of education followed the same patterns across their ALDs—if they reported a descriptor for the lowest category in mathematics, then they also reported it in ELA.

States: Reporting and Policy ALDs

Table 3 shows the number of states where policy or reporting ALDs could be located. It shows we located information for 46 states and could not locate information for five states. Table 3 shows that 43 of the 46 states reported something at the lowest category. The states belonging to PARCC (Colorado, New Mexico, and Washington, DC) do not report at the lowest performance category. Of the 43 states reporting at the lowest level, we located policy descriptors for 37 states and reporting descriptors for eight states (Alaska, Maine, Minnesota, Ohio, Pennsylvania, Texas, Utah, and Virginia).

ALD Status	Total Number of States	Number of States Reporting at Lowest AL	
Located	46	43	
Not Located	5		

Table 2	Numbors (of Statos	Poporting	at the	Lowost	Porformanco	Catogory
I able 5.	Number 5 (JI States	Reporting	allie	LOwest	renomance	Calegory

Table 4 lists the ALDs for grade 8 mathematics and grade 4 ELA. States that belong to the Smarter Balanced Assessment Consortium are grouped together. The policy ALDs are listed first, followed by the reporting ALDs. The policy descriptors appear to split between the use of negative language or positive language. Slightly more than half of the states with policy descriptors used negative language to describe student performance in this category. In other words, the descriptors stated the student "has not met" the achievement standard. For example, see Illinois and Indiana descriptors in Table 4. Slightly fewer than half use positive language to come to a similar conclusion. These states tend to use the phrase: "demonstrates minimal understanding." Many states also assert the students in the lowest achievement level will need academic support. For example, see New York or North Dakota descriptors for an example of positive language.



Table 4. Descriptors for Lowest Achievement Level

State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link			
Policy							
Smarter Balanced: California, Hawaii, Idaho, Michigan, Oregon, Nevada, Montana, South Dakota	Standard Not Met	The student has not met the achievement standard and needs substantial improvement to demonstrate the knowledge and skills in mathematics needed for likely success in future coursework.	The student has not met the achievement standard and needs substantial improvement to demonstrate the knowledge and skills in English language arts/literacy needed for likely success in future coursework.	https://www.cde.ca.gov/ta/tg/ca/sbac hievedescript.asp			
Smarter Balanced: Connecticut	Does Not Meet the Achievement Standard	The student has not yet met the achievement standard for mathematics expected for this grade. Students performing at this standard require substantial improvement toward mastery of mathematics knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in high school and college coursework or career training.	The student has not yet met the achievement standard for English language arts and literacy expected for this grade. Students performing at this standard require substantial improvement toward mastery of English language arts and literacy knowledge and skills. Students performing at this standard will likely need substantial support to get on track for success in the next grade.	https://portal.ct.gov/- /media/SDE/Student- Assessment/Smarter-Results- Resources/Interpretive_Guide_8_23 _17FINAL.pdf?la=en			
Smarter Balanced: Delaware	Minimal Understanding	The Level 1 student demonstrates minimal understanding of and ability to apply the English language arts and literacy (mathematics) knowledge and skills needed for success in college and career, as specified in the Common Core State Standards.	The Level 1 student demonstrates minimal understanding of and ability to apply the English language arts and literacy (mathematics) knowledge and skills needed for success in college and career, as specified in the Common Core State Standards.	https://www.doe.k12.de.us/cms/lib/D E01922744/Centricity/Domain/111/D eSSA%20Executive%20State%20S ummary%202017.pdf			



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Smarter Balanced: Vermont, Washington	Level 1	Student demonstrates minimal understanding of and ability to apply the knowledge and skills associated with college content readiness.	Student demonstrates minimal understanding of and ability to apply the knowledge and skills associated with college content readiness.	https://vt.portal.cambiumast.com/resources/test-blueprints/
Alabama	Level 1	The student has a minimal understanding of grade-level standards and is likely to need additional support at this level of learning as described in the Alabama Course of Study.		https://www.alsde.edu/sec/sa/Pages /relatedinfo- all.aspx?navtext=Resources
Florida	Inadequate	Highly likely to need substantial support for the next grade	Highly likely to need substantial support for the next grade	http://www.fldoe.org/core/fileparse.p hp/5663/urlt/Grade-LevelFS2021.pdf
Georgia	Beginning Learners	do not yet demonstrate proficiency in the knowledge and skills necessary at this grade level/course of learning, as specified in Georgia's content standards. The students need substantial academic support to be prepared for the next grade level or course and to be on track for college and career readiness.	do not yet demonstrate proficiency in the knowledge and skills necessary at this grade level/course of learning, as specified in Georgia's content standards. The students need substantial academic support to be prepared for the next grade level or course and to be on track for college and career readiness.	https://www.gadoe.org/Curriculum- Instruction-and- Assessment/Assessment/Document s/Milestones/ALD/ALDS_for_Grade 8 Milestones EOG Mathematics. pdf
Illinois	Did not yet meet expectations	Students performing at this level do not yet meet academic expectations for the knowledge, skills, and practices contained in the standards for ELA/L or mathematics assessed at their grade level. They will need academic support to engage successfully in further studies in this content area.	Students performing at this level do not yet meet academic expectations for the knowledge, skills, and practices contained in the standards for ELA/L or mathematics assessed at their grade level. They will need academic support to engage successfully in further studies in this content area.	https://www.isbe.net/Documents/Ne w-Meridian-Tech-Rpt-2019.pdf



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Indiana	Below Proficiency	Indiana students below proficiency have not met current grade level standards. Students may require significant support to develop the knowledge, application, and analytical skills needed to be on track for college and career readiness.	Indiana students below proficiency have not met current grade level standards. Students may require significant support to develop the knowledge, application, and analytical skills needed to be on track for college and career readiness.	https://www.doe.in.gov/assessment/i learn
Iowa	Not-Yet Proficient	Students performing at the not-yet- proficient level have not yet demonstrated the knowledge and skills to be classified as Proficient.	Students performing at the not- yet-proficient level have not yet demonstrated the knowledge and skills to be classified as Proficient.	https://itp.education.uiowa.edu/ia/do cuments/Research-Guide-Form-E- F.pdf
Kansas	Level 1	A student at Level 1 shows a limited ability to understand and use the skills and knowledge needed for post-secondary readiness.	A student at Level 1 shows a limited ability to understand and use the skills and knowledge needed for post-secondary readiness.	https://ksassessments.org/resources -and-training
Louisiana	Unsatisfactory	Students performing at this level have not yet met the college and career readiness expectations and will need extensive support to be prepared for the next level of studies in this content area.	Students performing at this level have not yet met the college and career readiness expectations and will need extensive support to be prepared for the next level of studies in this content area.	https://www.louisianabelieves.com/r esources/library/assessment
Massachusetts	Not Meeting Expectations	A student who performed at this level did not meet grade-level expectations in this subject. The school, in consultation with the student's parent/guardian, should determine the coordinated academic assistance and/or additional instruction the student needs to succeed in this subject.	A student who performed at this level did not meet grade-level expectations in this subject. The school, in consultation with the student's parent/guardian, should determine the coordinated academic assistance and/or additional instruction the student needs to succeed in this subject.	https://www.doe.mass.edu/mcas/tec h/



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Mississippi	Minimal Understanding	Students performing below the Basic level inconsistently demonstrate the knowledge or skills that define basic level performance.	Students performing below the Basic level inconsistently demonstrate the knowledge or skills that define basic level performance.	https://districtaccess.mde.k12.ms.us /studentassessment/Public%20Acce ss/Forms/AllItems.aspx?RootFolder =%2Fstudentassessment%2FPublic %20Access%2FStatewide_Assess ment_Programs%2FMAAP- Mississippi%20Academic%20Asses sment%20Program%2FMAAP%20R eport%20Interpretation%20Guides& FolderCTID=0x0120008C41041A90 7A304BA89A4587F88962BC&View =%7B5FB78E06-9076-48F4-9A3B- C02F433B41D0%7D
Missouri	Below Basic	Students do not yet demonstrate proficiency in the knowledge and skills necessary at this grade level/course of learning, as specified in content expectations. These students need substantial academic support to be prepared for the next grade level or course and to be on track for college and career readiness.	Students performing at the Below Basic level on the Missouri Assessment Program demonstrate a minimal command of the skills and processes identified in the Missouri Learning Standards. They demonstrate these skills inconsistently and/ or incorrectly in reading processes responding to literary and informational texts and in writing, listening, and speaking forms. Students performing at the Below Basic level use few strategies to comprehend and interpret texts, demonstrate little understanding of literary forms, and apply few strategies for accessing information. They demonstrate little or no ability to organize and/ or develop writing or exhibit little command of the conventions of standard English.	https://dese.mo.gov/sites/default/file s/asmt-gl-gir-spring-2019.pdf



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Nebraska	Developing	Developing learners do not yet demonstrate proficiency in the knowledge and skills necessary at this grade level, as specified in the assessed Nebraska College and Career Ready Standards.	With a range of texts with text complexity commonly found in Grade 4, a student performing in Developing can likely [Insert text from specific standard here]	https://www.education.ne.gov/asses sment/nscas-general-summative- assessment/nscas-mathematics/
New Hampshire	Level 1 Below Proficient	The Level 1 student is below proficient in applying mathematics knowledge/skills as specified in the standards. The student generally performs significantly below the standard for the grade level/course, is likely able to partially access grade-level content, and engages with higher order thinking skills with extensive support.	The Level 1 student is below proficient in applying English language arts knowledge/skills as specified in the standards. The student generally performs significantly below the standard for the grade level/course, is likely able to partially access grade-level content, and engages with higher order thinking skills with extensive support.	https://www.education.nh.gov/sites/g /files/ehbemt326/files/inline- documents/sonh/nhsas-v1-tech- report-2018-19.pdf
New York	NYS Level 1	Students performing at this level are well below proficient in standards for their grade. They demonstrate limited knowledge, skills, and practices embodied by the New York State P-12 Common Core Learning Standards for Mathematics that are considered insufficient for the expectations at this grade.	Students performing at this level are well below proficient in standards for their grade. They demonstrate limited knowledge, skills, and practices embodied by the New York State P-12 Common Core Learning Standards for English Language Arts/Literacy that are considered insufficient for the expectations at this grade.	https://www.engageny.org/resource/ performance-level-descriptions-for- ela-and-mathematics



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
North Carolina	Not Proficient	Students who are Not Proficient demonstrate inconsistent understanding of grade level content standards and will need support.	Students performing at this level have limited command of the knowledge and skills contained in the Common Core State Standards (CCSS) Reading Standards for Literature as assessed by referring to the text when drawing inferences, as well as when explaining what the text directly says; summarizing the text and determining the theme from details; using specific details to describe a character, setting, or event in a story; and determining the meaning of words and phrases as they are used in a text, including those words referring to mythological characters. Students will need academic support to engage successfully in this content area.	https://www.dpi.nc.gov/documents?fi eld_document_type_tid=388&field_d ocument_type_tid_op=or
North Dakota	Novice	The Level 1 student is below proficient in applying mathematics knowledge/skills as specified in the standards. The student generally performs significantly below the standard for the grade level/course, is likely able to partially access grade-level content, and engages with higher order thinking skills with extensive support	The Level 1 student is below proficient in applying English language arts knowledge/skills as specified in the standards. The student generally performs significantly below the standard for the grade level/course, is likely able to partially access grade-level content, and engages with higher order thinking skills with extensive support	https://www.nd.gov/dpi/districtsscho ols/assessment/ndsa
Oklahoma	Below Basic	Students have not performed at least at the Basic level.	Students have not performed at least at the Basic level.	https://sde.ok.gov/assessment- material



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Pennsylvania	Below Basic	The Below Basic Level reflects inadequate academic performance, and work at this level demonstrates a minimal command of and ability to apply the knowledge, skills, and practices represented in the Pennsylvania standards. Consistent performance at this level indicates extensive additional academic support may be needed for engaging successfully in further studies in this content area.	The Below Basic Level reflects inadequate academic performance, and work at this level demonstrates a minimal command of and ability to apply the knowledge, skills, and practices represented in the Pennsylvania standards. Consistent performance at this level indicates extensive additional academic support may be needed for engaging successfully in further studies in this content area.	https://www.education.pa.gov/K- 12/Assessment%20and%20Account ability/PSSA/Pages/DescriptorsCutS cores.aspx
Rhode Island	Not Meeting Expectations	A student who performed at this level did not meet grade-level expectations in this subject. The school, in consultation with the student's parent/guardian, should determine the coordinated academic assistance and/or additional instruction the student needs to succeed in this subject.	A student who performed at this level did not meet grade-level expectations in this subject. The school, in consultation with the student's parent/guardian, should determine the coordinated academic assistance and/or additional instruction the student needs to succeed in this subject.	https://www.ride.ri.gov/InstructionAs sessment/Assessment/RICASAsses sments.aspx
South Carolina	Does Not Meet Expectations	A student who does not meet expectations in the knowledge and skills necessary at this grade level of learning, as defined by the grade-level content standards, needs substantial academic support to be prepared for the next grade level and to be on track for college and career readiness.	A student who does not meet expectations in the knowledge and skills necessary at this grade level of learning, as defined by the grade-level content standards, needs substantial academic support to be prepared for the next grade level and to be on track for college and career readiness.	https://ed.sc.gov/tests/middle/sc- ready/



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Tennessee	Below	Performance at this level demonstrates that the student has a minimal understanding and has a partial ability to apply [Insert course here] knowledge and skills as defined by the Tennessee Academic Standards.	Performance at this level demonstrates that the student has a minimal understanding and has a partial ability to apply [Insert course here] knowledge and skills as defined by the Tennessee Academic Standards.	https://www.tn.gov/education/assess ment/tnready.html
Texas	Did Not Meet Grade Level	Performance in this category indicates that students are unlikely to succeed in the next grade or course without significant, ongoing academic intervention. Students in this category do not demonstrate a sufficient understanding of the assessed knowledge and skills.	Performance in this category indicates that students are unlikely to succeed in the next grade or course without significant, ongoing academic intervention. Students in this category do not demonstrate a sufficient understanding of the assessed knowledge and skills.	https://tea.texas.gov/student- assessment/testing/staar/staar- performance-level-descriptors
West Virginia	Does Not Meet Standards	The student generally demonstrates a minimal understanding of, and ability to, apply grade-level math knowledge, skills, and abilities relative to the West Virginia College- and Career-Readiness Standards for Mathematics.	The student generally demonstrates a minimal understanding of, and ability to, apply grade-level English language arts (ELA) knowledge, skills, and abilities relative to the West Virginia College- and Career- Readiness Standards for ELA.	https://wvde.us/assessment/scaled- score-information/wvgsa-in-grades- <u>3-8/</u>
Wisconsin	Below Basic	Student demonstrates minimal understanding of and ability to apply the knowledge and skills for their grade level that are associated with college content-readiness.	Student demonstrates minimal understanding of and ability to apply the knowledge and skills for their grade level that are associated with college content-readiness.	https://dpi.wi.gov/sites/default/files/i mce/assessment/pdf/WI Math Perf ormanceLevelDescriptors.pdf
Wyoming	Below Basic	Students performing at below basic level in Mathematics have minimal or no academic performance indicating understanding and little display of the knowledge and skills included in the Wyoming Content and Performance Standards.	Students performing at the below basic level in English Language Arts have minimal academic performance indicating minimal understanding and little display of the knowledge and skills included in the Wyoming Content and Performance Standards.	https://edu.wyoming.gov/educators/s tate-assessment/plds/



Table 4. Descri	iptors for Lowe	est Achievement L	_evel (Continued)

State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Reporting				
Alaska	Far Below Proficient	Students who score at this level may be able to recognize that irrational numbers are different from rational numbers, understand exponents as repeated multiplication, find the slope of a line using a graph, represent whole numbers in scientific notation, identify whether a relation is a function, recognize congruence and similarity, recognize single transformations of geometric figures, find the hypotenuse in a right triangle with sides whose lengths are whole numbers that are Pythagorean triples, and recognize associations in data that represent two quantities.	Students who score at this level attempt to read and minimally comprehend grade 4 text to identify main ideas and explicit details, determine meanings of basic words and phrases while identifying literal and figurative language, identify text features and structures used to organize a text, and identify relationships between parts of a text. When writing or revising, students attempt to use appropriate language and conventions, use strategies particular to a type of text, and structure a text to support a purpose or opinion.	https://education.alaska.gov/tls/A ssessments/Peaks/EducatorGuid e_Assessments_Reports.pdf
Maine	Well Below State Expectations	The student's work demonstrates a minimal understanding of, and ability to apply the mathematics knowledge and skills needed for achievement relative to the grade level Math Content and Practice Standards. The student solves some problems that require applying simple strategies to basic areas of mathematics without an understanding of the reasoning behind the strategies.	The student's work demonstrates a minimal understanding of the knowledge and skills needed to meet Maine's ELA/Literacy Content Standards with texts of appropriate complexity for the grade level.	https://www.maine.gov/doe/sites/ maine.gov.doe/files/inline- files/MEA_2018_ALDs%2BCut% 20Scores_10-09-18.pdf



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Minnesota	Does Not Meet the Standards	 A student at this level of mathematics succeeds at few of the most fundamental mathematics skills of the Minnesota Academic Standards. Some of the skills demonstrated may include: Number & Operation Recognizes fractions and terminating decimals as rational numbers 	 When interacting with literature and informational text, students at this achievement level demonstrate the following skills inconsistently and with minimal accuracy. Key Ideas and Details (Standards 1, 2, 3) Recall details from text Make simple predictions based on 	https://education.mn.gov/ MDE/dse/test/ald/
		 Algebra Recognizes linear functions in graphic presentations Translates linear representations from a table to a graph Identifies slope by counting whole number units on a graph Identifies patterns in a table of a linear function (e.g., recognizes patterns for x or y-values but not the relationship between x and y) Substitutes "easy" numbers and evaluates simple expressions Geometry & Measurement Recognizes parallel or perpendicular lines on a graph Data Analysis 	 Make simple predictions based on explicit text Identify a cause or an effect Identify obvious fact and opinion in explicit text Make general comparisons based on explicit text Locate explicit main idea and central message Identify basic sequence of events Craft and Structure (Standards 4, 5, 6) Recognize simple figures of speech Locate obvious context clues to understand word meanings Identify key words and phrases Recognize the features, format, and function of basic text structures (e.g., listing) and their impact on meaning State author's obvious purpose in explicit text 	
		 Generalizes the properties of the line of best fit of a graphed data set Displays data using scatterplots 	 Integration of Knowledge and ideas (Standard 8, Informational Text substrand only) Identify obvious evidence in text (e.g., logical connections between sentences and paragraphs) 	



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Ohio	Limited	A student performing at the Limited Level demonstrates a minimal command of Ohio's Learning Standards for Grade 8 Mathematics. A student at this level has an emerging ability to formulate and reason about expressions and equations, use functions to describe quantitative relationships, and analyze two- and three-dimensional space and figures using distance, angle, similarity, and congruence, and to understand and apply the Pythagorean Theorem.	A student performing at the Limited Level demonstrates a minimal command of Ohio's Learning Standards for Grade 4 English Language Arts. A student at this level has an emerging ability to determine the main idea or theme of a text and explain how it is supported by key details while providing a summary of the text, explain how an author uses evidence to support particular points in a text, and write an opinion or explanatory piece that introduces a topic, develops the topic with facts, details, and supported opinions, and links ideas with categories of information.	https://oh.portal.cambium ast.com/resources/reporting-resources/
Pennsylvania	Emerging/Below Basic	Students performing at this level identify rational numbers and locate approximate positions on a number line. They evaluate radical notation for perfect squares. Students solve one- and selected two-step equations in one variable with one solution. Students identify a function using a graph or table. They use the Pythagorean theorem or volume formulas to solve simple or routine problems. Students identify line of best fit or determine patterns of association in bivariate data.	A student performing at the below basic level demonstrates inadequate understanding of literary and informational texts. The student demonstrates minimal or no understanding of vocabulary, word meaning, and conventions of language. The student demonstrates minimal or no understanding of writing skills.	https://www.education.pa .gov/K- 12/Assessment%20and %20Accountability/PSSA /Pages/DescriptorsCutSc ores.aspx



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Texas	Did Not Meet Grade Level Performance	 Students achieving Did Not Meet Grade Level Performance can Solve problems using direct variation Solve application problems involving the Pythagorean theorem Use proportional and non-proportional relationships to develop foundational concepts of functions Approximate the value of irrational numbers 	 When reading texts of increasing complexity, students achieving Did Not Meet Grade Level Performance can Determine the meaning of unfamiliar words using explicit context Demonstrate a literal understanding of literary and informational texts Make plausible inferences about literary and informational texts 	https://tea.texas.gov/student- assessment/testing/staar/sta ar-performance-level- descriptors
Utah *Grade 6 ma	Below Proficient	Understands and represents situations with rates and ratios. Understands and uses negative numbers. Uses equations and expressions to solve problems, including equations and expressions that contain variables in place of numbers. Works with statistical data to find a measure of center, including mean, median, and mode.		https://www.schools.utah.gov /assessment/assessments? mid=1173&tid=7



State	Lowest AL Name	Grade 8 MA	Grade 4 ELA	Link
Virginia	Fail/Below Basic	 A student performing at this level should be able to <i>Reporting Category 1: Number, Number Sense, Computation, and Estimation</i> compare fractions and decimals identify natural numbers, whole numbers, and integers use manipulatives to recognize perfect squares solve practical problems involving fractions, decimals, and integers <i>Reporting Category 2: Measurement and Geometry</i> define and recognize acute, obtuse, right, and straight angles determine the areas of circles, triangles, and rectangles determine the volume/surface area of a rectangular prism given a labeled figure identify the image of a polygon resulting from a single transformation use manipulatives to describe the views (top/front/side) of a three-dimensional figure define the Pythagorean Theorem <i>Reporting Category 3: Probability, Statistics, Patterns, Functions, and Algebra</i> determine the probability of a simple event name the dependent and independent variables represented in a scatterplot apply the order of operations to numerical expressions simplify algebraic expressions 	 A student performing at this level should be able to Define context clues. Recognize that some words have prefixes and/or suffixes. Define synonym or antonym. Recall the purpose of word-reference materials. Recall literary terms. Recall plot events. Explain the characteristics of fiction and nonfiction. Define sensory words. Locate information in texts to answer literal questions. Define cause or effect. Locate text features. Recognize that an author has a purpose for writing. Define main idea or topic. Define fact or opinion. 	https://www.doe.virginia. gov/testing/scoring/perfor mance_level_descriptors /index.shtml



Each state uses a different strategy when writing reporting ALDs. Alaska uses "may be able" when describing the types of knowledge and skills students in the lowest achievement level possess, while Ohio discusses "emerging" knowledge and skills. Texas describes what students in the lowest level "can do." Pennsylvania uses action verbs to describe math skills for the students in the lowest level, while using "minimal or no understanding" to discuss reading skills.

The use of the phrases "may be able" or "can" is important when writing ALDs. The term "can" implies these are KSAs most students in the lowest category would be able to demonstrate. The term "may be able" does not carry the same meaning. "May be able" implies some students might have the KSAs while others do not.

States: Range ALDs

Range ALDs are detailed definitions of student performance that encompass the entirety of student performance in a particular achievement level. These are sometimes written at the beginning of the test development cycle and describe the types of KSAs students should be able to do. In other instances, these ALDs are based on items and describe the KSAs students can do. This ALD type was not the primary focus of our work; however, we encountered this ALD type as we searched for policy and reporting ALDs. These ALDs are developed for each content strand within a content standard. We have pulled examples from New York, Alaska, and Alabama for standards related to expressions and equations.

New York uses the following text to describe what a Level 1 student *should* do in relationship to cluster 8.EE.1,3,4 (Students work with integers): "Write simple numerical expressions involving whole number exponents and evaluate expressions with exponents of between 1 and 10."

Alaska describes "what a typical student scoring at each level can do" (ADEED, 2017). For Alaska Standards 8.EE.1-8.EE.8, the student in the lowest category (Far Below Proficient) is summarized as:

- A student at this level understands exponents as representing repeated multiplication.
- A student at this level finds the slope of a line using a graph.
- A student at this level represents whole-number multiples of ten in scientific notation.

Alabama also describes "what a typical student in each performance level can do... A student would not necessarily demonstrate all the skills listed a particular performance level." For Alabama Standards 8.EE.3-8.EE.10, a student in the lowest performance category is summarized as:

- estimates very large or very small quantities as a single digit times an integer power of 10,
- interprets the unit rate as the slope of a proportional relationship represented in a graph, and
- finds the slope of a line using a graph.

TIMSS and PISA

Achievement level descriptors are also used for the TIMSS and PISA assessments. Like NAEP, these assessments only produce scores at the group level and not for individual students. Unlike NAEP, neither assessment system uses a judgmental standard setting; instead, using a "pragmatic and empirically-based approach" (Olson & Nilson, 2017) normative cut scores are established. Both assessments use item maps to create item-level descriptors that are aggregated into ALDs. PISA splits the lowest level into two performance categories: 1a and 1b. The summary descriptors for the lowest performance levels for scientific literacy read:



At **Level 1a**, students are able to use every day content and procedural knowledge to recognise or identify explanations of simple scientific phenomenon. With support, they can undertake structured scientific enquiries with no more than two variables. They are able to identify simple causal or correlational relationships and interpret graphical and visual data that require a low level of cognitive demand. Level 1a students can select the best scientific explanation for given data in familiar personal, local and global contexts.

At **Level 1b**, students can use every day content knowledge to recognise aspects of simple scientific phenomenon. They are able to identify simple patterns in data, recognise basic scientific terms and follow explicit instructions to carry out a scientific procedure.

PISA uses "can" statements to describe the knowledge and skills demonstrated by students in the lowest proficiency level.

TIMSS describes performance at four benchmarks, including the lowest benchmark. For the 8th grade mathematics "Low Benchmark," the descriptor reads, "Have some knowledge of whole numbers and basic graphs." Consequently, TIMSS is describing, at a very high level, the skills possessed by students classified as "Low".

Other Research

Outside of psychometrics, other researchers have explored which students are in the lowest performance level. For example, Valencia and Buly (2004) examined 108 grade 5 students who scored in the lowest performance level of a grade 4 statewide reading assessment in a northwestern United States school district. The students were administered assessments that targeted expert-identified aspects of reading, including word identification, meaning (comprehension and vocabulary), and fluency (rate and expression). All students were measured individually on a battery of tests. Word identification was measured using the 1989 Woodcock-Johnson Psycho-Educational Battery-Revised and the 1995 Qualitative Reading Inventory-II (QRI-II). Comprehension was measured with the QRI-II and the 1981 Peabody Picture Vocabulary Test-Revised. Fluency was evaluated by measuring the reading rate of all passages from the QRI-II and state test. They also scored reading expression using the NAEP Oral Study rubric. All data were analyzed through a cluster analysis to look for similar patterns on word identification, meaning, and fluency.

The researchers created six profiles of prototypical students who fall in the lowest achievement level. These profiles include: Automatic Word Callers, Struggling Word Callers, Word Stumblers, Slow Comprehenders, Slow Word Callers, and Disabled Readers. For each prototypical student, they identified level of word identification, comprehension, vocabulary, expression, reading rate, and writing.

Demographic Composition of the Lowest Achievement Level

The main purpose of this paper is to explore how the lowest achievement level is described (or not described) by entities outside of the Governing Board. Even so, it is important to understand which student groups are classified in the lowest category of achievement on NAEP, for it is their performance that goes undescribed. Here, we examine 2019 NAEP Grade 8 Mathematics data for patterns of performance. Similar patterns are seen in other grade/content area combinations.

In 2019, the assessment performance of approximately 31% of all students was classified as below the NAEP Basic achievement level. When disaggregated by racial group, the 2019



assessment performance of 53% of Black students¹, 43% of Hispanic students, 49% of American Indian/Alaska Native, 14% of Asian/Pacific Islander students, and 20% of white students falls into the lowest category. When disaggregated by students participating in the free and reduced lunch (FRL) program (an indicator often used as a proxy for socioeconomic status), the assessment performance of 46% of FRL students falls in the lowest category while the performance of 18% of non-FRL students and 20% of unknown FRL status falls in the lowest category. In short, the test performance of students of color or students receiving FRL is more likely to fall into the lowest level than is the test performance of white students or non-FRL students. The graphics in Figure 1 show the relative stability of the percentages of Black students, Hispanic students, and white students in each category on the 2019 NAEP Grade 8 Mathematics assessment.







¹ The achievement gap between black and white students on NAEP is well-known (see USDOE, 2015).



The students comprising the lowest achievement level are disproportionately students of color or students living in poverty. Performance on NAEP, like any large-scale summative assessment, reflects societal inequities. Students of color and/or students in poverty often do not have access to the same educational resources as white, Asian, and/or wealthy students. Students' performance on NAEP may reflect systemic racism that exists in U.S. society.

Discussion

Like all well-designed assessments, the NAEP measures the content frameworks that underlie it. The NAEP is not intended to measure students' performance outside of the NAEP frameworks, and the Governing Board cannot report what is *not* measured. If the Governing Board decides to report the KSAs of the students in the lowest performance category, it may be necessary to add pre-requisite KSAs to the content frameworks to capture what students in this level can demonstrate on an assessment. This may culminate in a special test form specifically designed to capture the KSAs of the students in the lowest category. Or, NAEP could perform special studies to understand and report more about the knowledge and skills of students below the NAEP Basic level. If the Governing Board decides to create a descriptor for the lowest performance category, then a policy descriptor will also be needed for that level.

If the Governing Board decides to create descriptors, the range of performance in the lowest category will need to be addressed. This category covers a range from students who are unable to demonstrate skills to students who are almost Basic. There is not a clear best method for creating Below Basic ALDs. These are sometimes written to refer to what half of the students in the lowest level can do. In other cases, they are written to refer to the prerequisite skills students should have to enter the Basic category. If an ALD is created for this area, language that describes for whom the ALDs refer will need to be carefully crafted.

The reversal of current policy to create an ALD for the lowest achievement category would provide information on the KSAs of the nearly 50% of Black students and approximately 40% of Hispanic students represented in the lowest achievement level. This may become even more critical as the United States enters post-pandemic life. If U.S. students have experienced learning loss during the pandemic, then we can expect the percentages of students in the lowest category to increase.

Almost all states provide some sort of description of the lowest achievement category. In many cases, this is just a policy descriptor; however, a handful of states, as well as TIMSS and PISA, provide a reporting descriptor that lists skills students in the lowest category either can or should be able to do. States are creating the more detailed range ALDs that also list skills students in the lowest performance category either can or should be able to do. In short, there is precedence in the K–12 assessment space to create descriptors for the lowest performance category should the Governing Board decide to provide additional information about the knowledge and skills of students below the NAEP Basic level.



References

ACT (2017). Performance Level Descriptors.

https://actinc.my.salesforce.com/sfc/p/#30000000Wu5/a/1B000000QhZj/KQPOSfgRjBb vCP.9biXbH_v8btPvm8QgCjd5zj.7AqE

- Alaska Department of Education and Early Development (2017). Achievement Level Descriptors (ALDS) Mathematics Grades 6-8. https://docs.google.com/document/d/1GPbaGD7xqAt9H05eTB1oND3zdExEz5al1GtTRc PBHS0/edit#
- Olsen, R.V., & Nilsen, T. (2017). Standard setting in PISA and TIMSS and How these procedures can be used nationally. In S. Blomeke & J.-E. Gustafsson (Eds), *Standard Setting in Education* (pp. 69-84). Springer International Publishing.
- PISA (2015). Proficiency scale construction. <u>https://www.oecd.org/pisa/sitedocument/PISA-</u> 2015-Technical-Report-Chapter-15-Proficiency-Scale-Construction.pdf

TIMSS (2015). <u>http://timssandpirls.bc.edu/timss2015/international-results/wp-</u> content/uploads/filebase/full%20pdfs/T15-International-Results-in-Mathematics.pdf

Valencia, S., & Buly, M. (2004). Behind Test Scores: What Struggling Readers Really Need. *The Reading Teacher*, 57(6), 520-531. Retrieved February 26, 2021, from http://www.jstor.org/stable/20205395





Highest percentage of below *NAEP Basic* students in **2019 Grade 8** Mathematics







National Center for Education Statistics, Assessment Division

An Expert Panel Meeting on Performance Below the *NAEP Basic* Achievement Level:

A highlight of the meeting summary

May 2021





Summary Report

The Below NAEP Basic Panel

Background

As the responsible party for reporting National Assessment of Educational Progress (NAEP) results, the Assessment Division of the National Center for Education Statistics (NCES) has become increasingly concerned that it is overlooking students who perform below the *NAEP Basic* achievement level. The current NAEP reporting approach renders these students largely invisible and impedes efforts to provide evidence-based, targeted support to those most in need. To better inform policies to serve this population of students, NCES convened an expert panel on December 16–17, 2020, and charged its members with addressing questions in the three following areas:

- **Data:** What do we need to know about students who receive scores below *NAEP Basic*—who are they, what do they know, and what can they do in the various NAEP subject areas?
- **Measurement:** Are changes in NAEP procedures (e.g., design, administration, and analysis) needed to improve measurement and get a more accurate description of students who receive scores below *NAEP Basic*?
- **Reporting:** How can NAEP better describe the performance of students who score below *NAEP Basic* when reporting results?

Summary of Meeting

The panel was composed of nine experts, selected because of their strong background either in educational measurement, research, policy, or expertise in curriculum and teaching in mathematics or reading (see **pages 4–6** of this document for a brief bio of the panel members and the moderator). NCES hosted the expert panel meeting with support from the American Institutes for Research (AIR). The meeting, which consisted of two half-day sessions, opened with remarks from NCES Associate Commissioner Dr. Peggy Carr, followed by presentations of data on students at grades 4, 8, and 12 who performed below the *NAEP Basic* achievement level. Dr. George Bohrnstedt, AIR, moderated the panel discussions. In her opening remarks, Dr. Carr stated:

...high percentages of our students are performing below NAEP's basic level. We do not know enough about these students. Knowing more requires improving NAEP's processes for gathering data about them and describing them more completely. The American public, especially stakeholders in the education enterprise, should be alarmed about this growing group of underperforming students. NAEP needs to play a leadership role in better identifying who they are and what their educational needs are.

In addition, Dr. Carr differentiated the roles of NCES and the National Assessment Governing Board in NAEP assessments and procedures. She asked the panel "to be expansive and creative in thinking about the issues before us today."

NCES staff member Taslima Rahman, who organized the meeting, and two team members from AIR, Sakiko Ikoma and Markus Broer, presented data organized as three observations to give a comprehensive picture of those students who do not score, at a minimum, at the *NAEP Basic* achievement level. The first presentation, or Observation I, briefly described how NAEP results are reported to introduce the concept of "below *NAEP Basic*" and then showed results from the national, state, and district levels on students performing below *NAEP Basic*. The second presentation, or Observation II, focused on subscale and item-level data, and the third presentation, or Observation III,

Questions related to this content should be directed to Taslima Rahman at Taslima.Rahman@ed.gov.

Summary Report

The Below NAEP Basic Panel

focused on auxiliary data; that is, process data on what students did during the assessment and data from other NCES assessment surveys that could be linked to NAEP. Below are the takeaways from these three presentations.

- Results presented in Observation I showed that in 2019, the percentage of students scoring below NAEP Basic ranged from 20 to 40 percent across the three NAEP grades (4, 8, and 12) at the national level in reading and mathematics; the percentages were alarmingly high in some states (i.e., over 45 percent) and districts (i.e., over 60 percent) and they varied among student groups. Although the 2019 percentages were lower than the percentages seen in 1990, the percentage of students performing below NAEP Basic did not change much since 2003. This presentation also showed that the average scores of grade 8 and 12 students at the lower percentiles (i.e., 10th and 25th) were within the below NAEP Basic score range in both subjects and declined in recent years.
- The data shown in Observation II indicated that the percentages of students who received full credit for multiple-choice and constructed-response items were much lower for below NAEP Basic students than for students at any other achievement level; nevertheless, many below NAEP Basic students could answer some of these items correctly. The data in this presentation also showed that the score differences between below NAEP Basic and at or above NAEP Basic students across subcontent areas of mathematics and reading were similar when considered in standard deviation units.
- The results shown in Observation III indicated that although the not reached (NR) item percentage increased through the last few items for all students, students performing below *NAEP Basic* had the highest NR rate. This presentation also showed that more students among those performing below *NAEP Basic* compared to others had higher instances of not responding or responding incorrectly to practice prompts in the tutorial section.

Major Recommendations

Both the three guiding questions and the three presentations described above fostered dialogue throughout the two-day meeting. The generation of the recommendations came from overall panel discussion as well as a group process in which the panel was divided into two subgroups where each developed a set of recommendations that were then reviewed by the entire panel. The recommendations based on the panel discussion are as follows.

First, the panel recommended the development of achievement-level descriptors for students who

A Special Note:

It should be noted that, with respect to any specific recommendation, panel members did not vote individually on each recommendation; no consensus or priority ratings were sought from the panel. Therefore, it should not be implied that every panel member agreed with every recommendation and suggestion described in this document.

perform in the score range below the *NAEP Basic* cut point by outlining what students at this level know and can do. (The panel did note that achievement-level descriptions and cut points are set by NAGB.) The panel believes that the NAEP framework needs to carefully describe the construct of measurement and skill progressions required across all of the achievement levels, including what is now described as below *NAEP Basic*. This recommendation also underscores the need to name the level that is below

Questions related to this content should be directed to Taslima Rahman at Taslima.Rahman@ed.gov.
Summary Report

The Below NAEP Basic Panel

the *NAEP Basic* achievement level. Given the large range of scores below *NAEP Basic*, the panel also suggested giving consideration to including multiple levels below *NAEP Basic*, as is done in other large-scale assessment programs, such as the Program for International Student Assessment (PISA). Naming the below *NAEP Basic* score range and providing descriptions of what students who perform at this level know and can do would enrich the reporting of NAEP.

Second, the panel recommended that the distribution of items included in NAEP assessments correspond to the distribution of student ability, especially at the lower range. The current distribution of NAEP item difficulty is right-skewed and, therefore, lower performing students may become discouraged by what they see as inaccessible items. The panel suggested adding more items measuring the lower part of the NAEP scale so that the distribution of item difficulty more closely mirrors the entire distribution of student performance. The items of more appropriate difficulty will allow more precise measures of what students performing below *NAEP Basic* know and can do and add more insight into the performance of these students.

Third, the panel recommended that the NAEP reporting emphasis on students who perform below the *NAEP Basic* achievement level should, at a minimum, match the reporting emphasis for the three current achievement levels (*NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*). In addition, the panel suggested that further contextual information about students who perform below *NAEP Basic* be collected from teachers and schools so that policymakers, researchers, and the general public have a more robust set of variables from which to gain an understanding of these students' educational performance.

Other Recommendations

Other recommendations were offered during this two-day meeting to improve the NAEP program more generally. For example, the panel recommended that membership should overlap across the different committees that build the NAEP framework, develop the test specifications, write the items, and review the items both before and after reporting results. The panel believes such an overlap would promote better coordination among the various steps inherent in designing and administering the NAEP assessments, which in turn would ensure that the assessment clearly reflect the requirements laid out for the performance of students at all levels of achievement. This is an approach applied in some other large-scale assessments, such as PISA and the Program for the International Assessment of Adult Competencies (PIAAC). Other recommendations were to:

- a) consider adaptive testing to assess more precisely the performance of students at various score ranges on the NAEP scale,
- b) collect more information on instruction provided to students performing below *NAEP Basic* compared with students performing at or above *NAEP Basic*,
- c) take into consideration that the lack of reading skills and general background knowledge may hinder demonstrating the ability that is actually measured in certain items, especially for students performing below *NAEP Basic*, and
- d) conduct studies using process data to better understand the differences between those students performing below *NAEP Basic* and those performing at or above *NAEP Basic*, including how they approach items of varying type and difficulty and how motivated they are in taking the test.

Questions related to this content should be directed to Taslima Rahman at Taslima.Rahman@ed.gov.

Panel Member Bios

The Below NAEP Basic Panel

Moderator



George Bohrnstedt is a senior vice president and Institute Fellow at the American Institutes for Research (AIR). His earliest work at AIR was as Project Director of the congressionally mandated evaluation of Trial State NAEP done in collaboration with the National Academy of Education. He also chaired the National Center for Education Statistics' NAEP Validity Studies Panel (NVS) for 23 years before stepping down in October 2018. While chairing NVS, he authored or co-authored numerous studies undertaken by NVS, and provided oversight for the panel's research agenda and projects.

Panel members



Henry Braun is the Boisi Professor of Education and Public Policy and Director of the Center for the Study of Testing, Evaluation and Education Policy in the Lynch School of Education & Human Development at Boston College. After serving as an assistant professor of statistics at Princeton University, he joined the Educational Testing Service in 1979, where he served as vice-president for research

management from 1990 to 1999. He held the title of distinguished presidential appointee from 1999 until his retirement in 2006, when he moved to Boston College. A fellow of the American Statistical Association and AERA, he is an elected member of the National Academy of Education.



Ray Hart has more than 30 years of experience in research and evaluation. His work has spanned policy areas such as postsecondary success and college readiness, school improvement, teacher effectiveness, early childhood education, and adult and workforce literacy. He has worked with clients from the U.S. Department of Education, the U.S. Department of Housing and Urban

Development, the U.S. Department of State, the National Science Foundation, the National Academies of Sciences, Engineering, and Medicine, and many state and local departments of education. Dr. Hart currently serves as the Director of Research for the Council of the Great City Schools.

Panel Member Bios

The Below NAEP Basic Panel



Panel members



Hanseul Kang began her role as assistant dean and the first Anita and Joshua Bekenstein '80 B.A. Executive Director of The Broad Center at Yale School of Management in November 2020. Most recently, Kang served as the State Superintendent of Education for the District of Columbia from 2015 to 2020, providing strategic vision, clear direction, and steady leadership

to the District's state education agency, the Office of the State Superintendent of Education. Under Kang's leadership, the District of Columbia continued to make major strides in student achievement outcomes on NAEP, on its annual state assessments in math and English language arts, and on other key measures.



Irwin Kirsch is the Ralph Tyler Chair in Large-Scale Assessment and Director of the Center for Global Assessment at ETS in Princeton, NJ. Over the course of his career, Dr. Kirsch has worked in close collaboration with a number of state, national and international organizations, including the World Bank, UNESCO, the International Association for the Evaluation of Educational Achievement (IEA),

and the Organization for Economic Co-operation and Development (OECD), where he currently oversees the development and conduct of the two largest international assessments that provide policy-makers and key stakeholders with national and international comparative data on literacy and workforce preparedness: PIAAC and PISA.



Michele Mailhot joined the Maine Department of Education in July 2009 as the mathematics specialist. In this role, she works with a wide audience of stakeholders in education across the state providing professional learning to the field supporting research-based effective instructional practices in mathematics education. Before accepting the position of mathematics

specialist, Michele taught a combination of middle and high school mathematics for 11 years. In addition to her work in the classroom and at the state level, she also has worked with NASA developing curriculum for teachers across the nation with a focus on the integration of mathematics, science, and literacy.



Pamela A. Mason is a senior lecturer on education and the director of the Language and Literacy Master's program and the Jeanne Chall Reading Lab at the Harvard Graduate School of Education. Her professional and research interests encompass the role of culturally sustaining pedagogy in promoting literacy achievement, the interaction of text complexity and background knowledge, and

the intersection of literacy learning, culture, and multilingualism. She collaborates with colleagues nationally and globally on preparing reading specialist teachers and literacy coaches, developing the capacity of school leaders as literacy advocates, and evaluating schoolwide literacy programs.

Panel Member Bios

The Below NAEP Basic Panel

Panel members



Gary W. Phillips is Vice President and Chief Psychometrician at Cambium Assessment, Inc. Dr. Phillips provides senior-level guidance on all statistical and psychometric techniques and procedures used in assessment activities. This involves item response theory calibrating and equating, classical statistical analysis, differential item functions, standard setting, reliability and validity studies, generalizability studies, item banking, distributional

projections, and value-added studies involving hierarchical linear models and structural equation models. Dr. Phillips has published or presented over 200 papers, taught dozens of advanced graduate level statistics courses, and presented hundreds of workshops on advanced statistical and psychometric topics.



Jennifer Randall is an associate professor in the Research, Educational Measurement, and Psychometrics program, Director of Evaluation in the Center for Educational Assessment, and Associate Dean of Academic Affairs in the College of Education at the University of Massachusetts. Currently, she is particularly interested in the differential negative impact of both large and small-scale assessments on historically marginalized

populations in the U.S. and abroad; and the ways in which a culturally sustaining, anti-racist approach to instruction and assessment can mitigate these negative outcomes.



Lorrie A. Shepard is University Distinguished Professor in the School of Education at the University of Colorado Boulder. Dr. Shepard is past president of the American Educational Research Association and past president of the National Council on Measurement in Education. She was elected to the National Academy of Education in 1992 and served as president of the NAEd from 2005-2009. She has served on the NAEP

Validity Studies Panel from 1995 to the present. Dr. Shepard has received distinguished career awards recognizing her contributions in measurement, research, and teacher education respectively from NCME and ETS, AERA, and AACTE.







Highlights of the 2018 NAEP Oral Reading Fluency Study







Highlights of the 2018 NAEP Oral Reading Fluency Study

APRIL 2021

Sheida White American Institutes for Research

John Sabatini University of Memphis

Bitnara Jasmine Park American Institutes for Research

Jing Chen National Center for Education Statistics

Jared Bernstein Analytic Measures, Inc.

Mengyi Li American Institutes for Research

U.S. Department of Education Miguel Cardona *Secretary*

Institute of Education Sciences Mark Schneider *Director*

National Center for Education Statistics James L. Woodworth *Commissioner*

Assessment Division

Peggy G. Carr Associate Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

NCES, IES, U.S. Department of Education Potomac Center Plaza (PCP) 550 12th Street SW Washington, DC 20202

April 2021

The NCES Home Page address is <u>https://nces.ed.gov</u>. The NCES Publications and Products address is <u>https://nces.ed.gov/pubsearch</u>.

This publication is only available online. To download, view, and print the report as a PDF file, go to the NCES Publications and Products address shown above.

This report was prepared for the National Center for Education Statistics under Contract No. ED-IES-12-D-0002 with American Institutes for Research. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. government.

Suggested Citation

White, S., Sabatini, J., Park, B. J., Chen, J., Bernstein, J., and Li, M. (2021). *Highlights of the 2018 NAEP Oral Reading Fluency Study* (NCES 2021-026). U.S. Department of Education. Washington, DC: Institute of Education Sciences, National Center for Education Statistics. Retrieved [date] from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021026.

Content Contact

Jing Chen (202) 245-8324 jing.chen@ed.gov

Contents

	Page
List of Tables	1V
List of Figures	iv
Introduction	1
Purpose of the ORF Study	1
Importance of Measuring Oral Reading Fluency	1
Foundational Skills for Fluency	2
The Role of Language Comprehension in Reading Comprehension and Oral Reading	2
Method	3
Measures of Oral Reading Fluency and Foundational Skills	3
Operationalization of the Measures	3
Tasks	4
Scoring	4
Findings	5
New Data on the Reading Skills of Fourth-Graders Performing Below NAEP Basic	5
Characteristics of Students in the Below <i>NAEP Basic</i> Subgroups	5
Oral Reading Fluency and Foundational Skills for the Below NAEP Basic	
Subgroups	7
1: Oral reading fluency (passage reading)	7
ORF passage reading WCPM	7
ORF passage reading accuracy	8
ORF passage reading expression	10
Summary of findings on ORF passage reading and NAEP reading	
performance	10
2: Foundational skills (word and pseudoword list reading)	12
Word reading WCPM	12
Pseudoword reading WCPM	13
Conclusion and Implications	14
Conclusion	14
Implications	14
The NAEP reading framework and future assessments	14
Policy and research	15
References	16

List of Tables

Table

List of Figures

Figure

Page

Page

1.	Average passage reading WCPM, by NAEP reading achievement level and below <i>NAEP Basic</i> subgroup: 2018	8
2.	Average passage reading accuracy, by NAEP reading achievement level and below <i>NAEP Basic</i> subgroup: 2018	9
3.	Average passage reading expression, by NAEP reading achievement level and below <i>NAEP Basic</i> subgroup: 2018	11
4.	Average word reading WCPM, by NAEP reading achievement level and below <i>NAEP Basic</i> subgroup: 2018	12
5.	Average pseudoword reading WCPM, by NAEP reading achievement level and below <i>NAEP Basic</i> subgroup: 2018	13

Introduction

This publication highlights the key concepts and findings of the 2018 National Assessment of Educational Progress (NAEP) Oral Reading Fluency (ORF) study. For additional details, see its companion publication, *The 2018 NAEP Oral Reading Fluency Study* (White et al. 2021), which is available on the NAEP ORF website at https://nces.ed.gov/nationsreportcard/studies/orf/.

Oral reading fluency is defined as the ability to read text aloud with speed, accuracy, and proper expression. The 2018 NAEP ORF study was the first such NAEP study since 2002. It was administered to a nationally representative sample of over 1,800 fourth-graders from 180 public schools.

The students in the ORF study first completed the NAEP fourth-grade reading assessment. Then they moved on to the tasks that were administered for this study: reading out loud four short passages to assess oral reading fluency and two word lists to assess skills that provide the foundation for fluency. Both the NAEP reading tasks and the study tasks were administered to students on tablets, and students' responses were recorded on the tablets.

Purpose of the ORF Study

The purpose of the 2018 NAEP ORF study is to add new, policy-relevant information to the NAEP reading assessment. It includes, for the first time, a close examination of the oral reading fluency and foundational skills of fourth-grade public school students who perform below *NAEP Basic* on the NAEP reading assessment. "Foundational skills," word reading and phonological decoding, are defined under the heading, *Foundational Skills for Fluency*.

Importance of Measuring Oral Reading Fluency

Students who read aloud with appropriate speed, accuracy, and expression (i.e., students who have oral reading fluency) are more likely to comprehend connected text (Sabatini, Wang, and O'Reilly 2019) because they are able to conserve cognitive resources that can be applied to the comprehension of meaning (Perfetti 2007). Thus, oral reading fluency is a reliable and easily accessible indicator of overall reading competence—and a strong marker of progress in learning to read (Fuchs et al. 2001)—and its assessment has become one of the primary means of determining which elementary school students are on track toward meeting state reading standards and which students would benefit from additional services and intervention (McGlinchey and Hixson 2004; Reschly et al. 2009).

Foundational Skills for Fluency

The term "foundational skills" refers to two skills that are assessed in this study with two different word lists: (1) *word reading* (also known as word recognition)—the ability to read familiar words with accuracy and speed—and (2) *phonological decoding*—the ability to pronounce unfamiliar words based on knowledge of spelling-sound correspondences.¹ Research has established that fast and accurate word reading is a major driver of oral reading fluency (e.g., Eason et al. 2013; Metsala and David 2017; Silverman et al. 2013). In addition, phonological decoding is regarded by almost all reading researchers as a critical prerequisite for the development of skilled, fluent reading and reading comprehension. In essence, as children apply phonological decoding skills to the unfamiliar words that they encounter in text, they make a transition from being "novices" to being "experts" who read familiar words rapidly and automatically (Castles, Rastle, and Nation 2018; Share 1995).

The Role of Language Comprehension in Reading Comprehension and Oral Reading

Although the ORF study focuses on oral reading fluency, word reading, and phonological decoding, they are not the only factors that may affect performance on the fourth-grade NAEP reading assessment. One of the most important factors is language comprehension. "Language comprehension" is the ability to understand language based on knowledge of the meaning of words, sentence structure, and other aspects of language.² An extensive body of research and theory supports the view that language comprehension is necessary for reading comprehension (see, e.g., Foorman, Petscher, and Herrera 2018; Hoover and Gough 1990).

It is important to recognize that oral reading also involves language comprehension, just as silent reading and reading comprehension do. First, when students read a passage out loud with appropriate expression, they are using their ability to comprehend language as well as read the words in the passage. Second, when students read a passage out loud, they use their knowledge of word meaning and sentence structure to anticipate and recognize (read) the words in the text. This process is called "contextual facilitation of word recognition."³ Therefore, oral passage reading (fluency) tasks are measuring language comprehension in addition to fast and accurate word reading. This implicit measurement of language comprehension is one of the reasons why oral reading fluency assessments are valued by educators and widely used in elementary schools (Reschly et al. 2009).

¹ Many researchers consider phonological awareness to be another critically important foundational skill. It was not measured in this study because it is rapidly and fully acquired by normally developing readers in preschool, kindergarten, and first grade.

 $^{^{\}rm 2}$ Language comprehension is measured by tests that require no reading, such as orally administered vocabulary tests and listening comprehension tests.

³ Contextual facilitation has been extensively studied by researchers. Evidence comes from (1) experimental studies showing, for example, that coherent passages are read more rapidly than text containing the same words in random order; and (2) correlational studies showing that, for example, vocabulary and listening comprehension affect oral reading fluency when word reading skills are controlled statistically.

Method

Measures of Oral Reading Fluency and Foundational Skills

The 2018 NAEP ORF study included measures of oral reading fluency, word reading, and phonological decoding. The last two are regarded as foundational skills for fluency.

- Oral reading fluency (passage reading) refers to the ability to read connected text such as paragraphs and passages with appropriate rate, accuracy, and expression, which is an indicator of comprehension.
- *Word reading* (also known as *word recognition*) refers to the ability to recognize familiar written words with appropriate speed and accuracy, relying primarily on orthographic memory (memory of how the words are pronounced).
- *Phonological decoding* refers to the ability to pronounce unfamiliar words based on knowledge of spelling-sound correspondences.

As noted previously, many words that students initially pronounce by "sounding them out" eventually become automatically recognized as chunks of letters or whole words in a process that requires minimal conscious effort. This is why it is important to measure both the ability to phonologically decode unfamiliar words and the ability to recognize familiar words.

Operationalization of the Measures

Each of the above measures was operationalized in terms of two aspects of performance—rate and accuracy—as well as a combination of the two, words correct per minute.

- *Words correct per minute (WCPM)* refers to the total number of words correctly read divided by the amount of time taken to read the passages or word-level lists. This is the WCPM score.
- Accuracy refers to the percentage of words that was read accurately. For passages, the total number of attempted words⁴ in the passage was the denominator, and for word lists, the total number of words presented to students was the denominator.

⁴ Attempted words included words read correctly or incorrectly as well as those that were skipped.

Passage reading was operationalized in terms of one additional measure—expression—defined below:

• *Expression* refers to appropriate intonation, rhythm, emphasis, and pausing that groups words into phrasal and larger units in ways that express the meaning and structure of the text and enhance understanding and enjoyment in a listener.

Tasks

The following text materials were given to students to be read aloud:

- *Text passages,* consisting of 152–162 words, providing a measure of fourth-graders' ability to read words and sentences in connected text.
- *Word lists,* consisting of 24 English words arranged in increasing order of complexity, providing a measure of individual students' ability to recognize familiar words.
- *Pseudoword lists,* consisting of 18 made-up but pronounceable words (e.g., *jad*), providing a measure of students' ability to decode words they are unfamiliar with.

The word and pseudoword lists used in this study were developed based on principles derived from clinically valid measures of children's acquisition of word recognition and phonological decoding. Moreover, these word-level tasks along with the text passages were tested in cognitive laboratory studies administered by NAEP ORF team researchers to ensure they were within typical fourth-graders' ability to perform.

Scoring

In this study, NCES used a new automatic speech analysis/scoring system that calculated accuracy, rate, and WCPM variables to score recordings of students' reading. In preparation for scoring the tasks administered for this study, extensive work was done to ensure that correct word pronunciation would be scored reliably and that speakers of nonstandard varieties of English would not be unfairly penalized. The scoring system considered nonstandard pronunciations acceptable as long as they were consistent with the participants' general speaking pattern.

Scoring of the *Expression* variable, which is based on a detailed rubric,⁵ involved thorough training of human scorers and multiple levels of quality checks. To ensure reliability of scoring, supervisors spot-checked scores and provided feedback to scorers. In addition, a second scorer rescored 25 percent of all passage reading recordings to monitor interrater reliability (i.e., agreement between scorers on the scores assigned).

⁵ The scoring rubric for the Expression variable can be found in the companion publication, *The 2018 NAEP Oral Reading Fluency Study* (White et al. 2021) on the NAEP ORF website, <u>https://nces.ed.gov/nationsreportcard/studies/orf/</u>.

Findings

The findings of the 2018 NAEP ORF study are uniquely useful for exploring the question of how NAEP reading performance is related to oral reading fluency, word reading, and phonological decoding skills. As with all NAEP findings, it is important to remember that cause-and-effect relationships cannot be inferred from descriptive and correlational results. NAEP reading performance, oral reading fluency, word reading, and phonological decoding may be affected by a complex mixture of factors beyond the scope of the study.

New Data on the Reading Skills of Fourth-Graders Performing Below NAEP Basic

A major objective of the 2018 ORF study was to provide a nuanced picture of the reading performance of low-performing fourth-grade readers. To accomplish this, students performing below *NAEP Basic* were evenly divided into three groups based on the NAEP reading score distribution. The three groups were labeled below *NAEP Basic* **Low** (i.e., the bottom one-third of the students performing below *NAEP Basic*), below *NAEP Basic* **Medium** (i.e., the middle one-third of the students performing below *NAEP Basic*), and below *NAEP Basic* **High** (i.e., the top one-third of the students performing below *NAEP Basic*). Students' characteristics and oral reading performance were then compared across these subgroups.

Characteristics of Students in the Below *NAEP Basic* Subgroups

Overall, 36 percent of fourth-grade public school students performed below *NAEP Basic*, but 51 percent of Black fourth-grade students and 46 percent of Hispanic fourth-grade students performed below *NAEP Basic*.⁶ We found that Black students were also overrepresented in the lowest below *NAEP Basic* subgroup—i.e., below *NAEP Basic* Low. As shown in table 1, while 26 percent of the White students performing below *NAEP Basic* were at the lowest level of below *NAEP Basic*, 40 percent of the Black fourth-graders and 37 percent of the Hispanic fourth-graders who performed below *NAEP Basic* fell into this subgroup. Because 51 percent of Black students were in the below *NAEP Basic* group, this finding means that 20 percent of Black fourth-grade students (or one out of every five Black fourth-graders) performed at the lowest end of below *NAEP Basic* (51 percent × 40 percent = 20 percent). Similarly, 17 percent (or one out of six) of Hispanic fourth-graders were in the lowest in the lowest below *NAEP Basic* Intervent in the lowest below *NAEP Basic* (51 percent × 40 percent = 20 percent). Similarly, 17 percent (or one out of six) of Hispanic fourth-graders were in the lowest below *NAEP Basic* group, below *NAEP Basic* Low (46 percent × 37 percent = 17 percent).

⁶ Here we have reported the observed percentages for the ORF study sample, which are very close to the percentages for the operational NAEP sample. For Black students in the operational NAEP sample, the percentages of students performing below *NAEP Basic* were 50 and 53 for 2017 and 2019, respectively. For Hispanic students in the operational NAEP sample, the percentage of students performing below *NAEP Basic* was 46 in both 2017 and 2019.

Overall, 36 percent of fourth-grade public school students performed below *NAEP Basic*, but 50 percent of National School Lunch Program (NSLP)-eligible fourth-grade students performed below *NAEP Basic*. As shown in table 1, among students who performed below *NAEP Basic*, NSLP-eligible students were nearly equally divided among the three below *NAEP Basic* subgroups. About 35 percent of the NSLP-eligible students performed at the lowest below *NAEP Basic* level.

Student	below NAEP Basic	below NAEP Basic	below NAEP Basic			
characteristics	Low	Medium	High	Total		
All students	33	33	33	100		
Race/ethnicity						
White	26	35	39	100		
Black	40	31	28	100		
Hispanic	37	33	30	100		
NSLP eligibility						
Eligible	35	34	31	100		
Not eligible	27	32	41	100		

Table 1. Percentage of fourth-graders performing below *NAEP Basic*, by below *NAEP Basic* subgroup and selected student characteristics: 2018

NOTE: Rows may not sum to totals because of rounding. For National School Lunch Program (NSLP) eligibility, about 2 percent of the students lacked valid eligibility information. These students were also excluded because of small sample size. Learn more about the NAEP achievement levels <u>here</u>.

Oral Reading Fluency and Foundational Skills for the Below NAEP Basic Subgroups

As mentioned earlier, what this study adds to the previous studies of NAEP reading is a closer examination of the difficulties faced by fourth-grade students performing below *NAEP Basic* on the NAEP reading assessment. Because the 2018 NAEP ORF study participants had completed the NAEP reading assessment, it was possible to examine the relationship between reading achievement and each of the measures.

1: Oral reading fluency (passage reading)

ORF passage reading WCPM

As shown in figure 1, passage reading words correct per minute (WCPM) decreased significantly in moving down from the *NAEP Advanced* group to the *NAEP Proficient* group and *NAEP Basic* group.⁷ Also, and importantly, passage reading WCPM decreased significantly within the below *NAEP Basic* group. In moving down the subgroups, the average for students in the below *NAEP Basic* **High** subgroup was 108 WCPM, the average for students in the below *NAEP Basic* **Medium** subgroup was 95 WCPM, and the average for students in the below *NAEP Basic* **Low** subgroup was 71.

It is noteworthy that the passage reading WCPM difference between the lowest below *NAEP Basic* subgroup (below *NAEP Basic* Low) and the highest below *NAEP Basic* subgroup (below *NAEP Basic* High) is as large at 38 WCPM⁸ as the difference between the *NAEP Basic* and *NAEP Advanced* groups (37 WCPM).

The average passage reading WCPM across all levels was 120. To help put all of these numbers in perspective, based on the 2003 National Assessment of Adult Literacy (NAAL) (Baer et al. 2009), adult readers performing at the Intermediate and Proficient levels read orally at an average of 166 and 178 words correctly per minute (WCPM), respectively. This indicates that there is room for improvement even for fourth-grade students performing at the *NAEP Proficient* level (142 WCPM) and considerable room for improvement for fourth-grade students performing at the *NAEP Basic* level (123 WCPM).

⁷ All comparisons were conducted with an alpha level of 0.05, with multiple pairwise comparison adjustments applied when needed using the False Discovery Rate (FDR) procedure.

⁸ Unrounded numbers were used for calculating the differences between the estimates.



NAEP achievement level and below NAEP Basic subgroup

* Statistically significant difference compared to the next higher NAEP reading achievement level category, p < .05. All comparisons were conducted with an alpha level of 0.05, with multiple pairwise comparison adjustments applied using the False Discovery Rate (FDR) procedure.

NOTE: WCPM is an abbreviation for words correct per minute. The positions of the data points in the graphics are based on the unrounded numbers.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2018 Oral Reading Fluency study.

ORF passage reading accuracy

Perhaps the most noticeable difference among students performing below *NAEP Basic* is in passage reading accuracy. As shown in figure 2, the passage reading accuracy of students performing below *NAEP Basic* **Low** was 82 percent, about 9 percentage points⁹ and 12 percentage points lower than the below *NAEP Basic* **Medium** and **High** subgroups, respectively.

Eighty-two percent accuracy in practical terms means that students misread 1 out of every 6 words. Students who frequently misread words are likely to have difficulty understanding the text because the words are apt to be content words that are important for comprehension, not

8 Highlights of the 2018 NAEP Oral Reading Fluency Study: Findings

51

⁹ Unrounded numbers were used for calculating the differences between the estimates.

function words¹⁰ (e.g., the, and, on). Also, at 92 percent correct, the below *NAEP Basic* **Medium** group was missing 1 out of every 11 words, which is 1 word in nearly every sentence. The average percentage of words read correctly across all levels was 94 percent.



NAEP achievement level and below NAEP Basic subgroup

* Statistically significant difference compared to the next higher NAEP reading achievement level category, p < .05. All comparisons were conducted with an alpha level of 0.05, with multiple pairwise comparison adjustments applied using the False Discovery Rate (FDR) procedure.

NOTE: Accuracy refers to the percentage of words that was read accurately. The positions of the data points in the graphics are based on the unrounded numbers.

¹⁰ To hear an audio recording that illustrates the kind of words that were misread or read with difficulty, see <u>https://nces.ed.gov/nationsreportcard/studies/orf/illustrative_audio.aspx</u>.

ORF passage reading expression

Passage reading expression scores (figure 3) showed the same pattern as the passage reading WCPM. Performance declined steadily from *NAEP Advanced* to *NAEP Proficient* to *NAEP Basic* and continued to decline from below *NAEP Basic* **High** to below *NAEP Basic* **Medium** and below *NAEP Basic* **Low**. The average passage reading expression score for all fourth-grade students was at Level 4 on a scale of 0–5. That indicated that their oral reading expressed sentence structure and meaning, and that more than three-quarters of the words in the passage were read with appropriate expression.

For all readers performing below *NAEP Basic*, the average score was in the Level 3 range. That meant that their oral reading expressed the meaning of words, phrases, clauses, and a few sentences, and that they read more than half of the words in the passage with appropriate expression. For the lowest below *NAEP Basic* subgroup, below *NAEP Basic* **Low**, the average expression score fell below Level 3. That indicated that these students tended to focus on local word groupings, which means that they often paused in the middle of a phrase. For example, the sentence "Hawaii is a warm place, but parts of it are cold" would be read as [Hawaii] [is a] [warm place], [but parts of] [it are] [cold].¹¹

Summary of findings on ORF passage reading and NAEP reading performance

Overall, across all of the passage reading data described above (WCPM, accuracy, and expression), there is a strong and consistent relationship between the NAEP reading assessment performance and passage reading. The above figures also show that there is noticeable variation among the below *NAEP Basic* subgroups for every passage reading measure.

¹¹ Passage reading expression by a student in the below *NAEP Basic* **Low** subgroup is illustrated in an audio recording that can be found here: <u>https://nces.ed.gov/nationsreportcard/studies/orf/illustrative_audio.aspx</u>.



NAEP achievement level and below NAEP Basic subgroup

* Statistically significant difference compared to the next higher NAEP reading achievement level category, p < .05. All comparisons were conducted with an alpha level of 0.05, with multiple pairwise comparison adjustments applied using the False Discovery Rate (FDR) procedure.

NOTE: The positions of the data points in the graphics are based on the unrounded numbers. Expression score 0 = Insufficient passage reading sample for accurate rating; 1 = Reading is word by word; less than a quarter of the words are read with appropriate expression; 2 = Reading focuses on local grouping; less than half of the words are read with appropriate expression; 3 = Reading expresses the meaning of words, phrases, clauses, and a few sentences; more than half of the words are read with appropriate expression; 4 = Reading expresses sentence structure and meaning; more than three-quarters of the words are read with appropriate expression; 5 = Passage is read as if for a listener and is expressive throughout. For detailed passage reading expression score description, see *The 2018 NAEP Oral Reading Fluency Study* (White et al. 2021) on the NAEP ORF website, https://nces.ed.gov/nationsreportcard/studies/orf/.

2: Foundational skills (word and pseudoword list reading)

Word reading WCPM

By examining word-level reading apart from a passage, we could better understand the wordlevel processes that underlie fluency and passage comprehension. What we learned is that the foundational skills—word reading and phonological decoding—also varied widely within the below *NAEP Basic* subgroups (figures 4 and 5).

In word list reading, students read high-frequency words that have known meanings to most students in fourth grade. Performance on this task was regarded an indicator of accumulating knowledge of printed words and an increasing ability to read words rapidly and automatically without effortful decoding. As shown in figure 4, word reading declined across the *NAEP Basic* level through all the below *NAEP Basic* subgroups. The sharpest decline was between the below *NAEP Basic* Low subgroups.



NAEP achievement level and below NAEP Basic subgroup

* Statistically significant difference compared to the next higher NAEP reading achievement level category, p < .05. All comparisons were conducted with an alpha level of 0.05, with multiple pairwise comparison adjustments applied using the False Discovery Rate (FDR) procedure.

NOTE: WCPM is an abbreviation for words correct per minute. The positions of the data points in the graphics are based on the unrounded numbers.

Pseudoword reading WCPM

In pseudoword list reading (i.e., reading lists of made-up, but pronounceable words), students read made-up words that required them to use phonological decoding skills, the skills that enable a reader to pronounce sequences of letters based on knowledge of spelling-sound correspondences and orthographic patterns.

Like word reading skills, pseudoword reading skills declined across the NAEP reading achievement levels, including the below *NAEP Basic* subgroups (figure 5). The decline in mean performance was especially sharp between the below *NAEP Basic* **Medium** and below *NAEP Basic* **Low** subgroups. Moreover, there was a wide range within the below *NAEP Basic* subgroups. Fourth-graders in the below *NAEP Basic* **High** group read almost twice as many words correctly per minute (19) as those in the below *NAEP Basic* **Low** group (11), as shown in figure 5. The average number of pseudowords read correctly per minute was 22 for all fourth-grade students.





NAEP achievement level and below NAEP Basic subgroup

* Statistically significant difference compared to the next higher NAEP reading achievement level category, p < .05. All comparisons were conducted with an alpha level of 0.05, with multiple pairwise comparison adjustments applied using the False Discovery Rate (FDR) procedure.

NOTE: Pseudoword is a made-up but pronounceable word. WCPM is an abbreviation for words correct per minute. The positions of the data points in the graphics are based on the unrounded numbers.

Conclusion and Implications

Conclusion

The 2018 ORF study reveals that for an estimated 1.27 million¹² fourth-grade public school students performing below *NAEP Basic*, and particularly for an estimated 0.42 million¹³ fourth-grade students in the below *NAEP Basic* **Low** subgroup, fluent reading of connected text—sufficiently fast and accurate reading of sentences and passages—can be a major challenge. The study also shows that word reading and phonological decoding skills are underdeveloped in students performing below *NAEP Basic*, particularly for students in the below *NAEP Basic* **Low** subgroup.

Students in the below *NAEP Basic* **Low** subgroup not only have difficulty reading the words in the text quickly and accurately but also show a lack of appropriate expression in reading out loud, which is an indicator of poor comprehension. This makes it difficult for them to engage in the cognitive processes described in the 2017 NAEP reading framework. For an illustrative audio recording, see https://nces.ed.gov/nationsreportcard/studies/orf/illustrative_audio.aspx.

Implications

The NAEP reading framework and future assessments

First, the current reading framework does not describe any specific reading behaviors that characterize fourth-grade students performing below *NAEP Basic*. It states only that "These students are not necessarily nonreaders; many can complete some tasks on the assessment but are not able to attain the minimum score required for *Basic*" (National Assessment Governing Board 2017, p. 44). Based on the findings of this study, the new framework should incorporate a description of readers performing below *NAEP Basic*. It should acknowledge the fact that, compared to students performing at the *NAEP Basic* level or higher, students performing below *NAEP Basic* are more likely to have underdeveloped fluency, word reading, and phonological decoding skills. There should also be additional testing of fourth-grade students' oral reading fluency and foundational skills with a subsample of the students who take the main NAEP reading assessment. Such testing would provide much-needed information about the students who are performing below *NAEP Basic*.

 $^{^{12}}$ This number refers to 36 percent of 3.54 million (the number of public school, fourth-graders represented in the 2018 ORF study sample) = 1.27 million.

¹³ This number refers to a third of 1.27 million fourth-grade students who performed below *NAEP Basic*. Recall that students performing below *NAEP Basic* were evenly divided into three groups based on the NAEP reading score distribution.

Second, the framework (p. 4) notes that text comprehension is influenced by phonics knowledge and fluency; and, importantly, it recognizes that "without these foundational skills, comprehension will not occur." It goes on to state a goal or aspiration for fourth-grade students that is universally accepted by reading experts and reading educators: "By grade 4, when the NAEP Reading Assessment is first administered, students should have a well-developed understanding of how sounds are represented alphabetically and should have had sufficient practice in reading to achieve fluency with different kinds of texts" (p. 4). But what if this goal has not been met?

In the future, the framework should acknowledge that: "Although the majority of fourth-grade students do not have problems with fluency, word reading, and phonological decoding, these skills are not adequately developed for a significant percentage of readers performing below *NAEP Basic,*" as shown by the findings of the 2018 NAEP ORF report (White et al. 2021).

Policy and research

First, the problems of fourth-grade students performing below *NAEP Basic* highlighted by this report call for a solution-oriented discussion among education policymakers. The discussion may begin with recognition of the large income-based gaps in prereading skills that exist at kindergarten entry (Quinn 2015; Reardon and Portilla 2016) and proceed to a fresh and intensive look at programs of instruction in preschools and the early elementary grades, especially programs that enroll large numbers of Black and Hispanic children. Second, research is needed to determine the extent to which elementary schools teach accurate and efficient word reading skills, in systematic ways, as supported by existing research (e.g., Castles, Rastle, and Nation 2018). This is a topic that is being vigorously debated in policy circles at the present time.

References

- Baer, J., Kutner, M., Sabatini, J., and White, S. (2009). Basic Reading Skills and the Literacy of America's Least Literate Adults: Results From the 2003 National Assessment of Adult Literacy (NAAL) Supplemental Studies (NCES 2009-481). U.S. Department of Education. Washington, DC: National Center for Education Statistics, Institute of Education Sciences.
- Castles, A., Rastle, K., and Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest, 19:* 5-51.
- Eason, S.H., Sabatini, J., Goldberg, L., Bruce, K., and Cutting, L.E. (2013). Examining the Relationship Between Word Reading Efficiency and Oral Reading Rate in Predicting Comprehension Among Different Types of Readers. *Scientific Studies of Reading*, 17: 199-223.
- Foorman, B.R., Petscher, Y., and Herrera, S. (2018). Unique and Common Effects of Decoding and Language Factors in Predicting Reading Comprehension in Grades 1–10. *Learning and Individual Differences, 63:* 12-23.
- Fuchs, L., Fuchs, D., Hosp, M.K., and Jenkins, J.R. (2001). Oral Reading Fluency as an Indicator of Reading Competence: A Theoretical, Empirical, and Historical Analysis. *Scientific Studies of Reading*, *5*: 239-256.
- Hoover, W.A., and Gough, P.B. (1990). The Simple View of Reading. *Reading and Writing: An Interdisciplinary Journal, 2:* 127-160.
- McGlinchey, M.T., and Hixson, M.D. (2004). Using Curriculum-based Measurement to Predict Performance on State Assessments in Reading. *School Psychology Review*, 33: 193-203.
- Metsala, J.L., and David, M.D. (2017). The Effects of Age and Sublexical Automaticity on Reading Outcomes for Students With Reading Disabilities. *Journal of Research in Reading*, 40: S209-S227.
- National Assessment Governing Board. (2017). *Reading Framework for the 2017 National Assessment of Educational Progress.* Washington, DC: Author.
- Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11: 357-38.
- Quinn, D.M. (2015). Kindergarten Black–White Test Score Gaps: Re-examining the Roles of Socioeconomic Status and School Quality With New Data. *Sociology of Education, 88:* 120-139.
- Reardon, S.F., and Portilla, X.A. (2016). Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry. *AERA Open, 2:* 1-18.
- Reschly, A.L., Busch, T.W., Betts, J., Deno, S.L., and Long, J.D. (2009). Curriculum-based Measurement of Oral Reading as an Indicator of Reading Achievement: A Meta-analysis of the Correlational Evidence. *Journal of School Psychology*, 47: 427-469.

- Sabatini, J., Wang, Z., and O'Reilly, T. (2019). Relating Reading Comprehension to Oral Reading Performance in the NAEP Fourth-Grade Special Study of Oral Reading. *Reading Research Quarterly, 42:* 253-271.
- Share, D.L. (1995). Phonological Recoding and Self-teaching: Sine Qua Non of Reading Acquisition. *Cognition*, 55: 151-218.
- Silverman, R.D., Speece, D.L., Harring, J.R., and Ritchey, K.D. (2013). Fluency has a Role in the Simple View of Reading. *Scientific Studies of Reading*, *17*: 108-133.
- White, S., Sabatini, J., Park, B.J., Chen, J., Bernstein, J., and Li, M. (2021). *The 2018 NAEP Oral Reading Fluency Study* (NCES 2021-025). U.S. Department of Education. Washington, DC: Institute of Education Sciences, National Center for Education Statistics.





Studies to Review and Revise NAEP Achievement Level Descriptions (ALDs) for Mathematics, Reading, and Other Subjects

Background

On September 24, 2020, the National Assessment Governing Board (Governing Board) awarded contract# 91995920C0004 to Pearson (as a result of a competitive bidding process) for conducting studies to review and revise NAEP achievement level descriptions (ALDs) in mathematics and reading using the 2019 NAEP assessments at grades 4, 8, and 12¹. This work is intended to address the first recommendation of the <u>evaluation of NAEP achievement levels that</u> was conducted by the National Academies of Sciences, Engineering, and Medicine:

Recommendation #1: Alignment among the frameworks, the item pools, the achievement-level descriptors, and the cut scores is fundamental to the validity of inferences about student achievement. In 2009, alignment was evaluated for all grades in reading and for grade 12 in mathematics, and changes were made to the achievement-level descriptors, as needed. Similar research is needed to evaluate alignment for the grade 4 and grade 8 mathematics assessments and to revise them as needed to ensure that they represent the knowledge and skills of students at each achievement level. Moreover, additional work to verify alignment for grade 4 reading and grade 12 mathematics is needed.

The Board committed to conducting studies to review and revise the NAEP ALDs in its initial response to the evaluation that was formally adopted and sent to the Secretary of Education and Congress in December 2016. The Board's <u>Achievement Levels Work Plan</u>, adopted in March 2020, further describes the intention for this work: "Addressing Recommendation #1 should focus on the current reporting ALDs for mathematics and reading at grades 4, 8, and 12. The methodology will be similar to what was done to evaluate the alignment and revise the 2009 NAEP Reading ALDs for grades 4, 8, and 12 (Donohue, Pitoniak, & Beaulieu, 2010) and the 2009 NAEP Mathematics ALDs for grade 12 (Pitoniak, Dion, & Garber, 2010). This process will generate new reporting ALDs that comply with the revised Board policy statement" (p. 3).

According to Principle 1a of the Board policy on <u>Developing Student Achievement Levels for</u> <u>NAEP</u>, "Content achievement level descriptions translate the policy definitions into specific

¹ The base period of this contract includes the review and revision of ALDs in mathematics and reading at grades 4, 8, and 12; in addition, an option may be exercised for a second phase of the contract focusing on review and revision of ALDs in U.S. history, civics, science, technology and engineering literacy (TEL) at grade 8 based on data from the most recent administrations of those assessments in 2018 and 2019.

expectations about student knowledge and skills in a particular content area, at each achievement level, for each subject and grade. Content ALDs provide descriptions of specific expected knowledge, skills, or abilities of students performing at each achievement level. They reflect the range of performance that items and tasks should measure. When setting achievement levels, the content ALDs provide consistency and specificity for panelist interpretations of policy definitions for a given assessment. During reporting, content ALDs communicate the specific knowledge and skills represented by *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* for a given assessment" (p. 5).

Principles 3g and 4a of the Board policy apply specifically to this project of reviewing and revising the current ALDs and creating reporting ALDs (based on empirical data) that indicate what students at each achievement level *do* know and *can* do rather than what they *should* know and *should* be able to do². Additional details for carrying out the work described by principles 3g and 4a are included in the <u>Achievement Levels Procedures Manual</u>.

The basis for the evaluation of NAEP achievement levels (and subsequently for this project) is the *existing* NAEP frameworks and item pools, not the new NAEP Mathematics Framework currently scheduled for implementation in 2026 or the NAEP Reading Framework that is currently under development and consideration by the Board. In accordance with principle 4b of the Board policy, the achievement levels and/or ALDs will need to be reviewed again once the new frameworks are implemented. Such work is beyond the scope of this project.

Project Overview

Dr. Eric Moyer is the project director at Pearson and Dr. Jennifer Galindo is the assistant project director at Pearson. Pearson will conduct a pilot study and an operational meeting using scale anchoring studies where panels of content experts judge the alignment of the current mathematics and reading ALDs and produce a set of recommended reporting ALDs for the Governing Board to consider in reporting the results from the next regular administration of the NAEP reading and mathematics assessments at grades 4, 8, and 12. The Governing Board is expected to take action on the reporting ALDs for mathematics and reading at grades 4, 8, and 12 in advance of the next release of these results.

Based on careful review of the history of ALD development, review, and revisions for NAEP mathematics and reading, a model-based anchored approach for reviewing the alignment of the ALDs for NAEP mathematics and reading will be used. The methodology for this alignment review study is based on that of previous studies, including the ALD development and review meeting held in 2009. The methodology was specified by the Board's Achievement Levels Work Plan and was selected to reduce the potential for possible inconsistencies from the use of different methods. The process of the model-based anchored approach will result in organizing

² According to the Board policy, ALDs will continue to describe what students *should* know and *should* be able to do for the purposes of item development and standard setting; only the reporting ALDs will be written in terms of what students *do* know and *can* do.

specific NAEP items by achievement level, which will serve as a key referent for panelists in reviewing and revising the current ALDs.

The model-based anchored approach includes three stages. The first stage will involve conducting statistical analyses to determine the items from the subject and grade that are anchored to a level corresponding to the score range within cut scores set to represent the achievement level descriptors (ALDs). The second stage relies on panels of content experts for each individual assessment. The panelists individually review the items that are anchored to each performance level and create summary descriptions of what students in each level are expected to know and be able to demonstrate based on the knowledge and skills measured by the items. In the final stage, the panelists compare the current ALDs for the respective assessment with their summary descriptions. The panelists note the similarities and differences, to make a recommendation regarding whether the current ALDs accurately describe what students in each level are expected to know and be able to demonstrate or if revisions to the current ALDs are needed to improve alignment. The final alignment judgment will be used to report whether the panels determined that there exists alignment between the current ALDs and student expectations. The final panel summary descriptions will be used to revise the current ALDs to create reporting ALDs that indicate what students at each achievement level do know and can do.

There is a technical advisory committee (TAC) consisting of the following experts in ALDs:

Dr. Karla Egan (Principal, EdMetric)

Dr. Ellen Forte (CEO and Chief Scientist, edCount)

Dr. Susan Loomis (Independent Consultant)

Dr. Marianne Perie (President, Measurement in Practice)

Dr. Mark Reckase (University Distinguished Professor Emeritus, Michigan State University)

Dr. Lauress Wise (Principal Scientist, Human Resources Research Organization)

The TAC is scheduled to meet for more than 100 hours (approximately 4 hours per month, with additional meeting time following the pilot and operational meetings) to provide technical advice on all aspects of the project to review and revise the mathematics and reading ALDs; this is intended to help ensure that all procedures, materials, and reports are carried out in accordance with current best practices, providing additional validity evidence for the process and results. In addition to frequent meetings and reviews of materials, two TAC members will attend the pilot and operational meetings to observe and provide feedback on the process.

In response to previous COSDAM discussions, the project schedule was modified to account for conducting the panel meetings in person in late 2021 and early 2022. The pilot meeting will take place in Atlanta on October 25-28, 2021, and the operational meeting will take place in the same location on February 22-25, 2022. The resulting ALDs will be presented for Board discussion at the May 2022 Board meeting and Board action at the August 2022 Board meeting. The intention is for the ALDs from this project to be used in the reporting of NAEP results in fall 2022.

Project Update (August 2021)

During the May 2021 COSDAM meeting, the final Design Document for the NAEP ALD Review study was discussed by the Committee members and there were no recommended changes. Based on this approval of the study design, project staff continued to work with the TAC in reviewing procedures and materials for the study. During the TAC meetings over the past couple of months, the TAC has provided valuable feedback on the role of text complexity as part of the panelist item review process and the inferences that can be made from the ALD alignment review process. These discussions with the TAC were useful in evaluating and revising the materials for the study, to ensure that we are utilizing the most appropriate procedures to fulfill the purposes of the study.

An important part of the study process is the recruitment of panelists. The multi-phase recruitment process for this study will begin with a panelist nomination process, where individuals from different organizations and state departments of education will have the opportunity to nominate outstanding classroom and non-classroom educators to participate in the process. Materials and procedures for this nomination process were reviewed by the TAC to ensure that the recruitment process results in set of representative panelists. The nomination materials have been finalized and the nomination process will begin in July.

Next Steps

During the next few months, the project staff will finalize the meeting presentation materials and the study website and will start training facilitators in September in preparation for the October pilot meeting. The recruitment process will continue the nomination phase while also collecting information on nominated panelists, which will be used to select the set of representative panelists for the pilot study in October.

During the November COSDAM meeting, project staff will provide a brief preliminary update on the pilot study, which will be held in late October. After a more thorough review of the pilot study results, a more complete briefing will be provided during a COSDAM webinar to be scheduled for December or January.

Reporting and Dissemination Committee

July 22, 2021 3:00 – 5:00 pm <u>Zoom</u>



AGENDA

3:00 – 3:30 pm	Release Plans for Long-Term Trend and the NAEP High School Transcript Study: ACTION	Attachment A				
	Tonya Matthews, Chair					
3:30 – 4:00 pm	Review of Core Contextual Variables	Sent under separate cover				
	Jonas Bertling, ETS					
	Holly Spurlock, National Center for Education Statistics					
4:00 – 4:45 pm	State Mapping Study	Attachment B				
	Taslima Rahman, National Center for Education Statistics					
4:45 – 5:00 pm	General Updates and Next Steps on SES	See Executive				
	Marty West, Vice Chair	Committee tab for SV update				



NATIONAL ASSESSMENT GOVERNING BOARD RELEASE PLAN FOR THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP)

NAEP High School Transcript Study

Findings and data collected from the 2019 National Assessment of Educational Progress (NAEP) High School Transcript Study will be released to the public in October 2021.

BACKGROUND

In 2019, a nationally representative sample of grade 12 students took the NAEP Reading and Mathematics assessments in a nationally representative sample of America's high schools. In that year, the NAEP team requested that the sampled high schools provide transcripts for sampled students with complete transcripts, i.e., high school graduates.

The High School Transcript Study collects and reports data on the high school graduates' coursetaking patterns and rigor, credit accumulation, and grade point averages. The 2019 data can be compared to results from 1990, 1994, 1998, 2000, 2005, and 2009 transcript studies. The transcript data include demographic information on sampled graduates and can be linked to NAEP scores from 2019.

The High School Transcript Study (HSTS) bears little resemblance in scope, nature, and type of data to the reporting typical of the biennial Nation's Report Card. The transcript site follows a different layout and structure from traditional NAEP report cards. Unlike the Nation's Report Cards in which users may tailor tables, delve into the data easily through the NAEP Data Explorer, and interpret a vast array of findings readily, the complicated transcript data require more finesse. Simple questions often require complex analyses and may lead to confusion, e.g., discerning the rigor of courses from course titles, parsing standard courses from those more specialized, etc. A priority for the release of the HSTS data will be facilitating accurate and valid interpretations of the findings.

RELEASE PLAN

The release will focus on sharing results, stimulating conversation around high school coursework, and expanding the audience for these data.

This release will occur at approximately the same time when the Nation's Report Card is released biennially to cement the idea of NAEP Day in the last week of October. The release would combine a town hall approach with the feel of a moderated news talk show, e.g., C-SPAN, with an in-person component for speakers and a livestream for virtual attendees. The Board would tap its social media channels to crowdsource questions NAEP stakeholders want to know about high school graduates' schoolwork. Questions may cover high school course-taking trends, equitable access to rigorous courses, and concerns about academic preparations for postsecondary life.

The questions would be posed in a one-hour facilitated conversation that would 1) summarize HSTS results generally and 2) respond to specific questions from the field. Conversation must be emphasized; the approach should be interactive and not static. A dynamic facilitator will foster a robust conversation based on the selected questions and provide an opportunity for the National Center for Education Statistics' (NCES) Acting Commissioner Peggy Carr to share highlights from the data. Shining a spotlight on a few themes emerging from the complex data may help the audience grasp findings more easily.

A Governing Board member or two will introduce the event; secondary school principal representative Paul Gasparini has graciously agreed to participate in the release. The Board may consider inviting a few questioners to submit their queries via video.

In support of the release, the Board will produce and promote a video involving clips from interviews of high school seniors about their course-taking choices, to build interest in HSTS findings and connect the data to real life, not causally, but topically.

CENTRAL MESSAGES

Activities for the release will promote three primary messages. First, NAEP collects rich and valuable data beyond the familiar score and trend data reported in traditional report cards. Second, the HSTS is a useful resource for researchers, scholars, policy advocates, and educational administrators who seek deeper and broader context about high school graduates' experiences. Third, those who wish to understand changes in NAEP scores over time may find helpful insights from the contextual information provided by HSTS.

ACTIVITIES AFTER THE RELEASE

The Governing Board's communications contractor will work with Board staff to coordinate additional post-release communications efforts to target communities and audiences. To expose the results to a broader audience, the Governing Board will develop graphics featuring selected data and help either produce or promote a video (or graphic) about how to use and interpret the HSTS data.


NATIONAL ASSESSMENT GOVERNING BOARD RELEASE PLAN FOR THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP)

2020 Long-Term Trend

The national results of the 2020 National Assessment of Educational Progress (NAEP) Long-Term Trend (LTT) assessment for 9-year-olds and 13-year-olds will be released to the public in September 2021. Typically, this assessment includes data for 17-year-olds, however, public health restrictions due to COVID-19 prevented the administration of the assessment to that age cohort.

BACKGROUND

Beginning in the 1970s, the NAEP program has monitored student performance in mathematics and reading through the long-term trend (LTT) assessments. This assessment program within NAEP represents the nation's only continuous source of data on student achievement through a half-century. Nationally representative age-based cohorts of students participate in this assessment, which is administered through pencil and paper. The last LTT assessment was fielded in 2012 and reported in 2013.

The LTT assessment for which data will be reported this September occurred in 2020 and marked the last national assessment before schools were closed due to the COVID-19 crisis. Originally, the Governing Board expected that NCES would administer the LTT to 17-year-olds in spring 2022, to resume assessing the 17-year-old cohort who could not participate when assessment operations prematurely ceased in response to the pandemic.

However, at the August 2021 quarterly meeting of the National Assessment Governing Board, the Board will amend the current assessment schedule to administer the 9-year-old LTT assessment in the 2021-2022 school year instead. By assessing 9-year-olds immediately prior to school closures in 2020 and again this upcoming school year when the vast majority of schools will reopen with traditional full-time schedules, NAEP will capture student performance at two timepoints at the narrowest temporal boundaries of the COVID-19 potential impacts. Comparing

and contrasting scores in the 2019-2020 school year with 2022, given caveats about lack of causal inference, vastly changed circumstances in school experience, different sets of students due to changes in school enrollment, etc. may lend useful glimpses into how COVID-19 shaped student learning for the youngest participants in the LTT assessment.

RELEASE PLAN

Under the amended assessment schedule, the results from the LTT administration in the 2021-2022 school year may elicit much attention. The 2020 data to be released this September represents only 2/3 of the assessment. As such, the Governing Board will not host a virtual or inperson event to herald the release.

Instead, the Board will introduce the LTT report and propose creating and promoting one or two videos sharing and explaining the data. The video(s) could use graphics and possibly simple animation to help introduce and explain the data. Excerpts from interviews with Lesley Muldoon and Dr. Peggy Carr could provide context and highlight key findings.

The findings and the explanatory video(s) will be promoted through a press release, email announcement, social media promotion, and the monthly newsletter.

CENTRAL MESSAGES

Activities for the release will promote several messages. First, collecting and reporting data to chart a fifty-year trend line provides a helpful map to understanding general patterns in student achievement. Second, immense changes have occurred within the last fifty years, but the NAEP Long-Term Trend assessment is unique in not changing. NCES will trans-adapt this assessment to a digital platform in the coming years but strive to retain its scope and content to maintain the trend line. Third, the Board nimbly responded to the COVID-19 crisis by changing the LTT assessment schedule to produce what the Board expects will be useful insights into recent, unprecedented educational experiences.

ACTIVITIES AFTER THE RELEASE

The Governing Board's communications contractor will work with Board staff to coordinate additional post-release communications efforts to target communities and audiences through social media. The goal of these activities is to extend the life of the results and provide value and relevance to stakeholders.

Attachment B





Mapping State Proficiency Standards Onto the NAEP Scales

Since 2003, the National Center for Education Statistics (NCES) has compared each state's standard for proficient performance in grades 4 and 8 reading and mathematics by placing the state standards onto common scales from the National Assessment of Educational Progress (NAEP). The <u>report</u> released in 2021—the eighth in the series—highlights the mapping results from the 2018–2019 school year.

Since the reauthorization of the 1965 Elementary and Secondary Education Act in 2002, states have been required to define and report the percentage of students in grades 3 to 8 who meet the state performance standards in reading and mathematics. However, since each state independently develops or chooses assessments to measure its students' knowledge and skills, and sets standards for determining its students' performance, it is not possible to compare state standards unless they are placed on a common metric. Given that all states participate in NAEP for these grades and subjects, NAEP can be used as the common metric to compare the relative rigor of states' proficiency standards.

The NCES process of "state mapping," a form of equipercentile equating, shows where each state's standards fall on the NAEP scales and in relation to the NAEP achievement levels: *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*. The mapping analyses allow each state to compare the stringency of its standard of proficient performance with that of other states. Overall, in 2019, most state standards for proficient performance for both grades and both subjects mapped at the *NAEP Basic* achievement level. For both grades, more state standards mapped at the *NAEP Proficient* level in mathematics than in reading.

This study contributes to the conversation on national education policy by informing each state where its standards for proficient performance fall on the NAEP scales and allowing each state to compare its proficiency standards with those of other states. NCES mapping study results, however, are not indicative of the quality of state achievement standards; state and NAEP assessments are developed for different but related purposes and can vary in format and administration. The results of this study do not suggest that NAEP achievement levels are more valid or that states should emulate NAEP standards.

For more information, visit nces.ed.gov/nationsreportcard/studies/statemapping/.

Nominations Committee

Wednesday, July 28, 2021 5:30 – 6:30 pm (ET) Closed Session



Zoom Meeting

AGENDA

5:30 – 5:35 pm	Welcome and Agenda Overview
	Governor Jim Geringer, Chair
5:35 – 5:40 pm	Nominations for Board Terms Beginning October 1, 2021
	Lisa Stooksberry, Deputy Executive Director
5:40 – 5:55 pm	Board Vacancies for Terms Beginning October 1, 2022 Jim Geringer
5:55 – 6:20 pm	2022 Nominations Campaign
	Stephaan Harris, Assistant Director, Communications
6:20 – 6:25 pm	2022 Nominations Timeline
	Tessa Regis, Nominations Committee Liaison
6:25 – 6:30 pm	Next steps
	Jim Geringer
	Adjourn

ACTION: NAEP Reading Assessment Framework

Since its creation by Congress in 1988, the National Assessment Governing Board has overseen and set policy for NAEP, which includes determining the content and format of all NAEP assessments. The Board carries out its legislative mandate to determine the content and format of all NAEP assessments through its <u>policy</u> on Framework Development, which was revised in March 2018. The revised policy continues the Board's commitment to conducting a comprehensive, inclusive, and deliberative process to determine the content and format of all NAEP assessments, while adding details to address Board processes for framework review and updating. This commitment is met by developing framework recommendations through broadly representative framework panels and by refining these recommendations through collection of public comment.

As described in the Board policy, framework development and update processes are monitored and led by the Assessment Development Committee (ADC). In 2018, the ADC conducted a review of the current NAEP Reading Framework. In accordance with the Board policy, the ADC review included papers and discussions with an array of reading educators and experts. The ADC developed a Charge to the Reading Framework Panel that was unanimously adopted by the full Board in March 2019 (attached). The Charge included direction to develop recommendations that maximize the value of NAEP to the nation, while considering opportunities to extend the depth of measurement and reporting. In a competitive bid, the Board awarded a contract to WestEd to implement the updates to the Reading Framework.

In consultation with the ADC and Governing Board staff, WestEd selected and convened a broadly representative group of subject matter experts, practitioners, administrators, researchers, business representatives, and members of the general public – serving as the Visioning and Development Panels in accordance with Board Policy. The Development Panel met several times between November 2019 and March 2021 to develop and revise these recommendations.

The draft framework was posted for public comment from June 22 – July 23, 2020, and several webinars were held as part of the public comment process. The Board discussed and provided direction on iterative drafts of the framework during Board meetings held in July 2020, November 2020, March 2021, and May 2021.

During the May 2021 Board meeting, some Board members called for additional stakeholder engagement. In response to several outstanding concerns, Chair Haley Barbour and Vice Chair Alice Peisch convened a cross-committee Chair's Working Group consisting of Patrick Kelly, Tonya Matthews, Reginald McGregor, Marty West, Russ Whitehurst, and Carey Wright (in addition to Barbour and Peisch). The Working Group met several times in June and July 2021 to edit the framework with the goal of achieving greater consensus.

The Board released the Chair's draft to the public on June 25, 2021, and held several webinars with interested stakeholders and members of the public. A set of communications materials was developed to provide information about the Chair's draft to stakeholders. Those materials included frequently asked questions (FAQs), a comparison chart between the current and proposed reading frameworks, and a summary of the Chair's draft. Some stakeholders submitted written comments on the Chair's draft to the Board. The Chair's Working Group and the

Assessment Development Committee each met again in mid-July to review the additional stakeholder input and finalize the framework for Board action.

Board action on the 2026 NAEP Reading Framework is scheduled for August 5, 2021, as part of the quarterly Board meeting. The Assessment and Item Specifications will be reviewed throughout the fall and are scheduled for Board action during the November 2021 quarterly meeting.

Major milestones of the 2026 Reading Framework update are listed in the table below.

Milestone	Dates
ADC Framework Review	Spring/Summer 2018
ADC Framework Recommendation and Charge to the	March 2019
Visioning Panel Adopted by Governing Board	
Project Kickoff and Plan/Design Development	June – September 2019
Issues Paper and Resource Compilation Development	August – October 2019
Visioning Panel Meeting	October 2019
Development Panel Meetings	November 2019 – March 2021
Technical Advisory Committee (TAC) Meetings	2-3 weeks after each panel meeting
	and prior to submission of draft
	framework documents
Public Comment on Draft Framework	June – July 2020
Board Policy Deliberations on Draft Framework	July 2020 – May 2021
Development of Chair's Draft Framework	June – July 2021
Develop Final Versions of Framework Documents	July 2021
Board Action on Final Framework	August 2021
Board Action on Assessment and Item Specifications	November 2021

An updated FAQ document and comparison chart between the current and proposed updated frameworks are attached.

The National Assessment Governing Board Charge to the Visioning Panel For the 2025¹ National Assessment of Educational Progress (NAEP) Reading Framework

Unanimously approved on March 2, 2019

Whereas, The Nation's Report Card—also known as the National Assessment of Educational Progress (NAEP)—is mandated by Congress to conduct national assessments and report data on student academic achievement and trends in public and private elementary schools and secondary schools, and is prohibited from using any assessment to "evaluate individual students or teachers" or "to establish, require, or influence the standards, assessments, curriculum, ... or instructional practices of states or local education agencies" (Public Law 107-279);

Whereas, Congress specifically assigned the National Assessment Governing Board responsibilities to "develop assessment objectives consistent with the requirements of this [law] and test specifications that produce an assessment that is valid and reliable, and are based on relevant widely accepted professional standards";

Whereas, the Governing Board's <u>Strategic Vision</u> adopted in November 2016 established that the Board will, "develop new approaches to update NAEP subject area frameworks to support the Board's responsibility to measure evolving expectations for students, while maintaining rigorous methods that support reporting student achievement trends";

Whereas, the Governing Board established in its <u>Framework Development Policy</u> that the Board shall conduct "a comprehensive, inclusive, and deliberative process" to determine and update the content and format of all NAEP assessments;

Whereas, in accordance with the Governing Board's Framework Development Policy, the Board's Assessment Development Committee conducted a review of the current <u>NAEP Reading</u> Framework, which included <u>seven papers</u> from leading reading educators;

Whereas, based on the review of the NAEP Reading Framework conducted by the Assessment Development Committee, the Committee concludes that a substantial framework update is required to address digital platforms and new research, and recommends that the Board update the NAEP Reading Framework last updated in 2004 "to be informed by a broad, balanced, and inclusive set of factors" balancing "current curricula and instruction, research regarding cognitive development and instruction, and the nation's future needs and desirable levels of achievement," in accordance with the Framework Development Policy;

¹ The Reading Framework update initially was scheduled to be implemented in 2025 but the Congressional waiver provided in December 2020 resulted in the NAEP Reading and Mathematics assessments moving from odd years to even years on the NAEP Assessment Schedule.

Therefore,

- The National Assessment Governing Board staff, with appropriate contractor support and oversight by the Governing Board's Assessment Development Committee, shall conduct a framework update by establishing a Visioning Panel with a subset of members continuing as the Development Panel if necessary, in accordance with the Governing Board Framework Development Policy;
- All processes and procedures identified in the Governing Board Framework Development Policy shall be followed;
- The Visioning Panel will recommend necessary changes in the NAEP Reading Framework at grades 4, 8, and 12 that maximize the value of NAEP to the nation; and the Panel is 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;
- At the conclusion of the NAEP Reading Framework update process, the National Assessment Governing Board shall review recommendations from the Visioning Panel and Development Panel, if convened, and take final action on recommended updates to the reading framework, assessment specifications, and subject-specific contextual variables; and
- The framework update adopted by the Board will guide development of the 2025 NAEP Reading Assessment.



NAEP Reading Framework Updates for the 2026 Assessment FREQUENTLY ASKED QUESTIONS

July 23, 2021

The National Assessment Governing Board is an independent, nonpartisan organization whose 26 members include governors, state legislators, local and state school officials, educators, business representatives, and members of the general public. Congress created the Governing Board in 1988 to set policy for the National Assessment of Educational Progress (NAEP).

An important responsibility includes drafting and adopting frameworks, the map for each assessment's content and design. Here are highlights of our work.

Why is the Board updating the NAEP Reading Assessment Framework?

The Framework was last updated in 2004. Given NAEP's transition to digital-based administration in 2017 and changes in the testing and education landscape, it was time for an update.

How does the Board update a framework?

The iterative process takes about two years. It starts with the Governing Board's recommendation that an update is necessary. Once determined, the Governing Board convenes experts from research, policy, and practice to establish what content the assessment should cover and with what types of questions.

The experts on the Visioning Panel develop high-level recommendations for the update. Then, the experts on the Development Panel draft the framework. Next, the Governing Board offers their feedback on the draft framework. The Board also posts the draft framework in public forums to elicit feedback (i.e., through official federal channels for public comment as well as through outreach to stakeholder organizations). As part of this process, the Board conducted several webinars in Summer 2020 for a variety of audiences. In the case of this reading framework, hundreds of stakeholders submitted feedback, and an iterative process of review and revision resulted in several changes.

Most recently, the Board reviewed a revised draft and discussed outstanding issues at the May quarterly meeting. Following that meeting, the Board Chair, Haley Barbour, convened a working group of members from across the Board's standing committees who represent a cross-section of views expressed on the framework. The goal was to address remaining concerns, consider additional stakeholder feedback the Board actively sought out, and reach consensus to create a final revised draft to address remaining issues.

Date	Completed Milestone
March 2019	Board begins process to develop NAEP Reading Framework update
August – October 2019	Develop issues papers and resource compilation
October 2019	Visioning Panel convenes
November 2019	Development Panel begins process to draft recommendations
June 2020	Panel finalizes initial draft recommendations for public comment
June 22, 2020	Public comment period opens
July 2020	Webinars for state chiefs, state assessment directors, other SEA officials, and other stakeholders (e.g., district superintendents and their staff, disciplinary associations, researchers, policy organizations, etc.) ¹

The chart below highlights key milestones in the process to date.

¹ There were 8 webinars hosted by Governing Board contractor WestEd during the summer of 2020 for the following audiences: 1) stakeholders whose day-to-day work is directly affected by the Framework update (i.e., teachers, school administrators, teacher preparation). Co-hosted by the International Literacy Association (ILA); 2) stakeholders who are concerned with how reading is defined, taught, and assessed (i.e., reading/content experts, assessment experts, curriculum experts). Co-hosted by the Literacy Research Association (LRA); 3) For stakeholders who analyze or use NAEP data to inform their work (i.e., policy makers, researchers). Co-hosted by the National council on Measurement in Education (NCME); 4) Assessment Directors, Assessment English language arts (ELA) specialists, Chief Academic Officers, and ELA Collaborative. Co-hosted by the Council of Chief State School Officers (CCSSO); 5) Communication Directors, Deputies, and Chiefs. Co-hosted by the Council of Chief State School Officers whose day-to-day work is directly affected by the Framework update (i.e., teachers, school administrators, teacher preparation). Co-hosted by the National Council of Teachers of English; 7) stakeholders who are impacted by NAEP outcomes (i.e., employers, parents, general public). Co-hosted by the National School Boards Association; 8) all audiences, focusing on updates to the NAEP Reading Assessment Framework.



Date	Completed Milestone
July 23, 2020	Public comment period closes
October 2020	Review and analysis of public comment complete
November 2020	Board discusses public comment and provides feedback to the panel on draft recommendations; panel incorporates feedback into revised draft recommendations
March 2021	Board reviews and discusses revised draft recommendations; panel incorporates additional feedback into another revision
May 2021	Board reviews and discusses revised draft recommendations
May 2021	Chair Haley Barbour forms Working Group of Board members representing various viewpoints and the Board's standing committees to suggest revisions to address concerns
June 2021	Working Group meets to discuss recommendations and create an initial "Chair's Draft" of revised framework recommendations
June 2021	The Board's Assessment Development Committee, which oversees framework development, reviews draft and makes additional recommendations
June 2021	Board releases "Chair's Draft" of the revised framework recommendations
June-July 2021	Board conducts webinars and creates a special web page to collect stakeholder feedback for consideration to further revise framework
August 2021	Board is scheduled to vote on updated version of the proposed framework

Which parts of NAEP does the update change?

We see the update as an evolution from the current framework. For more information, see the <u>comparison chart</u> on the differences between the current NAEP Reading assessment and the proposed 2026 framework.

The initial proposed update more significantly departed from the current NAEP. Feedback from the Board, from the public, and from stakeholders steered the revised update to reflect more modest changes.



What about trend?

Gradual changes from the current assessment to the updated assessment (see comparison chart) should allow trend to be maintained. There is no guarantee, of course, but there is a good likelihood of preserving trend. (Maintaining NAEP's trend lines is always an empirical question and can only be confirmed after each test administration.)

The update reflects NAEP's emphasis on rigor, quality, and ability to chart trend.

Why care about trend?

The nearly thirty-year trend lines for NAEP Reading allow the public and the education field to understand how students' knowledge and skills in reading comprehension change over time. This points to where students are improving and performing well, which can help districts and states implement effective policy so all students can improve.

What's next?

The Board will vote on the final version of the 2026 NAEP Reading Framework at its August 2021 meeting. If the new framework is adopted, the Board will then consider detailed specifications to guide assessment and item development at its November 2021 meeting.

When do changes take effect?

Once approved by the Board, the update will initiate a multi-year effort that will involve development of specifications and then implementation of new item development and pilot testing before items are approved for use on NAEP. NAEP's reputation as the "gold standard" relies on a careful, methodical approach to implementing any changes to the assessments.

Throughout this assessment development process, the Board and the National Center for Education Statistics enact multiple stages of review from many stakeholders.

Finally, the update will take effect with the 2026 NAEP Reading assessment.





Similarities and Differences Between the 2019 NAEP Reading Assessment and Proposed 2026 NAEP Reading Framework

July 23, 2021

Area	2019 NAEP Reading Assessment (based on 2009 Framework)*	Proposed 2026 NAEP Reading Framework Update
Definition	 Reading is an active and complex process that involves: Understanding written text. Developing and interpreting meaning from text. Using meaning as appropriate to type of text, purpose, and situation. 	 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; Extract, construct, integrate, critique, and apply meaning in activities across a range of social and cultural contexts.
Comprehension Targets	Locate and Recall Integrate and Interpret Critique and Evaluate	Locate and Recall Integrate and Interpret Analyze and Evaluate Use and Apply
Disciplinary Contexts	Literary Text Informational Text – more general, but includes social studies, science, and other topics	Literature Contexts Social Studies Contexts Science Contexts

Similarities and Differences Between the 2019 NAEP Reading Assessment and Proposed 2026 NAEP Reading Framework – July 23, 2021

Area	2019 NAEP Reading Assessment (based on 2009 Framework)*	Proposed 2026 NAEP Reading Framework Update
Purposes	Specific purposes for each question communicated to students only for scenario-based tasks (introduced in 2019 following the transition to digital assessment)	 Broad Purposes Reading to Develop Understanding Reading to Solve Problems Specific purposes for each question communicated to students on <i>all</i> assessment tasks
Text Types	Literary Texts Informational Texts	Literature Texts Social Studies Texts Science Texts
Text Source	Authentic	Authentic except in rare instances
Text Format	 Digital texts as of 2017 Static – non-moving print, graphics, or images on screen Dynamic – navigation across modes (print, video, other) or nonlinear locations (hypertext link) 	 Digital texts Static – non-moving print, graphics, or images on screen Dynamic – expanded 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 sub-score	Language structures and vocabulary assessed No sub-score

Similarities and Differences Between the 2019 NAEP Reading Assessment and Proposed 2026 NAEP Reading Framework – July 23, 2021

Area	2019 NAEP Reading Assessment (based on 2009 Framework)*	Proposed 2026 NAEP Reading Framework Update
Universal Design Elements (UDE) – features that allow NAEP to be usable by <i>all</i> students	Tools and support features (implemented when assessment transitioned from paper to digital in 2017):	[Differences compared with current framework/assessment are listed in bold ; all others are already part of the assessment]
		Types of UDEs and possible examples:
	Task-based UDEs	Task-based UDEs
	 Highlighting and notetaking 	 Highlighting and notetaking
	 Text-to-speech on Directions and Help Screens 	 Text-to-speech on Directions and Help Screens
	 Zoom-in and selection of color schemes 	 Zoom-in and selection of color schemes
	 Sequential directions and transitions for reading collection of texts 	 Sequential directions and transitions for reading collection of texts
	 Look-back buttons to return to relevant section of text 	 Look-back buttons to return to relevant section of text
	- Graphic organizers	 Graphic organizers
	 Item foreshadowing 	 Item foreshadowing
	 Multi-part response frames 	 Multi-part response frames
	 Resetting by providing correct response to answered questions 	 Resetting by providing correct response to answered questions
		 Samples of student writing as examples
		Motivational UDEs
		 Explicit connections between broad and specific purposes
	 Task characters (avatars that act as partners in simulated settings) 	 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 	 Text providing brief topic previews
	 Limited pop-up notes for definitions of vocabulary 	 Limited pop-up notes for definitions of words or phrases

Similarities and Differences Between the 2019 NAEP Reading Assessment and Proposed 2026 NAEP Reading Framework – July 23, 2021

Area	2019 NAEP Reading Assessment (based on 2009 Framework)*	Proposed 2026 NAEP Reading Framework Update
Reporting	Overall scale score and achievement levels (NAEP Basic, NAEP Proficient, NAEP Advanced)	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	Disaggregation by all existing categories, adding:
		Disciplinary contexts
		 Socioeconomic status within race/ethnicity
		 Former English learners (ELs) as well as current ELs and non-ELs
	Data collected from student, teacher, and administrator questionnaires on contextual variables of interest	Data collected from student, teacher, and administrator questionnaires on expanded set of contextual variables
	Some data collected from students' test taking behaviors (process data)	Data collected from students' test taking behaviors (process data) on expanded set of contextual variables

Reading Framework for the 2026 National Assessment of Educational Progress

7/23/2021

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.

TABLE OF CONTENTS

TABLE OF CONTENTS	II
LIST OF EXHIBITS	IV
NAEP READING PROJECT STAFF AND PANELS	VIII
CHAPTER 1: OVERVIEW	1
Current NAEP Reading Assessment in a Digital Environment	2
Development of the 2026 NAEP Reading Framework	3
The Updated NAEP Reading Framework	4
Overview of the Updated NAEP Reading Framework's Key Components	6
Comprehension Targets	6
Other Key Components	6
Reporting 2026 NAEP Reading Assessment Results	6
Comparison of the 2009–2019 NAEP Reading Framework and the 2026 NAEP Reading	
Framework	7
CHAPTER 2: THE 2026 NAEP READING ASSESSMENT	. 10
The NAEP Definition of Reading Comprehension	. 10
Key Terminology in the Definition	. 10
Roots of the Definition	. 12
Updating the NAEP Reading Framework	. 13
The NAEP 2026 Reading Assessment and the Definition of Reading Comprehension	. 14
Comprehension Items: The Role of Comprehension Targets	. 15
Contexts and Purposes	. 17
Texts	. 19
Universal Design Elements	. 22
Contextual Variables	. 23
Summarizing the Relationship Between the Definition and Assessment Components	. 25
CHAPTER 3: DEVELOPING THE 2026 NAEP READING ASSESSMENT	. 28
Situating Readers Within Assessment Blocks	. 28
Designating Disciplinary Context	. 28
Designating a Broad Reading Purpose	. 29
Identifying Block-Specific Purposes and a Reader Role	. 30
Developing Assessment Tasks: Texts and Items	. 31
Selecting Texts	. 32
Developing Comprehension Items	. 35
Digital Assessment Features: The Role of Item Response Options, UDEs, and Process Data	. 38
Item Response Formats	. 42
Universal Design Elements (UDEs)	. 46
Process Data	. 50
Conclusion	. 52
CHAPTER 4: REPORTING NAEP 2026 RESULTS	. 53
Reporting Results	. 53
Legislative Provisions for NAEP Reporting	. 53
Achievement Levels	. 54
Reporting Results of the Updated NAEP Reading Assessment	. 54
Reporting Categories	. 55
Reporting by Disciplinary Contexts	. 55

Disaggregating Results Within Demographic Categories	55
Expanding Reporting Categories for English Learners	56
Contextual Variables	57
Reader Characteristics	58
Environmental Characteristics	59
Data Sources	60
Enhancing NAEP's Reporting Capacity	61
Conclusion	63
GLOSSARY	64
APPENDIX A: ADDITIONAL ASSESSMENT DESIGN FEATURES	70
APPENDIX B: ACHIEVEMENT LEVEL DESCRIPTIONS	80
NAEP Policy Definitions	80
Range ALDs	80
Organizational Features and Structures of the Reading Construct: Contexts, Purposes,	
Comprehension Targets, and Text Complexity	81
Comprehension Targets and Text Complexity	81
Broad and Specific Reading Purposes	81
NAEP Reading Achievement Levels: Grade 4	82
NAEP Basic	82
NAEP Proficient	83
NAEP Advanced	84
NAEP Reading Achievement Levels: Grade 8	85
NAEP Basic	85
NAEP Proficient	85
NAEP Advanced	86
NAEP Reading Achievement Levels: Grade 12	87
NAEP Basic	87
NAEP Proficient	88
NAEP Advanced	89
APPENDIX C: CONSIDERATIONS AND EXAMPLES FOR DEVELOPING BLOCKS	90
Employing the 2026 NAEP Reading Assessment Framework Principles: Assessment	
Components	90
Considering the Range of Variations Within Assessment Components and Across a Bloc	ck 91
Specific Guidelines for Block Development	94
Block Sketches	94
Hana Hashimoto, Sixth Violin, Grade	97
Hill District, Grade 1	106
E. B. White	121
APPENDIX D: ADDITIONAL EXAMPLES OF RESPONSE FORMATS AND COLLECTION OF PRO	CESS
DATA	133
APPENDIX E: ADDITIONAL EXAMPLES OF READING PURPOSES AND UDES	141
References	147

LIST OF EXHIBITS

Exhibit 1.1. Similarities and Differences Between the 2009–2019 and 2026 NAEP Reading
Frameworks7
Exhibit 2.1. Relationships Between the NAEP Definition of Reading Comprehension Definition
and the NAEP Reading Assessment
Exhibit 3.1. Principle and Provisional Distribution Targets for Sampling Disciplinary Contexts
Exhibit 3.2 Principle and Provisional Distribution Targets for Sampling Broad Reading
Purposes by Grade Level
Exhibit 3.3. Task-specific purposes presented at the beginning of a Grade 4 Reading to Develop
Understanding block using the text <i>Hana Hashimoto</i> , <i>Sixth Violin</i> (a short story) by
Chieri Uegaki and Qin Leng
Exhibit 3.4. Example of multiple texts readers encounter as part of one task on the ePIRLS
(2016) Grade 4 reading assessment
Exhibit 3.5. Seven Principles of Universal Design of Assessments (UDA)
Exhibit 3.6. Alignment of the 2026 NAEP Reading Assessment with Principles of Universal
Design of Assessments (UDA)
Exhibit 3.7. Example of Matching Response Format from PARCC Grade 8 Literature
Exhibit 3.8. Example of Grid Response Format from PISA
Exhibit 3.9. Flexible Distributions of Item Response Types Across Grade Level
Exhibit 3.10. Example of a Dynamic Search Engine Item from ePIRLS 2016 for Grade 4
Students
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
Exhibit 3.12. Teacher and student task characters remind the reader of the task goal for the
Second task
Exhibit 3.13. Example of Two Informational UDEs from NAEP's "Five Boiled Eggs" Block 50
Exhibit 5.14. Example of a Constructed Response field from ePIRLS 2010 for Grade 4 that
Conects Navigational Process Data. The Space Camp image and blast off button
who click on irrelevant details (i.e., advertisements) on a webpage rather than
attending to the comprehension item at hand
Exhibit 4.1 Generic NAEP achievement levels
Exhibit 4.2 Contextual Variables 61
Exhibit 4.1. Principle and Provisional Distribution Targets for Sampling Assessment Design
Elements: Text Formats and Modes 70
Exhibit A.2. Illustrative Examples of Texts and Other Media Across Single Static and Dynamic
Texts and Multilavered Digital Text Environments
Exhibit A.3. Commissioned Texts: Parameters and Constraints
Exhibit A.4. Passage Lengths for Grades 4, 8, and 12
Exhibit A.5. Typical Text Elements Across Disciplinary Contexts
Exhibit A.6. Text Structures and Features Within and Across Single Static and Dynamic Texts
and Multilayered Digital Text Environments
Exhibit A.7. Distribution of Cognitive Comprehension Targets Across Grade Level and Broad
Purposes

Exhibit A.8. Incl	lusion and Exclusion Criteria for Connected Language and Vocabulary
Exhibit C.1. Des	sign Components of a 2026 NAEP Reading Assessment Block
Exhibit C.2. Cor	ntinuum of Variation in Features of Assessment Components Within a Block 92
Exhibit C.3. Und	derlying Continuum of Variation in Assessment Components in the Block
Dea	sign for E.B. White, Hana, and Hill District Block Sketches
Exhibit C.4. Blo	ck Design for Hana
Exhibit C.5. Spe	cific purpose, reader role, and task characters serve to situate readers in a Grade
4 R	Reading to Develop Understanding block involving the short story Hana
Ha	shimoto, Sixth Violin by Chieri Uegaki and Qin Leng
Exhibit C.6. A C	Grade 4 RDU block illustrating a Locate and Recall multiple choice item. The
tea	cher 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 C	Grade 4 Locate and Recall item illustrating a fill in the blank short constructed
res	ponse item 100
Exhibit C.8. A C	Grade 4 Locate and Recall item illustrating a multiple select multiple choice
res	ponse format
Exhibit C.9. A C	Grade 4 Analyze and Evaluate short constructed-response item illustrating a task-
bas	ed UDE in the form of a look-back button that refers readers to the relevant
sec	tion of text
Exhibit C.10. Th	ne items for the first task help students develop an understanding of the events
and	l characters as in this Grade 4 Integrate and Interpret short constructed response
iter	n 102
Exhibit C.11. A	Grade 4 Integrate and Interpret Item for the first task using a single select
mu	Itiple choice format
Exhibit C.12. Te	eacher and student task characters remind readers of the second task goal in this
Inte	egrate and Interpret item
Exhibit C.13. Th	he student's responses from their completion of the previous item are carried over
to t	he next item as the completed notes. A graphic organizer with drag and drop
fea	tures 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
Inte	egrate and Interpret item
Exhibit C.14. Th	his 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-
res	ponse item format. Students use what they have learned from the text about Hana
as a	a person and apply that understanding to draw a conclusion about the kind of
per	son she is
Exhibit C.15. BI	ock Design for Hill District Sketch
Exhibit C.16. A	social studies context and reader role serve to situate readers in a Grade 12
Rea Ehihit C 17 S.	ading to Solve A Problem block involving several interconnected digital texts 10/
Exhibit C.17. Sa	ime-aged task characters and a task-based UDE in the form of four task-specific
pur	poses serve to guide and motivate readers in the KSP block
	Grade 12 KSP block illustrating the directions that readers are asked to follow as
the	y engage with texts and items. The task character reminds the reader of the
Spe	Grade 12 Locate and Recall item illustrating a multiple selection multiple shoirs.
EAIIDIL C.19. A	nonse format
ies	ponse ionna

Exhibit C.20. A Grade 12 Locate and Recall item illustrating a single-select multiple choice item
response format110
Exhibit C.21. Two Grade 12 items that ask readers to Integrate and Interpret (item 1) and Locate
and Recall (item 2) textual and visual information from an artist's rendering of the
site improvement plan published on a website
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
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
Exhibit C.24. A Grade 12 short constructed-response item that asks readers to integrate and
interpret information on a web page with a look-back button (task-based UDE). The
task character reminds readers of the specific purpose of the second task
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 114
Exhibit C 26. A Grade 12 item selected response zone item designed to capture process data
about how readers navigate through hyperlinked web pages 114
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
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
Exhibit C.29. A Grade 12 multiple-select response grid item with a task-based UDE (table) and
motivational UDE (task character) that serve to support readers as they engage in
one particular item in the RSP block
Exhibit 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 118
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 120
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
Exhibit C.35. Introduction to E. B. White
Exhibit C.36. Introduction to the grade 8 E. B. White literature block 123
Exhibit C.37. Task 1 would involve additional items 124
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 125
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 125
Exhibit C.40. Additional items accompany task 2 126
Exhibit C.41. An Integrate and Interpret item illustrating a matching item response format 127
Exhibit C.42. Integrate and Interpret item illustrating resetting of item responses from prior item
127
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
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
Exhibit C.45. The Second Text for the E. B. White Task: An Essay from the <i>New Yorker</i> 131
Exhibit D.1. Example of a Matching Selected Response Item for a Webpage Text from PISA's
Rapa Nui Block
Exhibit D.2. Example of a Matching Selected Response Item from a Grade 12 PARCC Block 134
Exhibit D.3. Example of a Zones Selected Response Item Format and the Use of Task Characters
from ePIRLS' Mars Block
Exhibit D.4. Example of a Grid Selected Response Item from PISA's Rapa Nui Block
Exhibit D.5. Example of a Zones Item for an Internet Text from ePIRLS "Elizabeth Blackwell"
$Block \dots 13/$
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 web Page
Exhibit D./. Example of a Short Constructed Response Item from PISA's Galapagos Islands
BIOCK
Exhibit E.I. Example of a Specific Reading Purpose and a informational UDE from PISA's
Exhibit E 2 Example of a Task based LIDE from the Smorter Dalanced Itams Dublished by The
Exhibit E.2. Example of a Task-based ODE from the Smarter Balanced items Published by The
Exhibit E 2 Example of a Motivational LIDE from NAEP's "Touch as Daisy" Plock
Exhibit E.5. Example of a Wollvational UDE, from NAEP's Tough as Daisy Block
Exhibit E.5. Two Examples of Informational UDEs non interaction of Dessage Introductions from a
Exhibit E.J. 1 we Examples of informational ODEs in the Porn of Passage inforductions from a Released NAED 2010 Block on E. P. White
Exhibit E 6 Example of Three Informational LIDEs in the Form of Dessage Introductions from
the Michigan Student Test of Educational Drogress
the whenigan student rest of Educational Progress 140

Visioning Panel

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

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

Carolyn Aguirre

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

Sarah Aguirre*

Field Education Specialist University of Texas, San Antonio San Antonio, TX

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

Nancy Brynelson*

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

Jinghong Cai

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

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

Byeong-Young Cho*

Associate Professor, Korean Language Education Hanyang University Seoul, Republic of Korea

Julie Coiro*

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, Learning and Professional Services American College Testing (ACT) Woodstock, MD

Robin Hall

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

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

Christy Howard

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

Panayiota Kendeou

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

Emily Kirkpatrick

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

Carol Lee*

Edwina S. Tarry Professor, Education and Social Policy Northwestern University Evanston, IL

Karen Malone

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

Marina Pacheco*

Associate Professor, Curriculum and Instruction University of Wisconsin, Madison Madison, WI

Cindy Parker

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

Jim Patterson

Executive Director The College Board Coralville, IA

P. David Pearson, Panel Chair*

Professor Emeritus, Education University of California, Berkeley Berkeley, CA

Sue Pimentel

Founding Partner Student Achievement Partners Fort Myers, FL

Alicia Ross*

Teacher / Educational Consultant Blue Ridge High School 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 New York, NY

Joan Herman

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

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 that will develop the NAEP Reading Assessment. In accordance with Governing Board policy, the 2026 NAEP Reading Framework focuses on "important, measurable indicators of student achievement to inform the nation about what students know and are able to do without endorsing or advocating a particular instructional approach."

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

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), National Assessment of Educational Progress Reauthorization Act (NAEPRA) of 2002, P.L. 107–279). The NAEP reading 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 2018). 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 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. In addition, a text-to-speech capability is available on the Directions and Help screens (but not available for the reading passages or questions). Texts or questions may include hyperlinks, such as pop-up notes to click for more information (typically a definition of a selected word), a look-back button that takes students back to the relevant sentence or location in the text, multi-part response frames, and more. Not all 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 Board charged the Visioning and Development Panels with developing recommendations for updating the framework as follows:

The Visioning and Development Panels will recommend to the Board necessary changes in the NAEP Reading Framework at grades 4, 8, and 12 that maximize the value of NAEP to the nation. The panels are also tasked with considering opportunities to extend the depth of measurement and reporting given the affordances of digital based assessment. The update process shall result in three documents: a recommended framework, assessment and item specifications, and recommendations for contextual variables that relate to student achievement in reading.

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

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

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 (Rogers, Lazarus & Thurlow, 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 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 U.S.

Overview of the Updated NAEP Reading Framework's Key Components

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

Comprehension Targets

Like its predecessors, the 2026 NAEP Reading Assessment engages students in reading texts and responding to questions that assess their comprehension of these texts. 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 Universal Design Elements 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 (1) socioeconomic status within race/ethnicity, whenever feasible¹ and (2) former English (ELs) learners 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: (1) cognition and metacognition and (2) engagement and motivation; and by two sets of environmental characteristics: (1) reports of school and community resources and (2) 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. Building from this previous framework and on digital innovations, updates include consideration of three additional, research-based concepts: (1) how social and cultural experiences shape learning and development; (2) how reading varies across disciplines; and (3) the increasing use of digital and multimodal texts.

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

	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

Exhibit 1.1. Simila	arities and Differences	Between the 2	2009–2019 and	2026 NAEP	Reading
Fram	eworks				

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 – non-moving print, graphics, or images on screen Dynamic – navigation across modes (print, video, other) or nonlinear locations (hypertext link) 	 Digital texts Static – non-moving print, graphics, or images on screen Expanded use of dynamic formats – navigation across modes (print, video, other) or nonlinear locations (hypertext link) 	
Text Complexity	 Determined by: Expert judgment Passage length Two or more research-based readability measures 	 Determined by: Expert judgment Passage length Quantitative and qualitative research- based complexity measures 	
Language Structures and Vocabulary	Vocabulary assessed Potential for subscore	Language structures and vocabulary assessed No subscore	
Universal Design Elements (UDE)	 Digitally-based assessment as of 2017 includes tools and support features: Highlighting and notetaking Text-to-speech on Directions and Help screens Zoom-in and selection of color schemes Sequential directions and transitions Look-back buttons to return to relevant section of text Graphic organizers Item foreshadowing Multi-part response frames Purpose statements Task characters (avatars that act as partners in simulated settings) 	 Types of UDEs and possible examples: Task-based UDEs Highlighting and notetaking Text-to-speech on Directions and Help Screens Zoom-in and selection of color schemes Sequential directions and transitions for reading collection of texts Look-back buttons to return to relevant section of text Graphic organizers Item foreshadowing Multi-part response frames Samples of student writing as examples 	

	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 Disciplinary contexts Socioeconomic status within race/ ethnicity, whenever feasible. Former English learners (ELs) as well as current ELs and non-ELs Data collected from student, teacher, and administrator questionnaires on expanded set of contextual variables Data collected from students' test taking behaviors (process data) on expanded set of contextual variables

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

- Chapter 2 presents the **2026 NAEP Reading Assessment**, including the definition of reading comprehension and major assessment components.
- Chapter 3 describes the **Development of the 2026 NAEP Reading Assessment**, including specific design elements.
- Chapter 4 explains the **Reporting of NAEP 2026 Results**, including the expansion of reporting categories, contextual variables, and reporting capacity.
The 2026 NAEP Reading Framework recommends updates necessary to deliver assessments that are relevant, fair, and valid measures of student achievement in the U.S. The 2026 Framework builds on the current NAEP framework and operational assessment, especially the advances made possible by digitally-based assessment, by drawing on current understandings of reading comprehension and assessment. Chapter 2 provides a detailed description of the components that will be included in NAEP Reading assessments that students will take beginning in 2026. The chapter begins with the 2026 NAEP Definition of Reading Comprehension, 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 to understand how readers make meaning in the presence of texts.

Contexts. Reading comprehension is a 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, first graders will use their knowledge of the stories they have listened to at home and in daycare settings to understand the stories they now have to read on their own. Adolescents in the U.S. 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 more challenging to 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, Radach, Mayer, & Lonigan, 2013). While these activities enable comprehension, they do not provide direct evidence of comprehension; thus, they are not directly assessed in the NAEP Reading Assessment.

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

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, Singer, & Trabasso, 1994; Kendeou, Van den Broek, Helder, & Karlsson, 2014; Kintsch, 1998; McCarthy & McNamara, 2021; Pearson & Cervetti, 2015; van den Broek, Risden, Fletcher, & Thurlow, 1996). These models represent the relationship between knowledge and comprehension as one in which existing knowledge is continually activated and integrated with textual information as readers develop a propositional understanding and, ultimately, a coherent mental representation of the text. Moreover, a large body of research has documented the impact of readers' topic knowledge and domain knowledge on reading comprehension across grade levels and text genres (e.g., Alexander, Kulikowich, & Schulze, 1994; Cervetti & Wright, 2020; Kendeou & O'Brien, 2016; Pearson, Hansen, & Gordon, 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, Risden, Fletcher, Thurlow, Britton, & Graesser, 1996) are still regarded as the most valid and useful cognitive accounts of reading comprehension. These models emphasize the multiple levels of meaning readers create, including a representation of the surface form that reflects accurate decoding; a text-base that includes all of the key ideas in the text plus the text-based inferences that link ideas within texts; and a situation model that represents the integrative links readers make between ideas expressed in the text and the knowledge they bring to reading.

Although earlier NAEP Reading frameworks were grounded in cognitive models of comprehension, they also acknowledged the importance of readers' purposes and the contexts in which they read and learned to read. In the first Reading Framework published in 1992, reading comprehension was defined as "... a complex process that involves an interaction among the reader, the text, and the context in which something is read" (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, 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, where 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, 2019, p. 3). The 2026 NAEP Reading Assessment will continue NAEP's longstanding focus on reading comprehension, rather than foundational skills or writing.

Updating the NAEP Reading Framework

The 2026 NAEP Reading Framework is updated to reflect three research-based developments that help 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, Becker, Rowe, & Pearson, 2016; Heath, 1983; Lee, 2017; Scribner & Cole, 1981; Smagorinsky, 2001; Street, 1984), including disciplinary communities (Goldman, et al, 2016; Moje, 2007). 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, p. 33).

This NASEM finding is also reflected in other large-scale assessments. PIRLS, the international assessment of reading for fourth grade students, notes that "social interactions about reading in one or more communities of readers can be instrumental in helping young students gain an understanding and appreciation of texts and other sources of information" (Mullis &

Marten, 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, 2007; Shanahan & Shanahan, 2008; Snow, 2010). These differences, which are related to the core activities in each discipline, require readers to employ different resources as they read and respond to text.

Also newly explicit in the 2026 Framework is recognition of the multimodal nature of texts used across all aspects of society. The widespread presence and rapid evolution of computers, smart devices, and software platforms have changed society's ideas about what counts as text and its uses. Students read digital/multimodal texts in and out of school. Even though there is a common thread to reading in print and multimodal texts, there are also substantial differences, particularly around navigation (Coiro, 2020; Hartman, Morsink, & Zheng, 2010; Serafini & Gee, 2017). The implication is that the NAEP Reading Assessment must sample multiple modes of text.

These updates allow the 2026 NAEP Reading Framework to account 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 added 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. These kinds of processing help 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, Stadtler, & Salmerón, 2018; 2020; Meola, 2004; Ostenson, 2014; Wineburg, 1991; Wineberg & McGrew, 2017). For example, readers might use information appearing in one text as the basis for evaluating the ideas or the use of language in a second text.

Use and Apply. The final 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, Palincsar, Biancarosa, and Berman, 2020). This set of targets reflects contemporary

understandings that comprehension may involve a series of processes that culminate in readers taking some kind of action in the world outside of text. As they engage in Use and Apply processes, readers must consider how to reframe ideas from their reading and experiences to create a new product for a specific purpose and audience (Marzano, 1988). As readers reflect on how to respond to items that require such processes, they take into account the reading purposes, the norms established by genre and disciplinary conventions, as well as 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 nonfiction narratives, and criticism, provides opportunities for enjoyment and for reflection and analysis around these themes, including how they shed light on their own experiences and social worlds. Literature also often provides opportunities to connect with cultures and experiences similar to or different from one's own, extending readers' understandings about the world. 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 otherauthors and literary traditions; and to situate problems in contemporary and historical contexts.

Science Contexts. Science contexts are primarily focused on observing and explaining the natural world. Although these scientific activities do not depend exclusively on reading, texts play an important role in learning about and communicating science ideas in school and non-school settings. Learning the concepts and processes of science in school involves the use of varied texts to describe, report, and articulate claims about the natural world (e.g., textbooks) and to record systematic efforts to act upon it (e.g., observation protocols, lab notes, experimental descriptions, journal articles). Outside of schools, individuals often access scientific information (e.g., in newspapers and on internet sites) needed to understand issues and solve problems. Moreover, the application of reading to understanding and acting upon the natural world calls on an array of reading strategies, as well as understandings about how scientists determine findings and what constitutes credible evidence for those findings.

Social Studies Contexts. Social studies includes history, geography, cultural studies, civics, and government, with less common 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 to create a product. For example, readers might be asked to use information across four different short texts to develop an argument for or against a city ordinance requiring bicycle lanes on all city streets with a certain traffic load.

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

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

Texts

Because texts are central to the NAEP Definition of Reading Comprehension, the 2026 NAEP Reading Framework recommends sampling from the large domain of texts that fourth, eighth, and twelfth graders are likely to encounter in school and non-school settings, as is described in more detail in the 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 non-literal interpretations (Appleman, 2017; Lee, Goldman, Levine, & Magliano, 2016; Rabinowitz, 1987).

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

Social Studies Texts. NAEP will also sample from the varied forms of texts common to the social studies. Selection will represent a wide array of text types, forms of representation, sources of information, and perspectives. These texts document human activity across cultures, societies, and time periods. They include newspaper articles, diaries, letters, speeches, records of sale, advertisements, official government documents, photographs, cartoons, maps, artwork, music, and video and audio recordings. They also include interpretive books and articles about events, time periods, or people, and classroom textbooks. Social studies texts may organize ideas chronologically or thematically to represent time periods, social structures, continuity and change, cause and consequence, and varied social or historical perspectives to consider how the

past influences the present (Charap, 2015; Seixas, 2010; Seixas, Gibson & Ercikan, 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 and PARCC). Like reading in many of today's classrooms, these assessments include print texts paired with audio clips, podcasts, infographics, and video segments. Even states that moved away from the CCSS and consortium assessments have retained standards and assessments that acknowledge widespread use of digital texts in homes, schools, and communities. Digital platforms offer a range of affordances in measuring reading comprehension (Coiro, 2020; Fitzgerald, Higgs, & Palincsar, 2020).

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

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

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

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

Universal Design Elements

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

UDEs are design elements of the assessment environment intended to help all test-takers access, organize, analyze, and express ideas when engaging in complex tasks, such as reading comprehension (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: taskbased, 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; Dejong, 2006; Zhang & Quintana, 2012). They maximize the likelihood that readers are able to cognitively engage with complex NAEP-designed reading experiences within the compressed time frame of an assessment. They might include a sequential set of directions to communicate expectations for how and why readers should engage with a collection of texts; they can also help readers plan

and monitor their work across multiple texts and tasks (de Jong, 2006). They might also include graphic organizers that allow readers to record and revisit their ideas, reduce time spent on searching and scrolling, and, thus, provide more time for students to read, evaluate, and engage with text content. These UDEs might also include simulated student work examples that offer models of approaches to tasks before students complete similar tasks independently (e.g., Sparks & Deane, 2014). 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, 2016). They may also maintain readers' interest by communicating explicit connections between the broader purpose for completing a task and the sub-tasks that need to be completed along the way. UDEs in the form of task characters provide written and/or oral directions or serve as experts or peers to provide information or moral support. Task characters may also serve as a simulated target audience with whom readers can communicate new understandings about what they have read and learned (e.g., Use and Apply).

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 but rather they provide orientations to topics, concepts, or obscure vocabulary that students may need to make meaning from text as they read (Kintsch, 1998; McNamara, 2021; 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. Universal Design Elements in the 2026 NAEP Reading Assessment reflect the NAEP Definition of Reading Comprehension in several ways. UDEs enable readers to engage with topics to be read about by providing brief previews and offering instructions on how to complete assessment tasks. They also include lookback buttons and definitions of some words (only those not measured on the assessment), thus reflecting the kinds of navigational aids and tools available in typical reading situations. In addition, UDEs clarify the nature and order of tasks and expected responses. 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 status², 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 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 <u>Chapter 4</u>.)

The digital format, which has been implemented starting in 2017, also allows NAEP to capture students' time on tasks and navigational moves as they complete the assessment. The process data now available because of the data-gathering assets of the digital platform can provide information about student journeys through the texts, directions, UDEs, and items students traverse during the assessment. From these data, NAEP can construct indicators about how students direct their attention (including moment-by-moment shifts in focus) and how long (or how little) they linger on different segments of the texts, the items, the UDEs, or the directions. These indicators can be used to help interpret performance 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.

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 National Center for Education Statistics are currently considering refinements of this approach that may affect the operationalization of socioeconomic status under the 2026 Framework.

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.

	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.

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

	Features	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, such as providing definitions of obscure words not measured on the assessment and offering lookback buttons.	Provide information that clarifies the nature and order of tasks and expected responses.	
Contextual Variables 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 that were used in previous NAEP Reading assessments (National Assessment Governing Board, 2019). These practices include incrementally augmenting current assessment design with features that are carefully tested and refined over time: a hallmark of NAEP development practices since the inception of the assessment.

The chapter is divided into three sections. The first section provides an overview of considerations related to developing block components of the 2026 NAEP Reading Assessment. This involves situating readers within a disciplinary context, a broad purpose, and a specific purpose and role for each block. The second section discusses the task components and how they can be used to expand the ways in which readers are asked to demonstrate their ability to engage in the comprehension processes outlined in <u>Chapter 2</u>. Task components include texts and comprehension items. The third section details considerations for leveraging digital assessment features, including item response formats, Universal Design Elements (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 session, test-takers engage in two grade appropriate blocks. The design of every block involves situating readers within a *disciplinary context*, a *broad purpose for reading*, and a *specific purpose* and *role* for the reader working through the block. See Exhibit 2 in Appendix C, which illustrates a range of design features that should be considered when designing assessment components. These features vary along a continuum within a block, from less to more dynamic and cumulative.

Designating Disciplinary Context

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

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

Principle for Sampling Disciplinary Contexts : The percentage of Literature decreases across grades as the 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. Reading to Solve a Problem activities entail developing understanding, or comprehending text, but in the service of using this understanding to take a specific action or create a product, such as a written explanation or a classroom presentation.

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

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

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

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

Specific purposes and reader roles are explicitly shared with test-takers as part of the directions at one or more locations in the block. Exhibit 3.3 depicts an example of what readers might see when they begin the Grade 4 Reading to Develop Understanding sample block in a literature context (see Appendix C). In this block, readers are invited to participate in a book discussion group about the short story *Hana Hashimoto, Sixth Violin* by Chieri Uegaki and Qin Leng (2014) with three other fourth grade student task characters (simulated avatar classmates). In addition to reading directions about the discussion goal, students are told they will read 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 disciplinespecific blocks for which both reading purpose and reader role are transparent to better simulate the situations in which most readers find themselves in school, workplace, and community situations.

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



Developing Assessment Tasks: Texts and Items

After readers are situated in the assessment block, they encounter two or more tasks, each with its own specific purpose. A task is a subunit within each block on the 2026 NAEP Reading Assessment. Each NAEP reading block has 2-3 tasks, one or more texts, and related comprehension items. Developers take into consideration time, total passage length, and grade appropriateness when determining the number of texts in each assessment block. Extended pieces of literature or a full argumentative essay might result in only one text with one or two tasks. Shorter texts such as a haiku poem, photograph, search engine result, or 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 Mars' surface, and a dynamic image of two planets spinning around the sun. The item is intended to assess fourth graders' understanding of how to use embedded hyperlinks to locate and recall important information about the passage.

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



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

- *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 non-fiction texts, and especially those in a modality other than print, include brief and purposeful topic introductions where appropriate?
- *Complexity*. Are the language features (vocabulary, syntax, discourse and rhetorical structures) representative of the specific grade and disciplinary context?

Several of these text selection criteria are elaborated in the following sections with a number of principles and design considerations.

Authenticity. Most texts included in NAEP Reading will be presented in their entirety, as students would typically encounter them. However, some texts may be excerpted from, 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 a RSP purpose. In the exceptional cases in which commissioned writing may be required, it should follow the text selection criteria applied to authentic texts. In very rare cases, then, commissioned texts may be used as part of a set of texts. Thus, such commissioned texts will not serve as the main, or anchor, text for a text set, nor will students be asked items focused on evaluating the credibility or accuracy of such texts. See Exhibit 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, Wasik, and Turner 1991).

Reflecting classroom practice, students in earlier grades generally read shorter texts while older students read longer texts. It is expected that in some cases, two or more texts (with static and/or dynamic textual features) will be used together to assess students' ability to compare, synthesize, and critique texts in terms of their content, themes, and stylistic features. In these cases, the total number of words will reflect the recommended passage length range for each grade.

Because 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 <u>Chapter 2</u>. 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 <u>Chapter 2</u> 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 details³.

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, complexity and "considerateness" (including both qualitative and quantitative measures; e.g., conventional readability criteria, reader-text connections, language structures and vocabulary considerations; Anderson & Armbruster, 1985) and reflect contemporary standards applied to digital texts and other contemporary media forms. Because readers use specific knowledge to identify important information in different types of texts, developers attend to variations in organization and cohesion in line with 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, Wixson & Pearson, 2014). Textual ideas in disciplinary contexts should be represented with appropriate vocabulary and, where needed, texts should have useful supplemental explanatory features such as definitions of technical terms or orthographic features (italics, bold print, headings) and connective signal words (e.g., first, next, because, however). Unfamiliar concepts should be defined with examples provided. Designers should aim for a flexible and diverse representation of language and structures across the blocks.

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

³ This document will be presented for Board action later in 2021

A potential difference between traditional and digital texts is the nature of text arrangement and the means with which readers navigate through and across texts (Cho, 2014). In selecting digital texts, it is important to attend to the features that allow for navigating 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 second), embedded video and links to other sources of information. Thus, it is important to determine that the ideas, perspectives and modes presented in digital media reflect what readers encounter in their academic and everyday lives.

Engaging experts in selecting texts that reflect authentic social and cultural traditions in a range of disciplinary contexts 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:

- *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 <u>Chapter 2</u>, 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 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 at various types of representation [e.g. directly stated in prose, embedded in a visual representation, or implied through symbolism] and across different locations in the text). On the other hand, items should not assess knowledge sources irrelevant to the items and associated comprehension targets in a given block. For example, items should not 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 targets. The single most important standard that the 2026 NAEP Reading Assessment will meet is asking questions about matters of substance in the texts. <u>Chapter 2</u> contains examples of what test items might ask readers to do with respect to each of the four comprehension targets.

Exhibit A.7 <u>in Appendix A</u> presents guidelines for distributing items mapped to comprehension targets across grade level and blocks. These flexible distributions allow for the possibility of varying the number of items for each target depending on block type. One broad principle is that the percentage of items designed to assess Integrate and Interpret or Analyze and Evaluate ideas increases across grades. In addition, in Reading to Solve a Problem (RSP) blocks, the percentage of items designed to assess Locate and Recall ideas decreases across grades as the percentage of Use and Apply ideas increases. Finally, the distribution targets should never outweigh the other principles in the bulleted list. In other words, for a given text, it is better to fall one item short in the number of items for a target than it is to include one 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; that is, the *type* of inference (a local, straightforward inference within a paragraph vs. a global inference across ideas in a text) combined with the *number* (one or multiple) and the *distance* of these inferences (within one text, across two texts, or beyond the text). 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 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 gradeappropriate 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 & Interpret) and Language Structures and Vocabulary.

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

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, Johnstone, & Thurlow, 2002, p. 6). This does not mean that the test will be less rigorous or that content should be altered. Rather, this is achieved through accessibility of content using diverse formats (e.g., item formats), technological tools (e.g., Universal Design Elements, or UDEs), and designs that include diverse test-takers.
2. Precisely Defined Constructs	Precisely defined constructs help to ensure that an assessment measures the construct it intends to measure rather than aspects not part of that construct, which creates construct-irrelevant variance. Without a precisely defined construct, it is hard to know whether items and other design features work towards measuring the intended construct or whether they might, in fact, be measuring something else.
3. Accessible, Non-biased Items	The purpose of this principle is to ensure that all test takers can access the content being assessed so that items measure the same construct for all

	Exhibit 3.5. Seven I	Principles of 1	Universal Design	of Assessments (UDA
--	----------------------	-----------------	------------------	---------------------

	students who take the assessment (i.e., items are "non-biased"). For example, if a passage contains a highly culturally-situated term that might be more familiar to some sub-populations of test takers (e.g., to boys more than to girls), this might 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	<i>Inclusive Population Assessed in NAEP Reading:</i> NAEP Reading aims to measure <i>reading comprehension</i> in a way that represents <i>all</i> students within the U.S. population at grades 4, 8, and 12 by not excluding any groups from sampling.
	<i>UDEs</i> 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, as readers would be during out-of-test reading activities. While assessments can never perfectly measure the constructs they intend to measure as those constructs exist in reality, assessments aim to do so to the extent possible (i.e., what is referred to as ecological validity). In so doing, this also supports construct validity, in alignment with the "precisely defined constructs" called for in UDA Principle 2. Situating the reader to access the content being measured because it activates 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 understandings relevant to those disciplinary contexts and 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.
	<i>Item Formats:</i> Thoughtful selection of item formats to measure particular comprehension targets within the context of the texts and specific purposes supports students' access to the test construct because they are able to focus limited cognitive resources on tasks aimed to measure the construct. This supports the assessment's ability to measure the construct it intends to measure (Principle 2) by facilitating <i>all</i> students' ability to access the construct (Principle 3).
3. Accessible, Non- biased Items	Regular NAEP Reading Research and Development Process: Item bias is tested through NAEP's regular item review and pilot testing procedures to ensure that items are not more or less difficult for students from particular subpopulations. To test item bias, the difficulty of items across different subpopulations of students (e.g., boys and girls) is compared to ensure that items measure the same construct across groups. Biased items are revised until they no longer demonstrate bias.
	<i>Disciplinary Contexts & Purposes:</i> 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.
	<i>Range of Texts and Tasks Represented:</i> 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	<i>UDEs and Item Formats:</i> 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	<i>Instructions:</i> 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.

	<i>Clear Comprehension Items and Tasks:</i> Similarly, items written using simple, clear language that is easily understandable regardless of a student's experience, knowledge, language use, interest support the student's ability to engage in the items that are measuring reading comprehension ability aligned to the comprehension targets.
	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	<i>Selection of Grade-Appropriate Texts:</i> 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	<i>Visual Layout:</i> 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 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

Chicken Forum Released Item #3

Chicken Forum		- 1	www.cnc.enneatri.	communicaspinin-chicken	5
Question 3.17 Refer to the Chicken Health Forum on the right. Click on the holoes in the fable to answer the guestion			Chicken Health		- -
			About	Forum	Pictures
tome marks on a forcing can be solariant to it	he lonis util	-	Giving Aspirin to Chickens		
ome posts on a forum can be relevant to the topic while ome posts are not. Gick on either Yes or No to indicate the bests in the table below are relevant to			Ivana SB THREAD ST	ARTER	Posted 28 October 18.12
vana_88's problem.	CTUIL IC		Helio everyone!	en? She is 2 years old an	d I think she hud ber leg.)
is the post relevant to ivana_88's problem?	Yes	No	can't get to the veterinarian until Monday, and the vet isn't answering the phone. My he seems to be in a lot of pain. I'd like to give her something to make her teel better until I can go to the vet Thank you for your help.		
NellieB79's post	0	0	• Nettients		Ensted 28 October 18 18
Monie's post	0	0	Lidon't know if genirin is safe fo	r hand or not I shows the	with my ust helping my dia
Avian_Deals's post	0	0	birds medicine 1 know that some drugs that are safe for humans can be very dangero for birds.		
Bob's post	Ó	0	NI MILUS.		
Frank's post	Ø	O	Mohie		Posted 28 October 18 52
			i gave an aspirin to one of my t day I went to the vet but she wa too much, so don't exceed the	tens when she was hurt. T as already better I think it dose limits! I hope she fee	here was no problem. The ne: might be dangerous if you give is better!
			Avian_Deals		Posted 28 October 19:07
			Hil Don't forget to check out my sale right now!	y super low deals on all bir	d supplies. I'm having a great
			2 Bob		Posted 28 October 19 15
			Can someone please tell me hi	ow to know if a chicken is s	ick? Thanks.
			Erank		Fewted 28 October 19:21
			Hello Ivana		

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

- Short constructed response Students respond by entering a short text in a response box that consists of a phrase or a sentence or two. 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.

	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%

Exhibit 3.9. Flexible Distributions of Item Response Types Across Grade Level

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, select, and order);
- Construct relationships among information (e.g., dynamic concept maps, multimodal representations); or
- Construct spoken responses (e.g., recorded spoken language in open-ended responses).

When selecting the format of any particular item, developers should be mindful of the cognitive and logistical demands of varied formats and how these may interact with reader familiarity and the time constraints of each activity.
Exhibit 3.10. Example of a Dynamic Search Engine Item from ePIRLS 2016 for Grade 4 Students



Universal Design Elements (UDEs)

Grounded in Universal Design of Assessments (Johnstone et al., 2006; Thompson et al., 2002), the NAEP 2026 Reading Assessment employs design features known as Universal Design Elements (UDEs). UDEs provide orientation, guidance, and motivation to sustain readers' journeys through the block. They are designed to mirror typical (non-testing) reading situations to improve the validity of the assessment.

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 (see <u>NAEP Accommodations</u>, last updated Oct. 2019). Examples of accommodations available on some assessments include extended time, options for responses in Braille or Sign Language, or having test-items read aloud. Universal Design of Assessments and the inclusion of UDEs are the means to enable *all* readers to validly demonstrate what they know and are able to do.

Types of UDEs. Examples of UDEs already exist in operational NAEP Reading (e.g., highlighters and look-back buttons) to reflect real-world experiences and how readers use 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 next. The next section clarifies the role of each UDE and offers some hypothetical examples of how these might appear in the 2026 NAEP

Reading Assessment. Additional details are provided in the *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, 2014).

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, in depth, a character, drawing on specific details in the text. To demonstrate this skill, the student needs to identify a character trait that is relevant, but selecting an accurate trait is insufficient to meet the construct measured. The student needs to be able to connect the selected character trait with a deeper interpretation of the character and the details of the text. In providing the word bank as a task-based UDE, 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., 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



Motivational UDEs. In the 2026 NAEP Reading Assessment, motivational UDEs are designed to facilitate students' interest in assessment content and persistence with challenging tasks (Alton & Proctor, 2008; Buehl, 2017; CAST, 2020; Guthrie & Klauda, 2015). Motivational UDEs might, for example, provide an engaging pre-reading preview 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 sub-tasks that need to be completed along the way. UDEs in the form of task characters may provide written and/or oral directions, or interact directly with readers as experts, teachers, or peers to provide information (see Exhibit 3.12). 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, Bon-Gettler, Kendeou, Carlson, & White, 2011).

Exhibit 3.12. Teacher and student task characters remind the reader of the task goal for the second task.



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 already will 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 when 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 <u>Chapter 2</u>) required to demonstrate their comprehension of challenging text (Alexander & Jetton, 2000; Buehl, 2017). Exhibit 3.13. Example of Two Informational UDEs from NAEP's "Five Boiled Eggs" Block



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."

Selecting appropriate locations for UDEs. Developers decide on appropriate locations in which to insert UDEs into each block of the assessment. Because some NAEP Reading 2026 tasks involve complexities in response to handling multiple tasks and texts, readers may be asked to check and reflect on their reading progress in the activity and allocate their attention accordingly. Intuitively designed transitions between each task, such as task characters, visual flow charts, or simple written statements may be used to guide readers through the task sequence and structure in any given block.

A major question for block developers is how to decide when to employ and when to forego the deployment of a specific UDE as the potential for added support is weighed against the potential for increased cognitive burden on the reader. Developers will also consider how to populate the grade-appropriate assessment space with UDEs while recognizing that readers have time limits within which to accomplish expected outcomes.

Process Data

Because 2026 NAEP Reading Assessment activities are situated in a fully digital environment, process data involving reader actions (e.g., number of mouse clicks, pathways

through a task or hypertext, transcribed voice responses, length of time spent engaged with reading material or responding to an item) can be easily collected in digital log files stored in a database. While these data are not reported for individual students, aggregations of these types of data hold potential power to measure levels of engagement in purpose-driven reading activities (e.g., capturing frequency, density, and intensity of engagement or identifying and comparing novice to expert level of practice). Process data from log files can be aggregated and interpreted to characterize how reader attributes or other 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, pages within passages, 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 fourth, eighth, and twelfth grade students. 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 has reported data for the nation as a whole, for participating states, and for large urban school districts that volunteer to participate in the NAEP Trial Urban District Assessment (TUDA). Results of the NAEP Reading Assessment administrations are reported in terms of average scores for groups of students on the NAEP 0–500 scale and as percentages of students who attain each of the three achievement levels (*NAEP Basic, NAEP Proficient*, and *NAEP Advanced*) discussed below. By design, the assessment reports results of overall achievement; it is not a tool for diagnosing the needs of individuals or groups of students.

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) legislation, states receiving Title I grants must include assurance in their state plans that they will participate in the reading and mathematics state NAEP at grades 4 and 8. Local districts that receive Title I funds must agree to participate in biennial NAEP reading and mathematics administrations at grades 4 and 8 if they are selected to do so. Their results are included in state and national reporting. Participation in NAEP does not substitute for the mandated state-level assessments in reading and mathematics at grades 3 to 8.

In 2002, NAEP initiated TUDA in five large urban school districts that are members of the Council of the Great City Schools (the Atlanta City, City of Chicago, Houston Independent, Los Angeles Unified, and New York City Public Schools Districts). Ten large districts

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

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

Achievement Levels

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

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.

Exhibit 4.1. Generic NAEP ac	hievement levels
------------------------------	------------------

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

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

- 1. Disaggregate scores for demographic subgroups in greater detail to provide a more accurate and dynamic description of student performance.
- 2. Expand the number of categories for reporting the achievement of English learners to better reflect the variability of English language proficiency within this population.
- 3. 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 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 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; socioeconomic status; 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 socioeconomic status within the student subgroup of race/ethnicity. In NAEP Reading, as in other large-scale assessments, lower levels of achievement historically are correlated with poverty. Disaggregating results by socioeconomic status within subgroups will reveal subgroup differences in reading achievement that are associated with socioeconomic status. 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 policy makers and stakeholders better understand reading performances in context. For example, these data may allow policy makers to consider how access to resources that support rich literacy opportunities may serve as an underlying driver of achievement.

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

Expanding Reporting Categories for English Learners

English learners (ELs) are defined by NAEP as students "who are in the process of acquiring English language skills and knowledge" (NCES, 2019). These students have not yet reached state-established standards for grade-level English proficiency and so are at the beginning or intermediate phases of acquiring English. In the prior NAEP reporting system, students were designated either as *not English learners* or *English learners* at the time of the assessment. The results for students who had been classified as ELs but who were no longer classified as such were reported along with students who had never been identified as ELs; hence, there was no way to disaggregate data to observe or track the successes and increases in achievement of former ELs.

The 2026 NAEP Reading Assessment results expand reporting categories in order to present data that is more attuned to the complex composition of today's student populations, and, thus, more informative for states and school communities (Durán, 2006; Hopkins, Thompson, Linquanti, August, & Hakuta, 2013; National Assessment Governing Board, 2014; Kieffer & Thompson, 2018). In keeping with the latest research and current requirements for state-level reporting under ESEA, Section 3121(a), the reporting system for the 2026 NAEP Reading Assessment disaggregates scores by three English proficiency categories for which school systems that participate in NAEP already collect data:

1. *Current English learners* – Students designated as English learners at the time of the assessment;

- 2. *Former English learners* Students who have reached grade-level standards of English proficiency within the last two years prior to the assessment and who have formally exited that status;
- 3. *Non-English learners* Monolingual students who speak only English; bilingual students who speak English and another language and who were never previously identified as English learners; bilingual students who reached grade-level standards of English proficiency more than two years ago.

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

Contextual Variables

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

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

The 2026 NAEP Reading Assessment uses an expanded set of research-based contextual variables (Guthrie & Klauda, 2015; Guthrie, Wigfield & Von Secker, 2000) to understand reading achievement (Solano-Flores, 2011; Solano-Flores & Nelson-Barber, 2001). Contextual variables are measurable, and some are also malleable (that is, 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, Afflerbach, Enciso, & Lesaux, 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, Brown, & Uccelli, 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, socio-demographic student characteristics, and student interests and experiences. Taken together, the network of contextual variables is intended to 1) correlate with performance on the outcome measure of reading comprehension; 2) be malleable (that is, influenced by differences in school and community settings); and 3) 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, Wixson, & Pearson, 2014). With respect to reader characteristics, the 2026 NAEP Reading Framework seeks to describe the role of students' perception of the interest, difficulty, and familiarity of texts, tasks, and contexts on their performances (Pintrich and Schrauben 1992; Eccles, O'Neil et al. 2005; Valencia, Wixson, & Pearson, 2014). Reader characteristic data to be collected from questionnaires and process data include the following:

Cognition and Metacognition

- 1. **Cognitive strategies** in reading comprehension refer to skills used to understand a text, such as drawing inferences to connect sentences together and checking to be certain that text information is fully understood (OECD, 2018).
- 2. **Metacognitive strategies** in reading comprehension refer to, for example, a student's use of a mental guidance system to perform such operations as deciding which sections of text are most relevant to an assigned reading goal, how to link two sections, and/or when to reread to seek more information or clarify understanding (Cho & Afflerbach, 2017).
- 3. **Topical knowledge** refers to students' use of their pre-existing knowledge of the reading topic to enable them to understand text information and construct new knowledge (O'Reilly, Wang, & Sabatini, 2019).

Engagement and Motivation

- 1. **Volume of reading** refers to the amount of reading a student does for personal interest, pleasure or learning (Schaffner, Schiefele, & Ulferts, 2013).
- 2. **Reading for enjoyment** refers to the goals, uses, purposes, reasons and benefits students have for reading in school and out of school (Pitzer & Skinner, 2017).
- 3. **Motivations for reading** refer to students' attention, effort, interest, and value for reading a particular text with a unique set of tasks and questions related to it (NAEP Reading Special Study, 2019).

Environmental Characteristics

Environmental characteristics are equally important in accounting for student performance. For example, students vary in their participation in cultural communities that may value reading in varied ways and integrate reading into their lives for different purposes (Skerrett, 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 (i.e., libraries, tutoring, out-of-school programs) is important, since as these environmental contexts shift, so do the roles of reading and texts in students' lives. The degree to which schools and communities offer access to out-of-school resources influences, to some degree, students' opportunities to learn, 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

- 1. **School social support** refers to the extent to which students perceive that their teachers and peers believe they contribute positively to classroom reading (through listening, speaking and interacting well with others) (Vaux, Phillips, Holly, Thompson, Williams, & Steward, 1986).
- 2. **Belonging in school** refers to the extent to which students perceive themselves to be accepted members of the school community (Faircloth & Hamm, 2005).

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

Self-Reports of Teacher, Instructional, and Classroom Supports

- 1. **Teacher support for reading engagement** refers to the extent to which students perceive their teacher(s) as providing materials and tasks that encourage the development of their reading competence and engagement (Afflerbach, Hurt & Cho, 2020).
- 2. **Teacher support for motivation** refers to the degree to which students perceive their teacher(s) to support their interests, and reading goals (Wigfield & Wentzel, 2007).
- 3. **Teacher support for students' background experiences** refers to the students' perceptions that their teacher recognizes and uses students' cultural, language, and social knowledge during reading instruction (Shin, Daly & Vera, 2007).
- 4. **Program and curricular support for reading development** refers to the extent to which teachers and administrators perceive that the school's reading program and curriculum enables them to support students' development of effective reading practices.

The NAEP 2026 Reading Framework expands collecting and reporting of contextual variables via use of refined survey item design, thereby allowing policy makers and stakeholders to gain more actionable insights regarding the variables' 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 performance (Guthrie, Wigfield & You, 2012). Students' positive perceptions of their teachers' support and classroom climate could also be associated with higher NAEP Reading performance (Pitzer & Skinner, 2017). If relations such as these emerge from NAEP, they could have meaningful implications for the need to attend to perceptions, identity, and affect to support reading comprehension and achievement (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 drawn from available state, district, or school records, information about some variables will be obtained from the *process data* (computer-generated records of navigational data collected automatically as students engage with the assessment) (Ho, 2017; Bergner & Davier, 2018). Exhibit 4.2 provides a list of variables, along with their source in the revised contextual variable plan.

Variables	Source		
		Teacher/	
	Ouestionnaire	Administrator	Process Data
Reader Characteristics	((
Cognition and Metacognition			
Cognitive strategies	\checkmark	\checkmark	\checkmark
Metacognitive strategies	\checkmark		\checkmark
Topical knowledge	\checkmark		
Engagement and Motivation			
Volume of reading	\checkmark	\checkmark	\checkmark
Reading for enjoyment	\checkmark	\checkmark	
Motivations for reading	\checkmark		
Environmental Characteristics			
Reports of School and Community Resources			
School social support	\checkmark	\checkmark	
Belonging in school	\checkmark	\checkmark	
Participation in out-of-school reading/literacy activities	\checkmark		
Reports of Teacher, Instructional, and Classroom Supports			
Teacher support for reading engagement	\checkmark	\checkmark	
Teacher support for motivation	\checkmark	\checkmark	
Teacher support for students' background experiences	\checkmark	\checkmark	
Program and curricular support for reading development	\checkmark	\checkmark	

Exhibit 4.2. Contextual Variables

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, both at the national and state level. When the NAEP Report Card for Reading is issued every two years, policy makers and the public pay attention, particularly to trend data. Yet, NAEP results have also been subject to misinterpretation (Linn and Dunbar, 1992; Jaeger, 2003; 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, either for 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.

- 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 Testing Commission, 2019; Task Force on Assessment of the International Reading Association, 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, importantly for the added category of former ELs. By reporting the performance of non-ELs and former ELs separately, it will be possible to determine whether the two groups perform at similar levels on the NAEP Reading Assessment.
- 5. Mine Process Data for Evidence of Cognitive and Metacognitive Processing. Initial forays evaluating the utility of the process (logfile) data for NAEP (Bergner & von Davier, 2018) and other digitally delivered assessments and instructional programs (Ho, 2017) suggest that there is substantial potential for using these navigational data as indirect indices of cognitive and metacognitive processes. These indices can be used, perhaps in triangulation with measures of the same variables from reading questionnaire responses, to understand comprehension performance more deeply. Simple bar graphs can be displayed in the Report Card, and data can be related to reading performance in the NAEP Data Explorer.
- 6. Enhance the Visibility and Utility of the NAEP Reporting Portfolio. An effort to expand, energize, and advertise the untapped resources of the NAEP reporting portfolio

would allow for more nuanced data analyses. The NAEP Data Explorer, for example, permits users to go online and generate more sophisticated analyses than typically appear in the Report Card, which, by its nature, can only provide foundational reporting. In the NAEP Data Explorer for the 2019 Reading Assessment, a user can query the database to obtain a report which, for fourth graders in the nation, breaks down the performance of low- versus high-SES students on the cognitive targets of Locate and Recall, Integrate and Interpret, and Critique and Evaluate when reading literary and informational text. For sound psychometric reasons, NAEP results are not reported separately for the comprehension targets; regardless, NAEP data can be used to obtain more in-depth, 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 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

Construction-integration model: Theoretical account that depicts the multiple models of meaning that readers create and employ to comprehend: surface level (accurate decoding or

literal meaning); text-based (key ideas and inferences within the text); situation model (the links that readers make between their knowledge and text ideas).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Operationalization: To put into action or to realize.

Opportunities to learn (OTL): Inputs and processes that enable student achievement of intended outcomes.

PISA: The Programme for International Student Assessment, an international assessment that measures 15-year-old students' reading, mathematics, and science literacy every three years.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Trait: A distinguishing feature or quality.

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

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

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

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

Vocabulary pop-up: 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. 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.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.2. Illustrative Examples of Texts and Other Media Across Single Static and Dynamic Texts and Multilayered Digital Text Environments

SINGLE STATIC TEXT

Examples of single static genres and forms of continuous prose, non-continuous prose, and everyday reading materials from which designers might sample as readers read to engage in literature, science, or social studies and history are found in Exhibit 2 in this appendix.

SINGLE DYNAMIC TEXT

Nonlinear text

Single text with hyperlinks that only connect to ideas within the same document; may also contain one or more dynamic media elements

Dynamic media

- Dynamic image
- Video
- Podcast
- Digital poster
- Infographic
- Interactive timeline
- Interactive chart or graph
- Data visualization
- Blog
- Simulation

MULTILAYERED DIGITAL TEXT ENVIRONMENT

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

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

Exhibit 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.3. Commissioned Texts: Parameters and Constraints

Guidelines for Using Commissioned Texts

The following guidelines seek to provide clarity about the circumstances under which commissioned texts might be used and the criteria with which developers should use such commissioned texts:

- Rare, never to exceed more than 5-10% of all texts included in NAEP at any grade level; 5% limit at 12th grade unless permission issues are encountered
- Only used when an appropriate authentic text cannot be located to include within a text set for a block, but never as an "anchor" text for a block
- Authored by writers within the discipline in which the block is situated and using specific criteria to meet strict guidance 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 items asking students to evaluate source credibility of such commissioned texts will be used
- Will meet the same complexity and other criteria for text selection as all texts for the NAEP Reading Assessment

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.

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

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

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.

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 Fanfiction Poetry Haiku, sonnet, ballad, dirge, epic, etc. Related Nonfiction Memoirs (Auto)biographies Literary analyses Reviews and recommendations 	 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 	 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 indicators (era, time, sequence, and tense)

Exhibit A.5. Typical Text Elements Across Disciplinary Contexts

	 Bench notes and science journals Procedures Published research articles Personal communications 	 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. Interpretive explanations or arguments about historical, social, and cultural phenomena and trends. Procedural texts, public service announcements 	 Linguistic frames and signals for organizing arguments, comparisons, and/or causal chains Lexical expressions that mark chronology or argument Abstraction and nominalization (e.g., to develop a chain of reasonings across events and happenings, e.g., this stance of brinkmanship) Rhetorical markers of persuasion 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 Event models (how historical events are described) Spatial (place, location) and temporal indicators (era, time, sequence, and tense)

Note: Many text types and elements are common across disciplines. All texts should include information about their sources and authors. In general, NAEP applies a standard of accuracy and trustworthiness to the texts it selects, especially in matters of scientific inquiry. For certain tasks, however, it is necessary to use texts with questionable, or at least different, levels of accuracy and trustworthiness if the purpose of a block, or a task within a block, is to engage students in analysis and critique of texts. It is even more likely that NAEP will employ texts that

represent different perspectives on an issue when students are asked to compare the multiple perspectives that texts/authors bring to a social or scientific issue.

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

SINGLE STATIC 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 through primarily static words, numbers, and/or visual graphics, such as those in a still photograph, diagram, or table.

SINGLE DYNAMIC TEXT

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 into 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 nonlinear elements both within a text (e.g., hypermedia or hyperlinks) that may lead to another text (e.g., another webpage within the same website or another webpage on a different website). Text features may include linked texts may contain either related or conflicting textual ideas. Multimodal elements (words, moving images, animations, color, music and sound) may appear in any or all texts.

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

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

Rules of Thumb	
\cdot 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	n grade level
All comprehension target	s are employed at each grade level.
 All comprehension target construct responses. What of and Integrate, to Analyze a text-based reasoning and the Moving across grades, the DDUI11 and a basis 	s require students to consult the text in order to select or changes across targets (from Locate and Recall, to Interpret nd Evaluate, to Use and Apply) is the sophistication of the ne inferences involved.
• RDU blocks, by definition	n, do not require the application of ideas to a new task. Hence
the bulk of Use and Apply	items will be in RSP blocks; nowever, NAEP should be open
to the possibility that an KI	JU 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.7 provides both the principles and ranges anticipated for the distribution of items for each comprehension target within blocks developed for each broad purpose (RDU and RSP) at grades 4, 8, and 12. Because item development is so greatly influenced by the affordances of the texts selected, the ranges for item types will vary from block to block, even within each broad purpose. Hence, as with previous frameworks, NAEP monitors the range of comprehension targets by looking at the total distribution across all of the blocks within a grade level for each disciplinary context.

Language Structures & Vocabulary Included / Excluded from Testing	Criteria
Included	 Words and language structures that appear across numerous texts, either across literary texts (e.g., <i>despise, benevolent</i>) or across social studies and natural sciences texts (e.g., <i>resolution, commit</i>) 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., <i>timid</i>). Words that include word parts (roots and affixes) useful to acquire and figure out the meaning of unfamiliar words (e.g., <i>disregard, counterargument</i>). Language that expresses logical relations between ideas (e.g., phrases that include connecting words such as <i>although, in contrast</i>) Expressions that refer to characters, events, or ideas previously introduced in the passage (e.g., <i>those alliances, this phenomenon</i>)
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.

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

Note: A total of 15-20 percent of items in any assessment block will assess passage-relevant Language Structures and Vocabulary knowledge while concurrently measuring a specific comprehension process.

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.

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

NAEP Policy Definitions

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

Range ALDs

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

Range ALDs should not be confused with <u>reporting ALDs</u>. The fundamental difference between the two is straightforward; range ALDs communicate <u>expectations</u>, and reporting ALDs convey <u>results</u>. 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 <u>should</u> 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 are constructed using empirical data during a later stage in the NAEP cycle, i.e., a live administration of the NAEP Reading Assessment. Further detail about the development of the reporting ALDs for NAEP is provided in the Governing Board's <u>policy statement on</u> <u>achievement level setting</u>.

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 (science, social studies, and literature). 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 his or her achievement level, will employ various knowledge types to accomplish the assessment's reading comprehension activities. In doing so, the reader will demonstrate achievement relative to four *comprehension targets:* (1) Locate and Recall; (2) Integrate and Interpret; (3) Analyze and Evaluate; and (4) Use and Apply. Students at each achievement level are expected to meet the demands of each comprehension target. However, as the complexity of texts increases on a given reading assessment, students, on average, are expected to demonstrate less competency with skills associated with higher-level comprehension targets, such as Use and Apply.

Broad and Specific Reading Purposes

Reading activities in an assessment block are situated within not only a disciplinary context but also 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. Reading to Develop Understanding (RDU) blocks ask students to *read and comprehend deeply* (analyzing, inferencing, interpreting, and critiquing) in or across disciplinary contexts. By contrast, Reading to Solve a Problem (RSP) blocks ask students to demonstrate understanding across multiple texts and related perspectives in order to solve a problem. Reading to Solve a Problem activities do involve comprehending text, but in the service of a specific action or product, such as a classroom presentation.

Both RDU and RSP blocks also have *specific* purposes with reader roles that shape how and why readers engage with the 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 flow charts, and follow step-by-step procedures to investigate scientific phenomena. Outside of school, students engage in reading science when participating in games, cooking, and crafts, and reading and viewing science and health news. NAEP simulates these Contexts of Reading to Engage in Science by providing test taskers with activities to respond to science and everyday texts like those read in and outside of school.

Social Studies Texts. People engage in reading social studies for the following purposes:

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

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

NAEP Reading Achievement Levels: Grade 4

NAEP Basic

Fourth-grade students performing at the *NAEP Basic* level should be able to locate specific pieces of information, identify relationships between explicitly stated pieces of information, make simple inferences and interpretations in static, dynamic, and multimodal texts, create 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, fourth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to identify or determine literary elements such as character point of view, theme or central message, problem, and setting. Readers should be able to explain how a text's illustrations contribute to what is conveyed by the text, explain the differences between poems, drama, and prose, and show understanding of vocabulary and simple figurative language. Readers should be able to produce a simple summary of a text and continue the narration of an incomplete story to a conclusion of their making.

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

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

NAEP Proficient

Fourth-grade students performing at the *NAEP Proficient* level should be able to make more complex inferences and interpretations, reconcile inconsistencies within and across 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, fourth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to describe in depth character, setting, and plot, and to explain how a theme or central message is conveyed through details in a text. Readers should be able to analyze how a printed version of a text relates to its multimedia version and show understanding of nuances in word meaning. Readers should be able to produce a detailed summary of a text and rewrite a story from a different character's perspective.

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including investigations), fourth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to explain events,

procedures, ideas, and concepts based on specific information in and across texts. Readers should be able to make predictions and to interpret an author's point of view or purpose, including in reference to a procedure or experiment and in comparison to another text's author. Readers should be able to develop a new procedure or experiment based on knowledge acquired from information gained from reading texts.

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

NAEP Advanced

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

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

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

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

NAEP Reading Achievement Levels: Grade 8

NAEP Basic

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

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

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), eighth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to determine the central ideas and conclusions of a text and explain how a text makes connections among and distinctions between individuals, ideas, and/or events. Readers should be able to integrate quantitative or technical information expressed in words in a text with a version of that information expressed visually (e.g., in a flowchart, diagram, model, graph, or table), show understanding of how to follow precisely a multistep procedure of an experiment, and show understanding of academic and domain-specific vocabulary, key terms, and symbols. Readers 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, eighthgrade readers performing at the *NAEP Basic* level should be able to determine the central ideas, determine and interpret an author's point of view or purpose, and distinguish between fact, opinion, and reasoned judgment in a text. Readers should be able to identify key steps in a text's description of a process related to social studies (e.g., how a bill becomes law). Readers should be able to produce a simple objective summary of a text and integrate information from multiple sources to produce a new text of informational or argumentative purpose.

NAEP Proficient

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

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

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

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

NAEP Advanced

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

When engaged in reading literature texts such as fiction, drama, film, poetry, and literary nonfiction, eighth-grade readers performing at the *NAEP Advanced* level should be able to use textual evidence as support to analyze how multiple literary elements in a text relate to each other and to analyze points of view of and between character(s) and the reader/audience. Readers should be able to analyze how a modern text draws on themes, patterns of events, or character

types from myths or traditional stories, and then evaluate how these elements resonate with society and their personal lives. Readers should be able to produce a literary text that adapts elements of a myth into a contemporary retelling based upon the reader's personal experience.

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

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

NAEP Reading Achievement Levels: Grade 12

NAEP Basic

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

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

When engaged in reading science texts such as exposition (including literary nonfiction), argumentation, and procedural texts (including experiments), twelfth-grade readers performing at the *NAEP Basic* level should be able to use textual evidence as support to analyze the specific results of a multistep procedure based on explanations in the text, explain how specific individuals, ideas, and/or events interact and develop over the course of a text, and analyze how

the text structures information or ideas into categories or hierarchies. Readers should be able to compare and contrast findings presented in a text to those from other sources and show understanding of general academic and domain-specific vocabulary, key terms, and symbols. Readers should be able to generate an alternative procedure or experiment based on knowledge acquired from information gained from reading texts.

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

NAEP Proficient

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

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

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

When engaged in reading social studies texts such as exposition (including literary nonfiction), argumentation, and documents of historical and literary significance, twelfth-grade readers performing at the *NAEP Proficient* level should be able to use textual evidence as support to

analyze how the central ideas interact and build on one another to produce a complex account. They should be able to analyze the themes, purposes, and rhetorical features of foundational U.S. documents and evaluate the effectiveness of the structure in the text's exposition or argument. They should be able to develop a detailed objective summary of a text. Readers should be able to evaluate multiple sources of information presented in different media or formats (visually or in words) in order to produce an argumentative text with evidence to structure and support a judgment.

NAEP Advanced

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

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

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

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

This 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 which, in turn, inform the comprehension items. Items are created in relation to item response formats, as different formats are used to collect different kinds of information. Similarly, all assessment components inform the use of UDEs because UDEs are used to help ensure that all students can gain access to the tasks required of them to complete the assessment and that the assessment measures students' reading comprehension of the texts and not something else (e.g., how well they can read or follow test directions). In this manner, a well-integrated block results, with all of the parts working in tandem.

Exhibit 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 sub-tasks, for readers to complete as they read one or more texts. For each task, there might be one or more comprehension items. UDEs are only employed as needed to bolster construct validity and ensure better measurement of the reading comprehension construct. Similarly, process data are only collected in places where developers think it might be useful for understanding why students perform the way that they do or for informing revision or future research and development.

As developers develop a block, they make decisions about each of the components described in Exhibit 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.





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.

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.	1	Purposes are paired with an essential inquiry question or problem to be examined throughout the task. All tasks and items within the block help readers work towards this theme, question, or problem.
Reader Role	Fewer parameters are specified for the reader's role. The reader is placed in a situation that provides fewer pieces of information about how to engage with the provided tasks and texts. The reader might be placed within a situation that contextualizes expectations for how to engage with provided texts and tasks. However, this situation provides less information about that role.	\$	More parameters are specified for the reader's role within the block. The reader is placed in a situation that provides multiple pieces of information about how to engage with the provided tasks and texts. Readers may be assigned a particular role, and their role may be more specified, particularly in relation to reading purpose(s) and expected outcome(s).
Tasks	Purpose-driven tasks and items are situated in line with disciplinary context, but tasks are less related to one another 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; more probability that tasks are interdependent; may have more

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

	across items within tasks; less need for resetting. Less involved culminating task, or no culminating task. Task not necessarily a determinant of all items in block.		need for resetting. More involved culminating task at the end of an activity that directly addresses the question or problem; major driver of the block.
Texts	<i>Number</i> : 1-3 topically related texts; excerpts may be included.	1	<i>Number</i> : 2-4 topically related and interconnected texts may be included. Readers may be asked to choose only some texts to engage with and in line with task purposes.
	<i>Dynamism</i> : More static texts with minimal dynamic features.		<i>Dynamism</i> : More texts with dynamic and/or or multimodal text features.
	<i>Linearity:</i> Fewer nonlinear structures to navigate within or across texts; less variation in structures across texts.		<i>Linearity:</i> More nonlinear structures to navigate within or across texts; more variation in structures across texts.
	<i>Features</i> : Texts include a narrower range of features and fewer types of media.		<i>Features</i> : Texts include a wider range of features and more types of media.
Items	Items are less connected to the overall specific reading purpose for the block and there are more opportunities for items to be related, but less connected, to this specific purpose and to the related tasks; Less dynamic item formats to support less complex tasks and items.	\$	Items are more connected to the overall specific reading purpose for the block. There are more opportunities for items to be more directly related to the specific reading purpose for the block and to the related tasks; More dynamic item formats to support more complex/multilayered tasks and items.
Universal Design Elements (UDEs)	Fewer cumulative reading purposes that may require UDEs for knowledge or motivation and potentially lesser need for task- based UDEs.	•	More cumulative reading purposes that may require UDEs for knowledge or motivation and potentially greater need for task- based UDEs.

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

Specific Guidelines for Block Development

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

- 1. Students deserve to know the tasks that lie ahead of them in the block. Guidance in the form of task-based UDEs is essential.
 - a. Both 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 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, O'Reilly, Weeks & Wang, 2019) and in the current 2019 operational NAEP SBT blocks offer an existence proof that these guidance features are manageable by 4th, 8th, and 12th graders. When these sorts of guidance features were included along with other UDEs in the 2017 special study, the enhanced blocks provided an overall comprehension performance advantage and resulted in higher motivational ratings by students, especially in the earlier grades. NAEP needs to monitor these matters with great vigilance.

Block Sketches

Sketches of three different blocks are provided to illustrate a range activity within assessment blocks that students might encounter when they participate in the 2026 NAEP Reading Assessment. To accomplish this goal, the Appendix offers three hypothetical sketches of blocks (showing only a sampling of items from each) that might be developed using the components (from Chapter 2) and the design principles (from Chapter 3) of the 2026 NAEP

Reading Framework. Importantly, these sketches are designed to exemplify key concepts from the framework and do not represent blocks or items that will be used on future NAEP assessments. Tasks presented with multiple sample items are provided to help readers of the framework envision how theoretical ideas in the framework might guide assessment design. However, these sketches do not represent fully expectations for enacting the NAEP style guide and other test specifications.

The first example (labeled *Hana* because it is built upon a short story text entitled *Hana Hashimoto, Sixth Violin* by Chieri Uegaki and Qin Leng) illustrates a block developed for the broad purpose of Reading to Develop Understanding (RDU). The second example (labeled *Hill District* because it is built upon a set of activities surrounding an authentic civic issue in the Hill District neighborhood of Pittsburgh, PA) illustrates a block developed for the broad purpose of Reading to Solve a Problem (RSP). 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, grade 4 readers read and interpret story excerpts from the short story, Hana Hashimoto, by Chieri Uegaki in preparation for a book discussion with three peers. First, students are asked to read to develop an understanding of the characters, key events, and author's 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 a 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 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 lends is otherwise classic RDU block, but lends itself to a Use/Apply culminating task (which is more characteristic of RSP blocks).
- The inclusion of the *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 Grade 4 readers develop understanding within a single, print 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 fourth 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 1) learn about important events in the story and characters' thoughts, feelings, and actions; and, 2) 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



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 towards 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 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 select 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)

Hana Hashimoto, Sixth Violin By Chieri Uegaki & Qin Leng			
When Hana Hashimoto announced that she had signed up for the talent show and that she would be playing the violin, her brothers nearly fell out of	What doe show?	es Hana want to do fo	r the talent
a tree. "That's just loopy," said Kenji. "You're still a	A O	Sing a song	Θ
"Stop kidding," said Koji. "You can barely play a note."	вО	Tell jokes	Θ
"It's a <i>talent</i> show, Hana." "You'll be a disaster!" Hana squared her shoulders and	٥ ،	Play the violin	Θ
took her violin and bow inside, leaving her brothers laughing like monkeys in	DO	Climb a tree	e

Exhibit C.7. A Grade 4 Locate and Recall item illustrating a fill in the blank short constructed response item

By Chieri Uegaki & Qin Leng	When Ojiichan plays his song about a crow
She pulled at the strings, letting them twang. It was true that she was still a beginner. She had only been to three lessons. The first time Hana held a real violin had been that summer, while visiting her grandfather in Japan. Long, long ago, her grandfather had been part of a great symphony orchestra in Kyoto. Ojiichan had been Second Violin and once played in front of the Imperial Family. Ojiichan usually played classical pieces by Mozart or Mendelssohn or Bach. But in the indigo evenings, Ojiichan would sit on the veranda and play requests. Hana always asked for a song about a crow cawing for her seven chicks. Whenever Ojiichan	cawing for her seven chicks, Hana feels:
played it, Hana would feel a shiver of happy-sadness	

Exhibit C.8. A Grade 4 Locate and Recall item illustrating a multiple select 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 on the underlined quote 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 select 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

Let's organize our	notes into		OUR	NOTES	
Good ideal He	re are	Note 1: Ha fun of her. anyway. Th practiced e	na's brothers made She practiced e text says, "Hana very day."	Note 3: Wh stage, she The text sa do her best	hen Hana is on decides to play. ys, "She would just t."
Move the notes from the notepad to sort the notes and prepare for t discussion.	s so far. Into the chart the class	Note 2: Wh stage, she is The texts sa swallowed medicine."	ien Hana gets on s feeling nervous. ays, "Hana her nerves like	Note 4: At story, Hana her violin in family. The happily pla again."	the end of the a is happy to play n front of her text says, "She yed her sounds
Hana's Thoughts	Hana's Fe	eelings	Hana's A	ctions	Hana Hashimoto
					Story NEXT

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 sub-groups of 4th graders) can do in each disciplinary context as they engage with texts and test items, while also being encouraged to draw from and use the knowledge, skills, and experiences they bring to that reading context.

Hill District, Grade 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 Reading to Solve a Problem block reflect design features that are more dynamic and cumulative in terms of content and format, as depicted toward the right side of the continuum in Exhibit 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 sub-tasks) are tightly structured so that one task builds on the previous, such that readers are asked to learn more about the project goals and get a general sense of the public's comments before they are asked to gain a deeper understanding of the historical significance of the proposed park.

The test block also includes opportunities for students to engage with several interconnected digital texts (e.g., excerpts from social media, search engine results, and multimedia websites and online news articles) that represent the perspectives of different kinds of community members and cuts across issues of contemporary and historical relevance.

Throughout the block, readers 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 Reading to Solve A Problem block involving several interconnected digital texts



Exhibit C.17. Same-aged task characters and a task-based UDE in the form of four taskspecific purposes serve to guide and motivate readers in the RSP block



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



Exhibit C.19. A Grade 12 Locate and Recall item illustrating a multiple-selection multiple choice response format



Exhibit C.20. A Grade 12 Locate and Recall item illustrating a single-select multiple choice item response format

TRIB LIVE	SEA	Righting a Wrong	Next City				
Pittsburgh C about \$32 m "cap" design Downtown v	ity Council hillion in sp ed to link t ha a three-	on Tuesday authorized ending for the interstat he Lower Hill District an acre park.	e 579		Accordir a large p	ng to the article, which organization part of the 'cap' project?	is funding
PennDOT wi july and end and Exhibition Allegheny Co	Il oversee in late 202 on Authori ounty,	the work expected to st. 21, according to the Spo ty of Pittsburgh and	art by ris	an den e	A O	Lower Hill District	\odot
Plans call for lawns, lands	r handicap caping, rec	ped-accessible pedestri reation areas, art and r	an pathways, an eplacement of wa	amphitheater, stage, Ikways in the area,	вО	Penn DOT	Θ
"The 'cap' pro- barrier and i services in D 1-579, known	oject will b re-establish owntown n as the Cro	e transformative for the ning connectivity to cen Pittsburgh," according to postown Expressway, ru	Hill District by re ters of employme to the SEA website	moving a physical nt, education and	¢٥	Cross Town Expressway	Θ
It will compli office redeve	essentially bs support ement the elopment p	ed by beams and pillars Penguins' long-awaited plan for the 28-acre form	\$450 million resi ner Civic Arena si	dential, retail and	٥o	Sports and Exhibition Authority	Θ
The SEA in 2 additional \$	016 receiv 5.2 million.	ed a \$19 million federal The remaining funding	grant for the wor is coming from t	k and is kicking in an ne state.			
Council una	nimously a	pproved the allocation	without commen				

Exhibit C.21. Two Grade 12 items that ask readers to Integrate and Interpret (item 1) and Locate and Recall (item 2) textual and visual information from an artist's rendering of the site improvement plan published on a website

SEA Righting a Wrong Next City Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan Site Plan				According	o the site plan, what read	will the park
SEA Righting a Wrong Next City Image: City Image: City Image: City <th>_</th> <th>I</th> <th></th> <th>replace?</th> <th>o the site plan, what road</th> <th>will the park</th>	_	I		replace?	o the site plan, what road	will the park
Site Plan • • • Highway I-579 • • • A park with trees • • • A park with trees • • • Highway I-579 • • • • Highway • • • • • • • • • • • • • • • • • • •	SEA	Righting a Wrong	Next City	A O	Bigelow Boulevard	0
Site Plan Site Plan Contral Ave. Highway Contral Ave. Highway Contral Ave. Highway Contral Ave. Highway Contral Ave. Highway Contral Ave. Highway Contral Ave. Highway Contral Ave. Highway Contral Ave. Highway Con		1000	Sports & Exhibition A	therity 0	Highway I-579	Θ
Site Plan Site Plan	P		These percent for the Annual Annual Statement Statement Statement Statement Statement Statement		Central Ave. Highway	Θ
According to the site plan, what are some of the features the park will offer? Select ALL that apply.		Site Plan		- O	A park with trees	Θ
A O Story Wall O Terrace C O Event Lawn	Sm Iwenovovicu	ts Planu - Reference	1	According to	the site plan, what are so	me of the
A ○ Story Wall - B ○ Terrace - C ○ Event Lawn -	6	- Landing Room	Constant Section	features the	park will offer? Select AL	L that apply.
© Event Lawn				*0	Story Wall	Θ
CO Event Lawn Θ	-				Terrace	Ξ
		NE -	130	• 0	Event Lawn	Θ

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



Exhibit C.23. A Grade 12 short constructed-response item with a look-back button (taskbased UDE) that asks readers to integrate and interpret information in an online newspaper article about the historical significance of the park's design

VE SEA Righting a Wong Head City Second Local South Cyter Add Line Business	Centure Us 1 9 0 0	What do you think the author, Mark Belko, means when he writes that the
A LO	Share and and	park effort attempts to " <u>right a wrong</u> "
ighting a wrong. New park over I-	and the second	that occurred half a century ago? Use
79 to reconnect Downtown and		evidence from the text to support your
he Hill District		answer.
MAX INLAS S Presure Paulo Salemi metalludysol gavetic.com ()	See what	
A \$32 million effort to "right a wrong" that occurred half a ombury ago gets its start Friday.	your monthly	
 A slew of politicians and stakeholders will gather in a parking lot near PPG Paints Arena to break ground on a three-aree park that will straddle Interstate 579/Crosstown Boulevard. 	income could look like in 30 seconds.	
To its supporters, the project is more than a green oasis surrounded by parking. It represents a literal reconnection of		
Downtown and the lower Hill District, a link severed to a large extent when the former Civic Arena was built in late 1950s.	CALCULATE	
The arena construction destroyed part of the Hill neighborhood,	-	

Exhibit C.24. A Grade 12 short constructed-response item that asks readers to integrate and interpret information on a web page with a look-back button (task-based UDE). The task character reminds readers of the specific purpose of the second task



Other potential items might ask readers to locate and evaluate the relevance of search engine results pertaining to the historical significance of some of the park's features (see Exhibit C.25) or locate (navigate to) and then analyze information from a website's menu to evaluate the expertise of the group responsible for publishing information about the park project (see Exhibits 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

TAS	K 2. Now, can you help us with some Interne	t research to find background information to share?
Dire abc use sup	ections: Read the list of Google [®] Search Resu ut the history of Pittsburgh's Hill District and the box to explain why that link is the best c port your thinking.	Its. Choose the link most likely to have information why this history is relevant to the park's plan. Ther hoice. Use evidence from the search engine results
https://www.bostonmagazi	te.com / news / 2017/04/05	
A Park Beneath the Apr 5, 2017 – A new park, o In June. It will stretch eight	I-93 Overpass Will Open in June - Boston alled Ink Underground, It scheduled to open under the 1-93 overpass acres and feature public art.	Your explanation:
https://archive.triblive.com - Pittsburgh City Coun- Jul 2, 2018 – A project to bui cap' linking the Lower Hill Dir office redevelopment plan for	cel i pitteburgh-allegheny : cil OKs 'cap' park over Crosstown	
https://pittsburghpa.gov / The Greater Hill Di The Neighborhood. As one continues to play an impor	kcp + hill-district I strict Master Plan pittsburghpa.gov - City of of Pittsburgh's earliest and largest neighborhoods, the HIII District tant role in the story of African-Americans	
https://www.post-gazette.cc Righting a wrong: 1 Jun 14, 2019 — A \$32 milli start Friday. A slow of politic Arena to break ground on a Boulevard.	m : 2019/06/14 > stories is lew park over I-579 to reconnect n effort to 'right a wrong' that occurred half a century ago gets its lans and stakeholders will gather in a parking tot near PPG Paints three-acre park that will straddle interstate 579/Crosstown	NEX

Exhibit C.26. A Grade 12 item selected response zone item designed to capture process data about how readers navigate through hyperlinked web pages



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

TR	IB LIVE SEA	Righting a Wrong Next Cit	у		Directio	ns. Select "SEA Hi	story" from the yellow
	ENT		- interest		answer	the question belo	W.
-	EN-	A	Sports & Exh	ibition Authority			
2	About Us Facil	ties Parking Park	s PublicArt 1-579 C	AP Relabed Projects B Opp	Do you	think the SEA is a	trustworthy source for
	SEA History	-			informa No. The	tion about the par n use the box to e	<pre>rk project? Select Yes or xplain vour choice using</pre>
	SEA Board Members	As a joint authority for the City and educational, cultural, civic, and so	nd County, the SEA provides venues f ocial events for the public. The Autho	or sporting, entertainment, rity owns and leases PNC Park,	details	rom the text.	, , ,
	SEA Staff	Heinz Field and CONSOL Energy	Center. The Authority owns and is res	ponsible for the operation of			
	MACIALITY ADDITIONEN PUNCTIV	the David L. Lawrence Convention	n Center (Convention Center). The SE	A also owns two parking	0	Yes	
	Stadium Authority Board I	sem facilities, riverfront parks, and va	rious associated infrastructure impro	vements.			

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 as well as functioning 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

Can you help us sort some of their	comments?	ark on social media.
Directions: Complete the chart by n perspective on the right.	noving each comment to accurate	ely match with a
A Cortland @cortland Wow – this will be a great place to bring my kids to play! #Hill District	Economic Perspective	
Jay Anderson @janders459 I don't understand why the city wants to spend their money on this park. I don't think this is a good use of our tax dollars. #Hill District	Environmental Perspective	(
Pedro Carano @caranofamily		
I like the idea of a park because it provides lots of trees and green space. But, why should be it built on a highway overpass? #PittCityPlanners #Hill District	Educational Perspective	
Ms. Peters @petersgrade8 I noticed in the park plan there were several signposts with a picture of a young girl named Keisha. Where can I read more about Keisha so I can talk with my students about how she fits in the planner's vision of the park? #PS57	Recreational Perspective	

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-select response format with a **task-based UDE (table) and motivational UDE (task character)** that serve to support readers as they engage in one particular item in the block. That is, the table is designed to first help readers focus their attention on relevant comments on the left side (rather than referring back to them in the original text) and then, 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**) ask readers to move each comment into the appropriate cells of a table that is part of a retractable digital notepad (**task-based UDE** marked near a blue arrow to illustrate how it can be minimized and maximized on the screen as needed). Readers use the notepad to store, organize, and recall important details as they read across multiple sources to solve the problem. 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-select response grid item with a task-based UDE (table) and motivational UDE (task character) that serve to support readers as they engage in one particular item in the RSP block


Exhibit 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

I found a lot of different opinions and persp organize these by topic, I'll add them to ou	ectives on the park ir r summary report for	n this website. If y the Mayor.	rou can help
Directions. Below are comments from two commu applies to each comment and if that person's comm Then select and drag each comment to the appropriet	nity members. Deterr nent would be consid riate box in the table	nine which persp ered a benefit or in your notepad.	ective best concern. Notepad
Longtime Hill District Resident Brenda Tate: For Brenda Tate, who has lived on the same block of Webster Avenue in the Hill for all of her 70 years, the park once again will give her the chance to traverse Wylie Avenue to the park then into Downtown and back. "There won't be separation. There will be a clear avenue to come back and forth. It's symbolic," she said.	Recreational Perspective	Benefits	Concerns
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)	Environmental Perspective Economic		
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	Historical Perspective		

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

Pittshurgh Chy Cauncil on Tuesday authorized about 532 million in spending for the interstate 579 card designed to link the Lower HIID District and Downtown via a three-acre park. What do people who like to exercise think about the proposed park plan? PennDOT will oversee the work expected to start by July and end in late 2021, according to the Sponson Allegheny County. Use your notepad to briefly describe a benefit and a concern about the park plan from a recreational perspective. Use details from the text and the video to support your answer. Notepad Plans call for handicapped-accessible pedestrian pathways, an amphitheater, stage, lawns, landscaping, recreation areas, art and replacement of walkways in the area. Benefits Concerns The Yap project will be transformative for the HIII District by removing a physical barrier and re-establishing connectivity to centers of employment, education and services in Downtown Expressivay, runs between the HIII and Downtown. The cap will essentially serve as a large bridge deck made of 8-inch-thick, reinforced concrete slabs supported by beams and pillars. Environmental Perspective Environmental Perspective It will complement the Penguins' long-awalted \$450 million residential, retail and office redevelopment julion for the 28-acre former Civic Arena site. Environmental Perspective Environmental Perspective Line Join Conder a \$10 million federal grant to the work and is kicking in an additional \$52.2 million. The remaining funding is coming from the state. Economic Perspective Historical Perspective Historical Perspective Perspective Historical Perspective <t< th=""><th>TRIB LIVE</th><th>SEA</th><th>Righting a Wrong</th><th>Next City</th><th></th><th>Directions: Click t Then answer this</th><th>he notepad to open question.</th><th>your notes.</th><th></th></t<>	TRIB LIVE	SEA	Righting a Wrong	Next City		Directions: Click t Then answer this	he notepad to open question.	your notes.	
Plans call for handicapped-accessible pedestrian pathways, an amphiheater, stage, lawns, landscaping, recreation areas, art and replacement of walkways in the area. The Cap injectivili be transformative for the Hill District by removing a physical barrier and re-establishing connectivity to centers of employment, education and services in Downtown Pittsburgh, "according to the 5EA website. 1-579, known as the Crosstown Expressway, runs between the Hill and Downtown. The cap will estantially serve as a large bridge deck made of 8-inch-thick, reinforced concrete slabs supported by beams and pillars. It will complement the Penguins' long-awalted \$450 million residential, retail and office redivelupment plan for the 28-acre former Civic Arena site. The SEA in 2016 received a \$19 million federal grant for the work and is kicking in an additional \$5.2 million. The remaining funding is coming from the state. Council unanimously approved the allocation without comment.	Plitsburgh (about \$32 n "cap" design Downtown PennDOT w July and end and Exhibiti Allegheny C	Ity Council hillion in sp ed to link t via a three- ill oversee i in late 20: on Authori ounty.	on Tuesday authorized bending for the Interstat the Lower Hill District ar acre park. the work expected to st 21, according to the Spo ty of Pittsburgh and	e 579 art by rrs		What do people w the proposed par Use your notepad a concern about t perspective. Use o video to support	vho like to exercise t k plan? I to briefly describe he park plan from a details from the text your answer.	chink about a benefit and recreational and the	Notepad
services in Downtown Pittsburgh," according to the SEA website. Is 79, known as the Crosstown Expression, runs between the Hill and Downtown. The cap will escentially serve as a large bridge deck made of 8-inch-thick, reinforced concrete slabs supported by beams and pillars. It will complement the Penguins' long-awalted \$450 million residential, retail and office redevelopment plan for the 28-arce former Civic Arena site. The SEA in 2016 received a \$19 million federal grant for the work and is kicking in an additional \$5.2 million. The remaining funding is coming from the state. Council unanimously approved the allocation without comment.	Plans call fo lawns, lands "The 'cap' pr barrier and	r handicap icaping, rei oject will b re-establis	ped-accessible pedestri creation areas, art and r e transformative for the hing connectivity to cen	an pathways, an amphitheate eplacement of walkways in the Hill District by removing a ph ters of employment, education	stage, area. ical and		Benefits	Concerns	
It will complement the Penguins' long-avaited \$450 million residential, retail and office redevelopment plan for the 28-acre former Civic Arena ate. The SEA in 2016 received a \$19 million federal grant for the work and is kicking in an additional \$5.2 million. The remaining funding is coming from the state. Council unanimously approved the allocation without comment.	I-579, know The cap will concrete sla	owntown n as the Cri essentially bs support	Pittsburgh," according to osstown Expressway, ru i serve as a large bridge ted by beams and pillar	o the SEA website. ns between the Hill and Down deck made of 8-inch-thick, rei 5.	wn. iorced	Recreational Perspective			
The SEA in 2016 received a \$19 million federal grant for the work and is kicking in an additional \$5.2 million. The remaining funding is coming from the state. Council unanimously approved the allocation without comment. Historical Perspective	It will compl office redev	ement the elopment (Penguins' long-awaited plan for the 28-acre form	\$450 million residential, retail ner Civic Arena site.	nd	Perspective			
Council unanimously approved the allocation without comment. Historical Perspective	The SEA in 2 additional 5	016 receiv 5.2 million	ed a \$19 million federal . The remaining funding	grant for the work and is kick is coming from the state.	ginan	Economic Perspective			
	Council una	nimously a	pproved the allocation	without comment.		Historical Perspective			

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.

TASK 3. the Mar three?	We are ready to summarize the views of different of yor. This is a big task, so can you help with two of the Thank you!	community members for our meeting with ne perspectives and then I'll finish the other
	Directions. Open your notepad to recall comments raise read about the I-579 Cap Project. Think about how their complete the items below. You can also select the notep view any of the sources you read.	ed by different community members in the texts you comments reflect different perspectives. Then pad to view your notes or click the links on the left to
Notepad	Part 1. Choose one perspective (safey, recreational, environmental, economic, or historical) and summarize briefly the benefits and concerns about the park proposal from that perspective. Be sure to cite one at least one piece of evidence from the texts you read to support each benefit and concern you listed.	Part 2. Choose a second perspective (safety, recreational, environmental, economic, or historical) and summarize briefly the benefits and concerns about the park proposal from that perspective, using evidence from the texts. We will include your written summary as part of our final report to the Mayor.
SEA Website Righting a Wrong Website		
Next City Website		NEX

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.

Task 4. No residents, summary	ow that you know more about the features of the park plan and the perspectives of Pittsburgh , you probably have your own opinion too! We'd love to include your own opinions in our final for the Mayor too.
Notepad.	Directions. Imagine you live in Pittsburgh and will attend the community meeting to express your views. Follow the directions to share your opinion. You can also select the notepad to view your notes or click the links on the left to view any of the sources you read. Choose the perspective (recreational, environmental, economic, or historical) that best relates to your own interests in the CAP Project and summarize briefly what you think about the park
Sources	proposal from that perspective. Support your thinking using evidence from one or more texts.
SEA Website	
Righting a Wrong Website	
Next City Website	

Exhibit C.34. This alternative format for the final Use and Apply item with openconstructed 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.

Task 4. resident summa	Now that you know more about the features of the parts, you probably have your own opinion too! We'd love ry for the Mayor too.	rk plan and the perspectives of Pittsburgh e to include your own opinions in our fina
Dir vie not	ections. Imagine you live in Pittsburgh and will attend the o ws. Follow the directions to share your opinion. You can als es or click the links on the left to view any of the sources y	community meeting to express your so select the notepad to view your ou read.
Sources	Choose the perspective (recreational, environmental, economic, or historical) that best relates to your own interests in the CAP Project and summarize briefly what you think about the park proposal from that perspective. Support your thinking using evidence	Type your answer in the box. OR
TRIB LIVE SEA Website	from one or more texts. You can choose to type your answer or make a voice	Click the blue microphone button to record
Righting a Wrong Website Next City Website	recording.	your answer.

Performance Evidence and Indicators. Scores from the Hill District block reveals 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 sub-groups 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 Reading to Solve a Problem 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

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 reading to develop understanding, by answering text-based questions that promote close reading of two texts as well as drawing inferences about how the ideas in the two texts inform one another. **Block Components (Disciplinary Context, Purposes, and Reader Role).** In this example, students read and answer questions about two texts representing common 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 informed about the role (working on their own) they will be asked to assume during the block, 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, 1) Reading and answering questions about the biographical sketch, *Not Just for Kids Anymore;* 2) Reading and answering question about the essay, *Twins*, and 3) Reasoning across the two texts to explain how what was learned in *Not Just for Kids Anymore* helps to understand E.B. White, the narrator of the essay, *Twins*. See Exhibit C.36, which shows task 1.

Exhibit C.36. Introduction to the grade 8 E. B. White literature block

Meet the author: E. B. White, the author of children's classics Charlotte's Web and Stuart	Why does [DiConsiglio think it is hard to label E. B. White?	
Little, was also a great essayist.	A O	White was as happy in a crowd in New York City as on a farm in New England.	Ξ
Not Just for Kids Anymore			
By John DiConsiglio	вО	White was a great writer.	(Ξ)
"I have a lot of the cat in me," said author E. B. White, "and cats are not joiners."	.0		
Perhaps that is why White, one of the country's	.0	white was well-liked by many people.	2
greatest writers, is so hard to label. His essays for <i>The</i> New Yorker appealed to an urbane crowd, but he is best remembered for his children's books. He loved	۰O	White could write more than one type of prose.	Θ
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.		NE	EXT

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—benign satire to biting critique.
- The statement near the end suggesting that his essays matched his personality.
- The very last statement, suggesting that he was an eminently likeable character. In terms of UDEs, note that there is 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:

- Getting 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

1 Idea from Not Just for Kide Anumora	2 How the idea applies to the perrotor of Twine
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.	

Exhibit C.42. Integrate and Interpret item illustrating resetting of item responses from prior item

No questions to answer on this screen. Below is from the Idea Box dragged into the correct spac you complete the next (and last) item in this blo	a the chart from the previous page with the phrases ses in the chart. You can refer back to this chart when ck.
1. Idea from No Longer Just for Kids Cats are not joiners.	2. How the idea applies to the narrator of <i>Twins</i> White and his companion stayed back from the
He could adapt to many settings.	Others who could see the moose. 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.
 When at the zoo, the narrator was able to sit He showed great respect for the animals at the He was capable of biting criticism. He graduated from Cornell University. 	Idea Box back and enjoy the birth of the twins he zoo

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

For the final task, you wil thoughts about how idea of <i>Twins</i> .	I use ideas from the biographical sketch to support your s from DiConsiglio's biographical sketch apply to the narrator
Use the <u>completed chart</u> to support your answer.	on the previous page or go back to either passage to get idea Fype your answer into the box below.

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

From *LITERARY CAVALCADE*, April 2000 issue. Copyright © 2000 by Scholastic Inc. Reprinted by permission of Scholastic Inc.

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.



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.

> Reprinted by permission of International Creative Management, Inc. Copyright © 1948 by E.B. White

Footnote

Sample items in the framework are being provided to exemplify key concepts in the framework and do not represent items that will be used on future NAEP assessments. These sample items may not represent accurately the full set of NAEP style guide and other test specifications. Tasks presented with multiple sample items are provided to help readers of the framework envision how theoretical ideas in the framework might guide assessment design, but they do not represent fully expectations for enacting the NAEP style guide and other test specifications.

APPENDIX D: ADDITIONAL EXAMPLES OF RESPONSE FORMATS AND COLLECTION OF PROCESS DATA

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

PISA			2 3 1		
Rapa Nui Question 6 / 7			Blog Book Review Science News		
Refer to all three so the tabs. Drag and drop the common, into the o theories.	ources on the right b causes, and the effe correct places in the	y clicking on each of act they have in table about the	SCIENCE NEWS Did Polynesian Rats Destroy Rapa Nui's Trees? By Michael Kimball, Science Reporter		
The Theories			In 2005, Jared Diamond published Collapse. In the book, he described the human		
Cause Effect Supporters of the Theory		Supporters of the Theory	settlement of Rapa Nui (also called Easter Island).		
		Jared Diamond	questioned Diamond's theory of what happened on Rapa Nui. They agreed that the huge trees had disappeared by the time Europeans first arrived on the island in the 18 th century, but they did not agree with Jared Diamond's theory about the cause of the disappearance.		
		Carl Lipo and Terry Hunt	Now, two scientists, Carl Lipo and Terry Hunt, have published a new theory. They believe that the Polynesian rat ate the seeds of the trees, preventing new ones from growing. The rat, they believe, was brought over either accidentally or purposefully on the canoes that the first human settlers used to land on Rapa Nui.		
The moai were carved in the same quarry.	Polynesian rats ate tree seeds and as a result no new trees could grow.	Settlers used canoes to bring Polynesian rats to Rapa Nui.	Studies have shown that a population of rats can double every 47 days. That's a lot of rats to feed. To support their theory, Lipo and Hunt point to the remains of palm nuts that show the gnaw marks made by rats. Of course, they acknowledge that humans did play a role in the destruction of the forests of Rapa Nui. But they believe that the Polynesian rat was an even greater culprit among a series of factors.		
The large trees disappeared from Rapa Nui.	Rapa Nui residents needed natural resources to move the moai.	Humans cut down trees to clear land for agriculture and other reasons.			

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.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.



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



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.

Rapa Nui Question 3 / 7			← → Q www.academicbookreview.com/Collapse		
Refer to the Review of Collapse on the rig shoices in the table to answer the question isted below are statements from the Review Are these statements facts or opinions? C Fact or Opinion for each statement.	ht. Clic n. iew of lick on	ck on the Collapse. either	Jared Diamond's new book, <i>Collapse</i> , is a clear warning about the consequences of demond's new book, <i>Collapse</i> , is a clear warning about the consequences		
Is the statement a fact or an opinion?	Fact	Opinion	civilizations that collapsed because of the choices they made and their impact on the environment. One of the most disturbing examples in the book is Rapa Nui.		
In the book, the author describes several civilizations that collapsed because of the choices they made and their impact on the environment.	ö	Ö	According to the author, Rapa Nui was settled by Polynesians sometime after CE. They developed a thriving society of, perhaps, 15 000 people. They can the moai, the famous statues, and used the natural resources available to the move these huge moai to different locations around the island. When the first Europeans landed on Rapa Nui in 1722, the moai were still there, but the tree		
One of the most disturbing examples in the book is Rapa Nui.	0	0	were gone. The population was down to a few thousand people who were struggling to survive. Mr. Diamond writes that the people of Rapa Nui cleared the		
They carved the moai, the famous statues, and used the natural resources available to them to move these huge moai to different locations around the island.	0	0	land for farming and other purposes and that they over-hunted the numerous species of sea and land birds that had lived on the island. He speculates that the dwindling natural resources led to civil wars and the collapse of Rapa Nui's society.		
When the first Europeans landed on Easter Island in 1722, the moai were still there, but the trees were gone.	0	0	The lesson of this wonderful but frightening book is that in the past, humans made the choice to destroy their environment by cutting down all the trees and hunting animal species to extinction. Optimistically, the author points out, we can choose not to make the same mistakes today. The book is written well and deserves to by		
The book is written well and deserves to be read by anyone who is concerned about the environment.	0	0	read by anyone who is concerned about the environment.		

Exhibit D.4. Example of a Grid Selected Response Item from PISA's Rapa Nui 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.5. Example of a Zones Item for an Internet Text from ePIRLS' "Elizabeth Blackwell" 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.6. Example of an In-line Choice Item from ePIRLS' Mars Block That Also Collects Process Data on Where Students Click on the Web Page



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



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.8. Example of a Fill in the Blank Item Response Format from ePIRLS' Mars 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.

APPENDIX E: ADDITIONAL EXAMPLES OF READING PURPOSES AND UDES

Exhibit E.1. Example of a Specific Reading Purpose and a Informational UDE from PISA's Rapa Nui 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.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.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.3. Example of a Motivational UDE, from NAEP's "Tough as Daisy" Block

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

Example 1



Example 2



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.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.

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.

- American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME), (2014). *Standards for educational and psychological testing*. Washington, DC: Author.
- Afflerbach, P., Hurt, M., & Cho, B. (2020). Reading comprehension strategy instruction. In D. L. Dinsmore, L. A. Fryer, & M. M. Parkinson (Eds.), *Handbook of strategies and strategic* processing (pp. 99-118). New York: Routledge.
- Afflerbach, P., Pearson, P. D., & Paris, S. G. (2007). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, 61(5), 364-373, doi.org/10.1598/RT.61.5.1
- Aguilar, G., Uccelli, P., & Galloway, E. P. (2020). Toward biliteracy: Unpacking the contributions of mid-adolescent dual language learners' Spanish and English academic language skills to English reading comprehension. *TESOL Quarterly* (early view), DOI: 10.1002/tesq.570
- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, *32*(4), 665-683, doi.org/10.1111/j.1559-1816.2002.tb00236
- Alexander, P. A., & Grossnickle, E. M. (2016). Positioning interest and curiosity within a model of academic development. In K. Wentzel & D. Miele (Eds.), *Handbook of motivation at school* (2nd ed., pp. 188–208). New York: Routledge.
- Alexander, P. A., & 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). Mahwah, NJ: Erlbaum.
- 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
- 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.
- Alvermann, D. E. & Wilson, A. A. (2011). Comprehension strategy instruction for multimodal texts in science. *Theory Into Practice*, *50*(2), 116-124.
- Anderson, R. C. (2019). Role of the reader's schema in comprehension, learning and memory. In D.E. Alvermann, N.J. Unrau, & R.B. Ruddell (Eds.), *Theoretical models and processes of literacy* (pp. 136-145). New York: Routledge.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, P. Mosenthal (Eds.), *Handbook* of Reading Research (Vol. I) (pp. 255-291). New York: Routledge.
- Appleman, D. (2017). Literary theory in the secondary school. In *English Language Arts Research and Teaching* (pp. 172-188). Routledge.
- Armbruster, B. B. (1984). Learning from content area textbooks: The problem of "inconsiderate text." In G. Duffy, L. Roehler, & J. Mason (Eds.), Comprehension instruction (pp. 202-217). New York, NY: Longman.
- 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). Mahwah, NJ: Lawrence Erlbaum Associates.
- Barron, N. (2015). *Words onscreen: The fate of reading in a digital world*. New York: 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). Hershey, PA: 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. V) (pp. 79-98). New York: 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.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press. (2nd. ed.). New York: 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*. New York, NY: 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.). Portsmouth, NH: 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. V) (pp. 237-260). New York: 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). New York, NY: 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). New York, NY: 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

- Council of Chief State School Officers and National Governors Association Center for Best Practices (2010). *Common core state standards for English language arts and literacy in history/social studies & science*. <u>http://www.corestandards.org/Standards/K12/</u>
- 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.), New York: 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). Mahwah, NJ: Lawrence Erlbaum Publishers.
- de Jong, T. (2006). Scaffolds for scientific discovery learning. Handling complexity in learning environments: Theory and research, 107-128.
- 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. <u>https://www.ksde.org/Portals/0/CSAS/CSAS%20Home/CTE%20Home/Instructor_Resource/FinalRealWorldReading.pdf</u>
- Durán, R. P. (December, 2006). State Implementation of NCLB Policies and Interpretation of the NAEP Performance of English Language Learners. Washington, DC.: *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, et al. (2005). Ability self-perceptions and subjective task values in adolescents and children. DOI: 10.1007/0-387-23823-9_15.
- Educational Testing Service. (January 22, 2019). *NAEP Reading Special Study: Scenario-Based Task and Discrete Task Report. Deliverable in Response to ID Task 9.2.1.* Princeton, NJ: Author. (Note: This report is not available to the general public because it contains secure information.)
- Every Student Succeeds Act, 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 aanalysis. *Journal of Adolescent & Adult Literacy*, 53, 587–597.
- Fitzgerald, M., Higgs, J., & Palincsar, A. (2020). *New media, new literacies: Implications for reading for understanding*. Washington, DC: 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. (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. <u>http://dx.doi.org/10.1080/00461520.2016.1168741</u>.
- Goldman, S. R., Greenleaf, C., Yukhymenko-Lescroart, M., Brown, W., Ko, M. L. M., Emig, J. M., ... & 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 SR, and Pellegrino JW. (2015). Research on Learning and Instruction: Implications for Curriculum, Instruction, and Assessment. Policy Insights from the Behavioral and Brain Sciences. 2015;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.
- Guthrie, J. T., Klauda, S. L. (2016). Engagement and motivational processes in reading. In Afflerbach, P. (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 41–53). Routledge.
- Guthrie, J. T., & Humenick, N. M. (2004). Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement.
- Guthrie, J. T., Wigfield, A., & You, W. (2012). Instructional contexts for engagement and achievement in reading. In S. Christensen, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 601-635). New York: Springer Science
- Guthrie, J. T., Wigfield, A., Von Secker, C. (2000) Effects of integrated instruction on motivation and strategy use in reading. Journal of Educational Psychology, 92(2), 331-341.
- Hartman, D. K., Morsink, P. M., & Zheng, J. (2010). From print to pixels: The evolution of cognitive conceptions of reading comprehension. *The new literacies: Multiple perspectives on research and practice*, 131-164.
- Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms.* Cambridge: Cambridge University Press.
- Heath, S.B. (2012). Words at work and play: Three decades in family and community life. Cambridge, UK: Cambridge University Press.
- Ho, A. (2017). Advancing Educational Research and Student Privacy in the "Big Data" Era. Washington, DC: National Academy of Education.

- 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.
- IRA/NCTE Joint Task Force on Assessment, International Reading Association, & National Council of Teachers of English. (2010). *Standards for the assessment of reading and writing*. International Reading Association and National Council of Teachers of English.
- International Test Commission (ITC) (2019) ITC Guidelines for the Large-Scale Assessment of Linguistically and Culturally Diverse Populations, International Journal of Testing, 19:4, 301-336, DOI: <u>10.1080/15305058.2019.1631024</u>
- Jaeger, Richard M. (2003). *Reporting the Results of the National Assessment of Educational Progress. NAEP Validity Study. Working Paper Series.* Available only online at: <u>https://nces.ed.gov/pubs2003/200311.pdf</u>.
- Johnstone, C. (2003). Improving Validity of Large-Scale Tests: Universal Design and Student Performance (NCEO Technical Report).
- 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. In *National Center on Educational Outcomes, University of Minnesota*. National Center on Educational Outcomes.
- Katz, M.L., Brynelson, N., and Edlund, J.R. (2019). "Enacting Rhetorical Literacies: The Expository Reading and Writing Curriculum in Theory and Practice." seventh edition, *Theoretical Models and Processes of Literacy*, D. Alvermann, N. Unrau, and M. Sailors (Eds.). New York: Routledge (pp. 533-562).
- Kendeou, P., & O' Brien, E. J. (2016). Prior Knowledge: Acquisition and revision. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Text and context* (pp. 151-163). New York: Routledge.
- Kendeou, P., Van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, 29(1), 10-16.
- Kieffer, M. J., & Thompson, K. D. (2018). Hidden Progress of Multilingual Students on NAEP. *Educational Researcher*, 47, pp. 391–398. doi: 10.3102/0013189X18777740
- Kincaid, J. P., Fishburne Jr, R. P., 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, 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). Mahwah, NJ: Erlbaum.
- Kress, G. (2013). Recognizing learning. In I. deSaint-Georges & J. J. Weber (Eds.), Multilingualism and multimodality (pp. 119–140). Rotterdam, ND: Sense Publishers.

- Lee, C. D. (1993). Signifying as a scaffold for literary interpretation: The pedagogical implications of an African American discourse genre. Urbana, IL, National Council of Teachers of English.
- Lee, C. D. (2005). Culture and language: Bi-dialectical issues in literacy. In P. L. Anders & J. Flood (Eds.), *The literacy development of students in urban schools*. Newark, DE: International Reading Association.
- Lee, C. D. (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*. NY, Routledge: 286-304.
- Lee, C. D. (2017). An ecological framework for enacting culturally sustaining pedagogy. *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*, 261-273.
- Lee, C. D. (2020). Social and cultural diversity as lens for understanding student learning and the development of reading. In E. B. Moje, P. Afflerbach, N. K. Lesaux, & P. Enciso (Eds.), *Handbook of reading research, Volume V* (37-56). New York: Routledge.
- Lee, C. D., 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). New York, NY: Routledge.
- Lee, C. D. and A. Spratley (2010). *Reading in the disciplines and the challenges of adolescent literacy*. NY, 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). London, UK: Routledge.
- Manderino, M. (2012). Disciplinary literacy in new literacies environments: Expanding the intersections of literate practice for adolescents. In P. Dunston, L. Gambrell, K. Headley, S. Fullerton, & P. Stecker (Eds.), *Sixty-first yearbook of the literacy research association* (pp. 69–83). Oak Creek, WI: Literacy Research Association.
- 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*, 1-19.
- McNamara, D. S. and W. Kintsch (1996). "Learning from text: Effects of prior knowledge and text coherence." <u>Discourse Processes</u> 22: 247-287.
- Measured Progress/ETS Collaborative. (2012). Smarter Balanced Assessment Consortium: Technology-enhanced item guidelines. <u>https://www.measuredprogress.org/wp-</u> 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. (2009). A call for new research on new and multi-literacies. *Research in the Teaching of English*, 43(4), 348-362.
- Moje, E. B. (2015). Doing and teaching disciplinary literacy with adolescent learners: A social and cultural enterprise. *Harvard Educational Review*, 85(2), 254–278. doi:10.17763/0017-8055.85.2.254
- Moje, P. Afflerbach, N. K. Lesaux, & P. Enciso (Eds.), *Handbook of reading research, Volume* V (37-56). New York: Routledge.
- 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. Boston, MA: International Association for the Evaluation of Education Achievement, & TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. http://pirls2021.org/wp-content/uploads/sites/2/2019/04/P21_Frameworks.pdf
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *ePIRLS 2016 International Results in Online Informational Reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <u>http://timssandpirls.bc.edu/pirls2016/international-results/</u>
- 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. Washington, DC: The National Academies Press. <u>https://doi.org/10.17226/23409</u>.
- National Academies of Sciences, Engineering, and Medicine (NASEM). (2018). How People Learn II: Learners, Contexts, and Cultures. Washington, DC: The National Academies Press. <u>https://doi.org/10.17226/2478</u>3.National Academies of Sciences, Engineering, and Medicine (NASEM). 2019. Monitoring Educational Equity. Washington, DC: The National Academies Press. <u>https://doi.org/10.17226/25389</u>.
- National Assessment Governing Board. (1992). NAEP Reading Consensus Project. *Reading framework for the 1992 National Assessment of Educational Progress* (Report No. 92-5002). Washington, DC: National Assessment Governing Board.
- National Assessment Governing Board. (2009). *Reading framework for the 2009 National Assessment of Educational Progress*. Washington, DC: 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. Retrieved from <u>https://www.nagb.gov/content/dam/nagb/en/documents/policies/naep_testandreport_studentswithdisabilities.pdf</u>
- National Assessment Governing Board. (2017). *Reading framework for the 2017 National Assessment of Educational Progress.* Washington, DC: National Assessment Governing Board and U.S. Department of Education.

https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/reading/2 017-reading-framework.pdf
- National Assessment Governing Board. (2018). Framework development policy statement. Washington, DC: Author, adopted March 3, 2018.
- National Assessment Governing Board. (2019). *Reading framework for the 2019 National Assessment of Educational Progress*. Washington, DC: National Assessment Governing Board and U.S. Department of Education.
- National Center for Education and the Economy [NCEE]. (2013). What does it really mean to be college and workforce ready? The English literacy required of first year community college students. Washington, DC: NCEE. Retrieved June 20, 2018 from http://ncee.org/wp-content/uploads/2013/05/NCEE _EnglishReport_May2013.pdf.
- National Center for Education Statistics. (2018). Going digital: NAEP assessments for the future. *National Assessment of Educational Progress*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences.
- https://nces.ed.gov/nationsreportcard/subject/dba/pdfs/2018_dba_brochure.pdf National Center for Education Statistics. (2019). NAEP Report Card: 2019 Reading Assessment.
- https://www.nationsreportcard.gov/highlights/reading/2019/.
- 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 Center for Education Statistics. (2020). Digitally based assessments. *National Assessment of Educational Progress*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences. <u>https://nces.ed.gov/nationsreportcard/dba/</u>
- National Council for the Social Studies. (2013). 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.
- 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*, Vol. 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
- OECD. (2019). PISA 2018 assessment and analytical framework. Paris: 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* (p. 135-165). New York: Routledge.

- Pacheco, M. (2018). Learning and becoming writers: Meaning, identity, and epistemology in a newsroom community of practice. *Mind, Culture, and Activity*, 25(2), 105-124.
- Paris, S. G., Wasik, B. A., & Turner, J. C. (1991). The Development of Strategies of Readers. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research* (Vol. 2, pp. 609-640). Mahwah, NJ: Lawrence Erlbaum Associates.
- Pearson, P. D., & Cervetti, G. N. (2015). Fifty years of reading comprehension theory and practice. In P. D. Pearson, E. H. Hiebert, N. K. Duke (Eds.), *Research-based practices for teaching Common Core literacy* (pp. 1-24). New York: Teachers College Press.
- Pearson, P. D., & Hiebert, E. (2014). The state of the field. *The Elementary School Journal*, *115*(2), 161-183.
- Pearson, P. D., Hansen, J., & Gordon (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., Palincsar, A. S., Biancarosa, G., & Berman, A. (2020). Reaping the rewards of the Reading for Understanding Initiative. Washington, DC: 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. *The neurocognition of language*, *167*, 208.
- 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. <u>Special Issue: The Science of Reading: Supports, Critiques, and</u> Questions, *Reading Research Quarterly*, 55 (1), 331-345. doi.org/10.1002/rrq.359
- Pintrich, P. R. and B. Schrauben (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. <u>Student perceptions in the classroom: Causes and consequences D</u>. Schunk and J. Meece. Hillsdale, NJ, Erlbaum: 149-183.
- Pitzer, J & Skinner, E. (2017). Predictors of changes in students' motivational resilience over the school year: The roles of teacher support, self-appraisals, and emotional reactivity. *International Journal of Behavioral Development*, 41(1) 15–29.
- Rabinowitz, P. J. (1987). *Before reading: Narrative conventions and the politics of interpretation*. Ithaca, NY: Cornell University Press.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND. https://www.rand.org/pubs/monograph_reports/MR1465.html
- 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. (2019): Engineering a twenty-first century reading comprehension assessment system utilizing scenario-based assessment techniques. *International Journal of Testing*, DOI:10.1080/15305058.2018.1551224

- 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). New York, NY: Routledge.
- Serafini, F., & Gee, E. (Eds.). (2017). *Remixing multiliteracies: Theory and practice from New London to new times*. Teachers College Press.
- Scalise, K. and 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. and Cole, M (1981). *The Psychology of Literacy*. Cambridge, Mass.: Harvard University Press.
- 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, and N. Lesaux. (Eds.). *Handbook of reading research* (Vol. V) (pp. 328-344). New York, NY: 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 <u>http://sampleitems.smarterbalanced.org</u>
- 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). New York: Routledge.
- Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 38(5), 553-573.
- Sparks, J. R., & Deane, P. (2014). *Cognitively-based assessment of research and inquiry skills:* Defining a key practice in the English language arts. Manuscript in preparation.
- Stenner, A.J. (1996). Measuring reading comprehension with the Lexile framework. Durham, NC: Metametrics Inc.
- Street, B. V. (1984). Literacy in theory and practice. NY: Cambridge University Press.

- Taft, M. L., & Leslie, L. (1985). The effects of prior knowledge and oral reading accuracy on miscues and comprehension. *Journal of Reading Behavior*, *17*(2), 163-179.
- Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). Universal design applied to large scale assessments. *Minneapolis: 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 Hashimoto, sixth 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. Washington, D.C.: 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., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (p. 123–139). IAP Information Age Publishing.
- van den Broek, P., & 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. *Models of understanding text*, 165-187.
- van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. Science, 328, 453–456.
- 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.
- 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.
- 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). 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.

Across the Board: Understanding Recent NAEP Results

Over 2018 and 2019, the NAEP program assessed students in nearly every major subject and across all three NAEP grades. In 2018, NAEP assessed eighth graders in civics, geography, and U.S. history. In 2019, NAEP assessed students in fourth, eighth, and twelfth grades in mathematics, reading, and science. The results revealed a growing divergence in knowledge and skills between the nation's highest and lowest achievement students over the last decade (see this discussion of the divergence in mathematics and reading at grades 4 and 8).

NCES will present a summary of key findings from the 2009-2019 NAEP assessments vis-à-vis diverging trends and performance changes across student groups and geographic areas. Following the presentation, Board members will discuss reactions to the key findings and steps the Board might consider to (a) better inform education stakeholders about these cross-cutting trends, and (b) deepen understanding/insights into these patterns (e.g., by commissioning special research studies, etc.).

Overview of the NAEP Technology and Engineering Literacy (TEL) Assessment

Background

In March 2010, the Governing Board adopted the <u>NAEP Technology and Engineering Literacy</u> (<u>TEL</u>) Framework; the <u>TEL assessment</u> was administered to a national sample of eighth-grade students in 2014 and 2018 and is next <u>scheduled</u> to be administered in 2024. The framework calls for the use of many scenario-based tasks (SBTs) in addition to discrete (stand-alone) items.

TEL measures students' knowledge and skills in three interconnected areas: Technology and Society, Design and Systems, and Information and Communications Technology. There are three cross-cutting practices as well: Understanding Technological Principles, Developing Solutions and Achieving Goals, and Communicating and Collaborating. The framework defines literacy as the level of knowledge and competencies about technology and engineering needed by all students and citizens to function in a technological society. The framework defines technology and engineering literacy as "the capacity to use, understand, and evaluate technology as well as to understand technological principles and strategies needed to develop solutions and achieve goals."

Relating to national efforts in science, technology, engineering, and mathematics (STEM) fields, the NAEP Technology and Engineering Literacy Assessment was developed to measure the "T" and "E" in STEM, augmenting the long-standing NAEP assessments in science and mathematics. The NAEP Science Framework, last updated in November 2005 for implementation in the 2009 assessment, does not measure technology or engineering. The decision not to include technology and engineering in the NAEP Science Framework adopted in November 2005 was one of the rationales for creating a separate NAEP TEL Framework. The development of both of these frameworks preceded the release of the Next Generation Science Standards (NGSS) in 2013. The NGSS includes both science and engineering.

Overview

The purpose of this session at the August Board meeting is to provide background on the TEL Framework and assessment to serve as a foundation for upcoming Board decisions on the NAEP Assessment Schedule and the NAEP Science Framework. Presenters will include:

- Assessment Development Committee member Christine Cunningham will introduce the TEL Framework and discuss current considerations for TEL in the context of the upcoming NAEP Science Framework update; and
- William Ward of the National Center for Education Statistics will provide an overview of the TEL assessment and current operational issues and challenges.

GROUND TRANSPORTATION OPTIONS NAGB Quarterly Board Meeting – August 5-6, 2021 Ritz Carlton Tysons Corner 1700 Tysons Blvd McLean, VA 22102 (703) 506-4300

App Based Ride Services

BWI Thurgood Marshall Airport (BWI)

App-based ride services pick up and drop off passengers at the terminal curbs on the Departures/Ticketing Level between doors 9 and 11.

Dulles International Airport (IAD)

Passenger pick-up is located on the ground level outside of Baggage Claim, accessible via Doors 2, 4 and 6. Your driver will communicate the specific arrival door via in-app messaging.

Ronald Reagan National Airport (DCA)

Private vehicle pick-up is located on the third (outer) curb outside Terminal A and on the second (outer) curb outside Terminal B/C Baggage Claim (arrivals level). Passengers coordinate directly with the driver.

Shared Ride Service

Super Shuttle provides shared ride service to and from BWI Thurgood Marshall Airport (BWI). For pick up, claim your luggage, and proceed to Ground Transportation/Shared Ride Vans. Reservations are required for all transportation services to and from the hotel and should be booked 24-48 hour in advance. Go here to book on-line or call 1-800 BLUEVAN/ (800) 258-3826. The one-way fare is approximately \$96.00. Super Shuttle does not provide service to the hotel from Dulles International Airport (IAD) and Ronald Reagan National Airport (DCA).

Taxi Service

Arrivals and Departures via BWI Thurgood Marshall and Ronald Reagan National Airports

Several taxi companies provide service from BWI Thurgood Marshall Airport (BWI) and Ronald Reagan National Airport (DCA). The one-way trip from BWI takes approximately one hour and the fare is approximately \$100 - \$120. The one-way fare from Reagan is approximately \$30 and travel time is approximately 20 minutes. Taxi stands are located outside the airport and hotel.

Arrivals and Departures via Dulles International Airport

Washington Flyer Taxi Service (703) 572-8294 provides exclusive taxi service from Dulles International Airport. The oneway fare is approximately \$50 per person and travel time is approximately 25 minutes. Upon arrival at Dulles, proceed to the baggage claim/arrivals area on the lower level of the main terminal and proceed to the Washington Flyer taxi stand at Door 2 or 6. A customer service representative will assist you with coordinating service.

Public Transportation-Metrorail

Ritz Carlton Tysons is accessible by Metrorail via the Tysons Corner Metro station via the Silver line. Exit the Tysons Corner station and walk west on Chain Bridge Rd/VA-123 toward Tysons Blvd. Turn right onto Tysons Blvd. Keep left at the fork to continue onto Tysons Blvd to the Ritz. The walk takes approximately 7 minutes and is convenient if you have a small carry-on.

Parking

Valet parking is available in the hotel's parking garage at a rate of \$25 per day, \$45 per night. Valet hours are from 7:00 am until 11:00 pm. Daily self-parking is available at a rate of \$12.00. No in and out privileges.

NATIONAL ASSESSMENT GOVERNING BOARD											
Travel Expense Report											
Name:						FY 2021 Per diem rates (DC)					
Trip Purpo	se:	August 2021 Quarterly Board Meeting				Daily Per Diem		\$76.00			
Start on						Meal & Incidental Expense Breakdown		\$18 Breakfast, \$19 Lunch, \$34 Dinner, \$5 Incidentals			
inp dutes	End on					Expenses at a glance					
Notes						Other Expenses, (e.g. tolls, Internet access)		\$0.00			
						TOTAL EXPENSES		\$0.00			
DATE	From (trip origin)	To (destination)	Airline/Train (if purchased own)	Lodging (if not pre-paid)	Per Diem (based on meals provided)	Taxis	# of miles	Total mileage (0.56/mile)	Parking	Other expenses (enter description in line)	TOTAL
Arrival			\$0.00	Pre-paid							
Return											
Detail Other Expenses											
						\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00
						\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00
						\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00
						\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00
						\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00
						\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00
						\$0.00	0	\$0.00	\$0.00	\$0.00	\$0.00
			\$0.00	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00