2028 NAEP Science Assessment Framework

Contract # 91995922C0001

THE 2023–2024 NATIONAL ASSESSMENT GOVERNING BOARD

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

Honorable Beverly Perdue, Chair
Former Governor – Democrat
North Carolina
New Bern, North Carolina

Representative Alice Peisch, Vice Chair
State Legislator – Democrat
Massachusetts House of Representatives
Wellesley, Massachusetts

Lisa Ashe
Curriculum Specialist
Mathematics Consultant
North Carolina Dept. of Public Instruction
Wake Forest, North Carolina

Shari Camhi
Local Superintendent
Superintendent of Schools
Melville, New York

Michelle Cantú-Wilson
General Public Representative (Generalist)
Creative Director & Speaker
Marquee Consulting, Inc.
Houston, Texas

Tyler W. Cramer
General Public Representative (Generalist)
CEO and Executive Manager
Remarc Associates LLC
San Diego, California

Christine Cunningham
Curriculum Specialist
Senior Vice President of STEM Learning
Museum of Science
Boston, Massachusetts

Viola Garcia
Local School Board Member
Aldine Independent School District
Houston, Texas

Angélica Infante-Green
Chief State School Officer
Commissioner of Elementary and Secondary Education
Rhode Island Department of Education
Providence, Rhode Island

Patrick L. Kelly
12th Grade Teacher
AP U.S. Government and Politics
Richland School District 2
Columbia, South Carolina

Anna King
General Public Representative (Parent Leader)
Former President - National PTA
Oklahoma City, Oklahoma

Suzanne Lane
Testing and Measurement Expert
Professor of Research Methodology
University of Pittsburgh
Pittsburgh, Pennsylvania

Scott Marion
Testing and Measurement Expert
President and Executive Director
The National Center for the Improvement
of Educational Assessment, Inc.
Dover, New Hampshire

Reginald McGregor
Business Representative
Vice President – Government Relations Rolls Royce
Corporation
Indianapolis, Indiana

Michael A. Pope
8th Grade Teacher
Zama American Middle School
Camp Zama
Department of Defense Education Activity
Sagamihara, Japan

Julia Rafal-Baer
General Public Representative (Parent Leader)
Managing Partner ILO Group
Cranston, Rhode Island

NATIONAL ASSESSMENT GOVERNING BOARD

For further information, contact:

National Assessment Governing Board

800 N. Capitol St., NW

Suite 825

Washington, DC 20002-4233

www.nagb.gov

January 2024

The 2028 NAEP Science Framework was developed for the National Assessment Governing Board by WestEd under contract 91995922C0001.

This document includes descriptions of science and engineering practices, crosscutting concepts, and disciplinary core ideas from *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, and excerpts from *Next Generation Science Standards: For States, by States* and associated appendices, with permission granted by the National Academies Press, Washington, DC.

TABLE OF CONTENTS

TABLE OF EXHIBITS

ABOUT THE NATION'S REPORT CARD

What Is NAEP?

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 science, mathematics, reading, and several other subjects. NAEP results enable comparisons of what sampled students know and are able to do among states and jurisdictions, among various demographic groups, and over time. By law and by design, NAEP does not produce results for individual students or schools. NAEP scores are always reported at the aggregate level.

In 1988, Congress created the National Assessment Governing Board (Governing Board) as an independent, nonpartisan organization responsible for setting policy for NAEP. The 26 members of the Governing Board include governors, state legislators, state and local school officials, educators, researchers, business representatives, and members of the general public who are appointed by the U.S. secretary of education. Development, administration, scoring, and reporting of the NAEP assessment are carried out by the National Center for Education Statistics (NCES), located within the U.S. Department of Education's Institute of Education Sciences (IES).

As the ongoing national indicator of the academic achievement of U.S. students, NAEP regularly collects information on representative samples of students and periodically reports on student achievement in reading, mathematics, writing, science, and other subject areas. NAEP assessments are administered to students in grades 4, 8, and 12 at the national level and sometimes also for states and districts that volunteer to participate at the state level or in the Trial Urban District Assessment (TUDA) program.

The NAEP Authorization Act of 2002 (NAEP, P.L. 107-279) is the governing statute of NAEP. This law stipulates that NCES develops and administers NAEP and reports NAEP results. Under the law, the Governing Board's responsibilities include determining the assessment schedule, developing the assessment frameworks that provide the blueprints for the content and design of the assessments, and setting the achievement levels.

By law, NAEP assessments shall not evaluate personal beliefs or publicly disclose personally identifiable information, and NAEP assessment items shall be secular, neutral, nonideological and free from racial, cultural, gender, or regional bias. Although broad implications for academic subject matter may be inferred from the assessment, NAEP does not specify how any subject area should be taught; nor does it prescribe a particular curricular approach to teaching any subject.

The NAEP program is strongly committed to equity and advances this goal through the design, administration, and reporting of assessments that strive to be inclusive and accessible for all participating students. NAEP assessments align with current educational measurement standards[1] for fair and unbiased assessments. Through contextual questionnaires, NAEP gathers and reports data that may enhance understanding of factors related to differential student achievement.

NAEP data can be used as a tool for researchers and policymakers by providing reliable information on student achievement in reading, mathematics, science, and other subjects. The NAEP website provides subject-matter achievement results (as both scale scores and achievement levels) for various subgroups; results of surveys taken by students, teachers, and school leaders to provide information on contextual factors such as school facilities and teaching methods; the history of state and district participation; publicly released assessment questions; and scoring guides. The website also contains user-friendly data analysis software to enable access to all aspects of NAEP data, perform significance tests, and create customized graphic displays of NAEP results.

Frameworks and Specifications Documents

The development of a new or updated NAEP assessment begins with the creation of a framework that describes the subject matter to be assessed for students in grades 4, 8, and 12 and the assessment questions to be asked, as well as the assessment's design and administration. In accordance with Governing Board policy, a 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."

Each framework is accompanied by an item specifications document that serves as the "assessment blueprint" with additional information about item development. Unlike frameworks that are intended for a general audience, specifications documents are intended for a more technical audience, including NCES and the contractors who will develop the assessment items.

The broad-based process used in the development of the frameworks and specifications documents means that current thinking and research are reflected in the descriptions of what students should know and be able to do in a given subject. Therefore, these documents are frequently used as resources and models for the development of state assessments.

[1] 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.

EXECUTIVE SUMMARY

Introduction

The field of science is continuously evolving as new discoveries are made and more information becomes available: the emergence of artificial intelligence and other technological advancements; new vaccines and health care procedures; measuring the cosmos; and so much more. All of these scientific discoveries continue to have significant impacts on our daily lives. And all of them are grounded in science.

Having highly qualified individuals to continue making these discoveries is necessary for global development. These individuals all start as students in the classroom, so it is essential that today's students have the appropriate skills to foster their knowledge and understanding of the various science disciplines.

The 2028 NAEP Science Framework was updated to reflect the most current science information and practices. With these updates, policymakers, educators, school and district leaders, and other education stakeholders will have a better understanding of students' science knowledge and skills.

2028 NAEP Science Framework

Process Updates

The National Assessment Governing Board determines the assessment frameworks for NAEP based on recommendations from panels of content experts. As part of the development of the 2028 NAEP Science Framework, there were several changes made to the development process compared to past frameworks. These changes included:

- Multiple opportunities for input from science experts and public at the beginning of process
- Panel focus on substantive outline rather than narrative text
- Public comments on working draft of framework
- An Educator Advisory Committee in addition to a Technical Advisory Committee
- Panel leadership team rather than Panel Chair

The Governing Board had multiple opportunities for seeking input at the beginning of the process—both from science experts and the broader public. In addition to commissioning consultant papers from individuals and national science organizations, there was an initial public comment period on whether and how the existing framework should be changed. The Governing Board also worked with a strategic communications contractor to conduct in-depth interviews with several different stakeholders. These interviews provided essential insights from experts in the field, laying the groundwork for the framework revisions.

The Governing Board asked the panel to focus primarily on developing a substantive outline of the framework. This outline was the foundation of the working draft that the public was able to review and provide feedback on during the public comment period. The public comment period was held earlier in the process so that more time could be spent revising the document in response to important feedback. In order to solicit a wide array of feedback from the general public, with specific focus being on those in the fields of science, science education, and education, the Governing Board worked to ensure access to the public comment period and understanding of the working draft. During the public comment period, the Governing Board and panel members gave multiple presentations to various stakeholders and organizations, where they provided an overview of the process and recommended updates, while simultaneously requesting feedback from NCES.

An Educator Advisory Committee (EAC) was assembled for the first time to increase input and participation from teachers and other non-classroom educators in science. Finally, there was a panel leadership team of four Development Panel members instead of a single Panel Chair.

Content Updates

The National Research Council (NRC) of the National Academy of Sciences developed *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Concepts*, which was first released in July 2011. This framework provided an evidence-based foundation for assessment standards due to its integration of current scientific research and the three interrelated and equally important dimensions:

- Disciplinary Core Ideas
- Science and Engineering Practices
- Crosscutting Concepts

In order to update the 2028 NAEP Science Framework to reflect the more sophisticated expectations of understanding scientific ideas and practice reflected in most state standards, the Steering Panel formed to provide direction for the framework changes made the following initial set of recommendations:

- Update the construct of science to be assessed
- Update NAEP Science disciplinary concepts and practices, and add crosscutting concepts
- Expand the science construct to include aspects of technology and engineering
- Describe how the NAEP Science Assessment should assess the three dimensions of science
- Describe how student performance should be reported in light of science-specific contextual variables

Summary of Changes in the 2028 NAEP Science Framework*

Topic	Change
NAEP science construct	The framework defines the construct of science achievement and explains how this construct is operationalized using the three dimensions of science.
Three dimensions of science	NAEP science content has been redefined as any knowledge and reasoning skills that students need to know and be able to do on the NAEP Science Assessment. The content now includes the three dimensions of science: • NAEP Disciplinary Concepts • NAEP Science and Engineering Practices • NAEP Crosscutting Concepts
Technology and engineering	Technology and engineering concepts that are relevant to science achievement have been integrated into the updated science and engineering practices.
Assessment design	The framework calls for students to use the three dimensions of science.
	The framework provides expanded recommendations and guidance on the following: use of diverse tasks, phenomena, and contexts for items; and considerations for language complexity and cultural relevance. This framework also eliminates concept maps and hands-on tasks (HOTs).
	The framework calls for an even distribution of items across the three disciplines (Physical Science, Life Science, Earth and Space Sciences) across grades 4, 8, and 12.
	The framework includes a complexity framework to ensure that items are accessible to a wide range of learners.
Reporting results	Subscale reporting categories in Physical Science, Life Science, and Earth and Space Sciences have “Sensemaking in” added to each.
	Recommendations for science-specific contextual variables have been updated and prioritized.

**See Exhibit 1.1. in Chapter 1 for a comprehensive table of changes and the corresponding rationale.*

Conclusion

The 2028 NAEP Science Framework offers educators, state and district leaders, policymakers, and those in the science and business communities a guide to what students can and should know and do when it comes to science learning. With updated NAEP assessments informed by this framework, we will have a better understanding of what students currently understand about science and what they still need to know.

NAEP SCIENCE PROJECT STAFF AND COMMITTEES

The Science Framework for the National Assessment of Educational Progress is the result of extraordinary effort and commitment by hundreds of individuals across the country, including some of the nation's leading scientists, science educators, policymakers, and assessment experts. Every attempt has been made to be accurate in the description and representation of the science content. Existing framework development policies and procedures ensure periodic reviews and revisions, if necessary, to reflect advancements in scientific knowledge. Because this is a public endeavor, the Governing Board welcomes suggestions for future versions of the framework.

Steering Panel

* indicates the subgroup who drafted this framework as part of the Development Panel
^ indicates member of the Panel Leadership Team

Daniel Alcazar-Roman*
Associate Director, Center for K-12 Science
Lawrence Hall of Science
University of California, Berkeley
Berkeley, CA

Aneesha Badrinarayan*^
Director of State Performance Assessment Initiatives
Learning Policy Institute
Washington, DC

Tina Cheuk*
Assistant Professor California Polytechnic State University, San Luis Obispo
San Luis Obispo, CA

Jenny Ferrell Christian*^
STEM Director of K–12 Science and Wellness
Dallas Independent School District;
Council of the Great City Schools District Representative
Dallas, TX

Lakeitra Davis-Carter*
4th/5th Grade Teacher
Wilkinson County Elementary
Woodville, MS

Richard Duschl*
Executive Director of the Caruth Institute for Engineering Education
Southern Methodist University
Dallas, TX

Jennifer Greever*
TCAP Development Coordinator for Science
Tennessee Department of Education
Chattanooga, TN

Debra Hall
K-5 Science Consultant
North Carolina Department of Public Instruction
Raleigh, NC

Michael Heinz*
Science Coordinator
New Jersey Department of Education
Trenton, NJ

Kelley Hodges
Science Intervention Teacher
Patronis Elementary School
Panama City Beach, FL

Nancy Hopkins-Evans*^
Associate Director for Program Impact
Senior Science Educator
BSCS Science Learning
Wayne, PA

Joseph Krajcik*^
Lappan-Phillips Professor
Michigan State University College of Education and the College of Natural Science
East Lansing, MI

Okhee Lee
Professor of Childhood Education
New York University
New York, NY

Lawrence Lerner
Professor Emeritus
California State University, Long Beach
Woodside, CA

Ramon Lopez
Professor of Physics
University of Texas at Arlington
North Richland Hills, TX

Michael Lowry*
Science Department Chair
The McCallie School
Chattanooga, TN

Heather Morley*
High School Science Teacher
Champlain Valley Union High School, University of Vermont
Jericho, VT

Blessing Mupanduki*
Assessment Specialist
Research, Accountability, and Evaluations Division
Department of Defense Education Activity (DoDEA) Headquarters
Alexandria, VA

Tiffany Neill*
Research Scientist
University of Washington, Institute for Math + Science Education
Oklahoma City, OK

Jessica North*
8th Grade Science Teacher
Waunakee Community School District
DeForest, WI

Martin Osae
Middle School Teacher
West Dallas Junior High School (Dallas ISD)
Dallas, TX

Eric J. Pyle*
Professor, Geoscience Education
Past President, National Science Teaching Association
James Madison University
Harrisonburg, VA

Helen Quinn*
Professor of Particle Physics and Astrophysics, Emerita
SLAC National Accelerator Center
Stanford, CA

Ashlyn Razzo Laudel
Senior Director of Science Achievement, 9-12
Achievement First
Greenwich, CT

Philip A. Reed*
Professor, STEM Education and Professional Studies
Old Dominion University
Norfolk, VA

Brian Reiser*
Professor, Learning Science
Northwestern University
Evanston, IL

Yvette Selby-Mohamadu*
Lifetime Member
National Society of Black Engineers
DC Professionals–Pre-College Programs
Washington, DC

Sharon Sikora
Director of Middle School Curriculum and High School Chemistry Teacher
Sacred Heart Schools, Atherton
Atherton, CA

Iris Wagstaff
Founder and Executive Director
Wagstaff STEM Solutions
Washington, DC

Jason Zimba
Executive Vice President and Chief Academic Officer, Mathematics
Amplify
New York, NY

Educator Advisory Committee

Jaclyn Austin
Instructional Facilitator, Secondary Science
Howard County Public School System
Ellicott City, MD

Mihir Datta
High School Teacher
Holmes County Consolidated School District
Lexington, MS

Kellie Finnie
Director of Curriculum and Innovation
Dearborn Heights School District 7
Dearborn Heights, MI

Genevieve Garcia
Middle School and High School Science Teacher
Kotlik School, Lower Yukon School District
Kotlik, AK

Francis Panion
Science Instructional Coach
Miami–Dade County Public Schools
Miami, FL

Karen Pollari
Elementary School Teacher
Sidney Public Schools
Sidney, MT

Nicolette Roque
District Secondary Science Specialist
Duval County Public Schools
Jacksonville, FL

Wade Whitehead
Teacher
Crystal Spring Elementary School
Roanoke City Public Schools
Roanoke, VA

Technical Advisory Committee

[#] appointed to the Governing Board and recused from TAC as of October 1, 2023

Catherine Close
Vice President, Psychometrics
Renaissance Learning
Madison, WI

Karla Egan
Principal
EdMetric LLC
Malta Bend, MO

Bonnie Hain
Sr. Director of Education Solutions
Derivita
Salt Lake City, UT

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

James Pellegrino
Professor Emeritus, Psychology
Founding Co-Director, Learning Sciences Research Institute
University of Illinois Chicago
Chicago, IL

Guillermo Solano-Flores[#]
Professor of Education
Stanford University Graduate School of Education
Stanford, CA

Governing Board Staff

Rebecca Norman Dvorak
Assistant Director for Psychometrics
National Assessment Governing Board
Washington, DC

Stephaan Harris
Assistant Director for Communications
National Assessment Governing Board
Washington, DC

Sharyn Rosenberg
Contracting Officer's Representative
Assistant Director for Assessment Development
National Assessment Governing Board
Washington, DC

Anthony White
Contracting Officer
Contract Specialist
National Assessment Governing Board
Washington, DC

WestEd Staff

Nicole Cervantes
Administrative Assistant
WestEd
Washington, DC

Molly Faulkner-Bond
Science Measurement Specialist
Senior Research Associate
WestEd
Washington, DC

Quintin Love
Measurement Specialist
Senior Research Associate
WestEd
Las Vegas, NV

Mark Loveland
Project Director
Senior Research Associate
WestEd
San Francisco, CA

Randy Mangubat
Administrative Assistant
WestEd
San Francisco, CA

Lydia Martinez Rivera
Science Measurement Specialist
Science Content Specialist
WestEd
San Antonio, TX

Taunya Nesin
Science Content Lead
STEM Networking and Partnership Director
WestEd
Washington, DC

Danielle Oberbeck
Science Content Team Coordinator
Senior Research Associate
WestEd
San Francisco, CA

Marianne Perie
Measurement Lead
Director, Assessment Research and Innovation
WestEd
Stilwell, KS

Christopher Ruperto
Fiscal Manager
Division Financial Manager
WestEd
San Francisco, CA

Megan Schneider
Project Manager
WestEd
Austin, TX

Steven Schneider
Senior Advisor
Senior Director of STEM Research and Entrepreneurship
WestEd
San Francisco, CA

Jennifer Childress Self
Science Content Specialist
Senior Research Associate
WestEd
Washington, DC

Matt Silberglitt
Science Measurement Specialist
Manager, Science Assessment
WestEd
Martinez, CA

Jill Wertheim
Science Content Specialist
Senior Research Associate
WestEd
Washington, DC

Consultant

Cary Sneider
Cary I. Sneider Consulting
Portland, OR

Contractor

Safal Partners
Houston, TX

Communication Consultants to the Governing Board

Kate Johnson
Partner
FINN Partners

Jacqui Lipson
Partner and NAGB Project Director
FINN Partners

Isabella Perales
Account Supervisor
FINN Partners

Marina Stenos
Senior Partner
Practice Manager and STEM Advocacy and Education Expert
FINN Partners

Leah Van Blaricom
Assistant Account Executive
FINN Partners

Lauren Empson
Contractor
Empson Communications

Note: This list of project staff, panels, and advisory committees reflects professional affiliations during the project period for framework development.

CHAPTER ONE: Overview

1A. A Brief History of NAEP Science

Science achievement data have been part of NAEP assessments since 1969. Science achievement results were reported 10 times between 1969 and 1999. In 2004, the Governing Board adopted a revised NAEP Science Framework to address the rapidly changing nature of science and science education, as well as advances in assessment methodologies. The updated NAEP science assessment was administered for the first time in 2009, and the same framework was the basis for NAEP science assessments in 2011, 2015, and 2019. The 2009–2024 NAEP Science Framework (NAGB, 2019), as now known, will remain the basis for the 2024 assessment.

1B. The 2028 NAEP Science Assessment Framework Development Process

The Governing Board in 2021 conducted a review of the current science framework to determine the need for updates to assessments in 2028 and beyond. In accordance with Board policy, the review included an open comment period and commissioned papers and discussions with science educators and experts. Based on this review and other relevant research, the Governing Board determined that the NAEP Science Assessment Framework needed to be updated.

In accordance with the Governing Board policy on Assessment Framework Development for NAEP, new and updated frameworks are developed by Steering and Development Panels consisting of educators, state and local school administrators, policymakers, researchers and technical experts, assessment specialists, and other content experts and users of assessment data.

The Steering Panel formulated high-level guidance about the state of the field and how to implement the Board charge. The Steering Panel includes 30 members, of whom 20 members extended their service as members of a Development Panel. The Development Panel represented the larger group as it worked with Governing Board staff and members to develop the framework and specifications documents.

The Governing Board conducted an open call for panelist nominations from mid-June through mid-July 2022. Extensive and targeted outreach was conducted to hundreds of stakeholder groups and individuals representing education, policy, industry, assessment, research, and other science-related areas. The Board evaluated applications to serve on the panels with the goal of constructing a diverse and representative panel of stakeholders. The following factors were prioritized: (a) individuals specifically nominated to represent a national organization with key constituencies; (b) panelist role; (c) experience and expertise overall and in the specific sub-content areas covered by the framework; (d) demographic characteristics, including race, gender, and geography; (e) previous knowledge and expertise with both Next Generation State Standards (NGSS) and non-NGSS state science standards; and (f) varied perspectives on issues relevant to the Board charge. Thirty individuals were invited and accepted to serve on the Steering and Development Panels.

The Governing Board unanimously adopted the following charge to the Steering and Development Panels that would subsequently be convened to develop an updated science framework:

- NAEP must account for greater convergence in state science standards but cannot endorse the standards of any particular state or group of states.
- NAEP should remain forward-looking and consider what students should know and be able to do in science to be successful in college and careers.
- Updates should consider whether the definition of student achievement in science needs to incorporate relevant aspects of the 2014 NAEP Technology and Engineering Literacy (TEL; NAGB, 2018) Framework.[2]
- Updates to the NAEP Science Assessment Framework should prioritize relevance, utility, and validity over the need to maintain trend lines, but continuing the trend lines is desirable if possible.
- Updates should balance the emphasis on content and practices to ensure that the measurement of skills does not occur in isolation from content knowledge.
- Updates should be bound by considerations of feasibility, including technical issues (i.e., ensuring that the framework can be operationalized), cost (e.g., accounting for scenario-based tasks being more expensive than other item types), and the NAEP legislation (including but not limited to the requirements for NAEP to be nonsectarian).
- Updates should support the development of assessment items reflective of students who have a wide range of knowledge and skills in science.

In July 2022, the Board awarded a contract to WestEd through a competitive bidding process to convene the panelists, conduct meetings, and assist in creating the new framework and item specifications documents. Additional assistance was provided by a Technical Advisory Committee (a group of six measurement experts who provided feedback on technical issues) and an Educator Advisory Committee (a group of eight science educators who provided feedback on issues particularly relevant to practitioners).

The panelists were tasked with developing a substantive outline of the framework that would invite public comment at an earlier stage of the process compared to prior science frameworks to allow ample time to address substantive feedback. Public comment was solicited from March to April 2023, and the framework has been revised in response to the feedback received. The Panel Leadership Team presented updates and engaged in discussion with the Governing Board at

[2] The National Assessment of Educational Progress in Technology and Engineering Literacy (NAEP TEL) is a computer-administered assessment that measures problem-solving abilities related to design and systems, the use of digital tools for collecting and communicating information, and students' understanding of issues related to technology and society. The NAEP TEL was administered to a nationally representative sample of 20,500 grade 8 students in 2014 and again to 15,400 grade 8 students in 2018.

every quarterly meeting beginning November 2022. The Board adopted the framework at the November 2023 quarterly meeting.

1C. The Changing Construct of Science Achievement

Although NAEP has measured science achievement since its inception in 1969, the definition of science achievement has changed considerably over the decades. A major purpose of the new science framework is to anticipate how K–12 science achievement should be defined—that is, how the construct of science achievement is to be operationalized for assessment—in 2028 and beyond. As stated in the Governing Board policy, "The framework shall determine the extent of the domain and the scope of the construct to be measured for each grade level in a NAEP assessment."

The definitions of science achievement in early NAEP assessments emphasized knowledge of scientific concepts and theories, and the interpretation of natural phenomena based on that knowledge. Gradually, capabilities of scientific inquiry, such as observation, inference, and experimentation, came to be recognized as valuable and measurable, and more recently the ability to apply science principles in understanding and improving technologies. These changes and others were reflected in two influential documents developed in the 1990s: *National Science Education Standards* (National Research Council [NRC], 1996) and *Benchmarks for Scientific Literacy* (American Association for the Advancement of Science [AAAS], 1993). These documents influenced the development of state science standards during the early 2000s, which were mandated by a federal law, passed with overwhelming bipartisan support in 2001, that came to be known as No Child Left Behind (NCLB). Consequently, the 2009–2024 NAEP Science Framework drew heavily from these two documents.

The 2009–2024 NAEP Science Framework called for students at all three grade levels to be assessed on their understanding of concepts and theories in physical science, life science, and Earth and space sciences, and for items to be constructed to demonstrate students' knowledge through four types of science practices: identifying science principles, using science principles, using science inquiry, and using technological design. The percentages of items in the 2009–2024 framework varied across the subscales. For example, recognizing that Earth and space sciences were rarely taught at the high school level, a larger percentage of items in Earth and Space Sciences was called for at the grade 8 level (40%) versus at the high school level (25%). And, in recognition of the fact that few students learned about technology and engineering in any grade, only 10 percent of items were to have students apply science through technological design.

Since the 2009-2024 NAEP Science Framework was developed for implementation in 2009, the practice of science education in the nation's schools has undergone yet another fundamental change, guided largely by the release of the seminal document *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Concepts* (NRC, 2012), which was developed by a blue-ribbon panel of scientists, engineers, educators, and researchers under the auspices of the National Research Council of the National Academy of Sciences. *A Framework* *for K-12 Science Education* (NRC Framework) provides a sound, evidence-based foundation for assessment standards by drawing on current scientific research—including research on the ways students learn science effectively—and identifies the science all K–12 students should know and be able to do. The NRC Framework includes several innovations, leading to a new definition of the construct of science achievement that includes:

- updates to disciplinary core ideas in Physical Science, Life Science, and Earth and Space Sciences as endpoints of instruction at grades 2, 5, 8, and 12.
- the introduction of crosscutting concepts that apply to nearly all fields of science and engineering.
- identification of specific and measurable practices common to science and engineering in place of the more amorphous inquiry skills.
- a call for the teaching and assessment of science to integrate all three dimensions of science—disciplinary core ideas, crosscutting concepts, and practices of science and engineering—to make sense of natural phenomena, and to solve challenging problems in real-world contexts.
- recognition of the interrelation among science, engineering, society, and the environment.

Over the past decade, a great majority of states have patterned their standards and assessments on the essential ideas and specific definitions in *A Framework for K-12 Science Education.* Although full adoption of the new ways of teaching has been slow to take effect (Banilower et al., 2018), the development of new curriculum materials and professional development programs aligned with the new standards continue to move the nation in a common direction.

1D. Science Achievement in the 2028 NAEP Science Framework

The Steering and Development Panels used *A Framework for K-12 Science Education* (NRC, 2012) as a foundational resource for recommendations for updates to the assessment construct for NAEP Science. Consistent with that document, the Panels defined the construct that the 2028 NAEP Science Framework will measure as follows:

> ***Science achievement*** *is the ability to use relevant disciplinary concepts (Physical Science, Life Science, Earth and Space Sciences), science and engineering practices, and crosscutting concepts to identify and address problems, make sense of phenomena, and evaluate information to make informed decisions.*

The Panels also identified the following claims that they wanted the new framework to be able to support. Students should be able to do the following:

- reason scientifically using disciplinary concepts in combination with science and engineering practices and crosscutting concepts.
- address problems in the natural and designed world.
- make sense of phenomena in the natural and designed world.
- evaluate information and make decisions.

Working with those claims led to a focus on sensemaking and the need for multiple dimensions, which are themes throughout the framework. Three-dimensional science assessments require students to make sense of phenomena and problems using disciplinary concepts, science and engineering practices, and crosscutting concepts. Built into the proposed construct is the idea of scientific sensemaking, an essential aspect of all test items on the 2028 NAEP Science Assessment. In contrast to items that measure a student's ability to recall rote knowledge, NAEP Science Assessment items will require students to engage in scientific sensemaking: actively applying disciplinary concepts, science and engineering practices, and crosscutting concepts to figure out a phenomenon or address a real-world problem. The role of phenomena and problems in scientific sensemaking, along with the lessons learned about how assessments elicit sensemaking, is illustrated in detail in Chapter 3 in Exhibit 3.4.

1E. Opportunity to Learn and an Expansive Understanding of Contextual Variables

NAEP testing is not intended to report results for individual students or even schools; rather, it is intended to evaluate the state of science learning across the nation and regions. The NAEP Science Assessments are intended to be administered to students at the national level and sometimes also for states and large urban school districts that volunteer to participate in the Trial Urban District Assessment (TUDA). This framework defines what should be measured about science learning and seeks also to place those results in the context of what can be gleaned about the students, their learning opportunities, and their personal contexts from student, teacher, and school administrator questionnaires. These factors are called contextual variables by NAEP. While NAEP student results are reported by categories required by law such as gender, race, English Learner status, or socioeconomic status, these variables are highly interrelated, and these categories alone do not tell us the full story of what the results are saying about the state of science learning in our nation's schools.

Opportunity to learn is generally understood to refer to inputs and processes that shape student achievement, including the school conditions; time and spaces devoted to science learning; teacher knowledge and perspectives about science learning; and the curriculum, instruction, and resources to which students have access. When opportunity to learn was first used as a concept, Carroll (1963, 1989) emphasized the time allowed for learning. For the past 60 years, the concept of opportunity to learn has continued to evolve, as have efforts to measure in-school opportunities to learn, with the majority of scholars focusing on the classroom as the unit of analysis and instruction as central. Research has documented, for example, the negative effects on achievement of policies and practices that are often found in schools serving children who live in poverty or have special needs, including an inadequate supply of science teachers with strong knowledge and skills, a tendency to offer few advanced science courses, and a common practice of tracking these students disproportionately into low-level courses that restrict their learning opportunities (e.g., Ferguson et al., 2007; Kohlhaas et al., 2010), all of which can be understood as instructional resources that shape what students learn.

In recent years, there has been significant research on science learning and on the conditions and contexts that affect it. Two NRC reports have summarized much of what is known in this domain, namely *Science and Engineering for Grades 6–12* (National Academies of Sciences, Engineering, and Medicine [NASEM], 2019) and *Science and Engineering in Preschool Through Elementary Grades* (Davis & Stephens, 2022). These reports have noted the historical tendencies to view science learning, particularly at the high school level, as in service to the production of scientists and engineers, and thus intended for a select group of students. They and other studies, including *A Framework for K-12 Science Education* (NRC, 2012), have argued that a strong science education is needed for all students as preparation for life and community membership in the world of today, where many personal and community decisions require everyone to be able to interpret and apply scientific ideas and practices in the context of their daily lives. The NAEP Science Assessment, along with contextual information about the experiences of the participating students, is intended to measure how well that need is being met. Contextual information is critical to interpreting its results. Priorities for science-specific contextual variables are included in Chapter 4.

1F. Challenges of Developing a NAEP Assessment

Once a framework is completed and approved by the Governing Board, the next step is for the National Center for Education Statistics to develop the assessment. Here we discuss three major challenges: (a) measurement constraints and the nature of the items included on the assessment, (b) time and resource constraints and how much can be assessed in NAEP, and (c) the timeline for the framework and the difficulty of developing a framework with the rapid explosion of knowledge in the Information Age. Although these challenges apply to assessments in all subjects, they are especially challenging in the area of science, due to the rapidly changing nature of the subject and wide diversity of potential topics. Each of these challenges is discussed below.

What NAEP Measures

The NAEP Science Framework is an assessment framework, not a curriculum framework. A curriculum framework is designed to inform instruction, to guide what is taught, and often, to guide how it is taught. It represents a very wide universe of learning outcomes from which teachers choose what and how they teach. An assessment framework provides content and measurement guidance to assessment developers who then create items that can be assessed within time and resource allocations. An assessment framework does not cover all relevant content for each grade level; some concepts, practices, and activities in school science are not suitable to be assessed on NAEP, although they may well be important components of a school curriculum. The content to be assessed by NAEP has been identified as (a) disciplinary concepts that are central to the physical, life, and Earth and space sciences, (b) science and engineering practices, and (c) crosscutting concepts that are valued by educators and the science and business communities.

Assessment experts on the Development Panel and staff of the Governing Board also considered the feasibility of recommendations. For example, hands-on performance tasks, which were called for in the 2009–2024 NAEP Science Framework, have been eliminated due to concerns about cost and feasibility of implementation. Some of the knowledge and skills that had previously been assessed by hands-on performance tasks will be assessed through other means, including scenario-based tasks administered on the digital platform.

Administration Conditions

NAEP is an "on-demand" assessment. It ascertains what students know and are able to do in a set amount of time (60 minutes per student) and with limited access to resources (e.g., students will not have feedback from peers and teachers or opportunities for reflection and revision). Assessment frameworks describe the content and format of the assessment and indicate what students should know and be able to do. Due to the conditions of the assessment administration, certain elements of science learning may not be feasible to measure or evaluate. The absence of extended inquiry in NAEP, for example, is not intended to signal its relative importance in the curriculum; there is not sufficient time during the assessment administration for individual students to engage in the entire scientific process, but the range of the scientific process is covered across the full sample of students and items.

Current and Future Standards and Curricula

The framework attempts to strike a balance between what students are likely to encounter in their curriculum and instruction now and in the near future. It is a significant challenge to write a framework for the future. Cutting-edge science research creates new knowledge and investigative practices at the intersection of disciplinary boundaries. For example, research on human and natural systems has generated new understanding about environmental science that is closely linked to knowledge generated in Physical Science, Life Science, and Earth and Space Sciences. Although the framework is organized into these traditional areas, features of current science research are woven throughout. Another example of burgeoning knowledge is the rapid development of technologies, such as the transformation of our energy infrastructure from fossil fuels to renewable resources; new developments in artificial intelligence; and more advanced tools to measure, observe, and model warning systems for storms, tornadoes, and hurricanes.

The framework is intended to be both forward-looking (in terms of the science content that will be of central importance in the future) and reflective (in terms of current school science instruction). Because it is impossible to predict with certainty the future of school science instruction, the choices made for this framework should be revisited in response to future developments.

A summary of the changes in the 2028 NAEP Science Framework, compared with the 2009–2024 NAEP Science Framework, is shown in Exhibit 1.1.

Exhibit 1.1. Summary of Changes in the 2028 NAEP Science Assessment

Topic	Change	Rationale
NAEP science construct	This framework defines the construct of science achievement and explains how this construct is operationalized using the three dimensions of science. 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.	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 toward measuring the intended construct or whether they might, in fact, be measuring something else.
Three dimensions of science	NAEP science “content” has been redefined as any knowledge and reasoning skills that students need to know and be able to do on the NAEP Science Assessment. The content now includes updated and renamed science content statements (now disciplinary concepts) and science practices (now science and engineering practices), along with the addition of crosscutting concepts. These are now referred to collectively as the “three dimensions of science.”	The 2009–2024 NAEP Science Framework organized what students should know and be able to do into two buckets: science content and science practices. Based on research presented in the NRC Framework, it is recommended that the science content covered on the NAEP Science Assessment now consist of science disciplinary concepts, science and engineering practices, and crosscutting concepts.
	NAEP Disciplinary Concepts are well-tested theories and explanations developed by scientists organized into three major groupings: Physical Science; Life Science; and Earth and Space Sciences.	While the science ideas are still organized into three broad disciplinary groupings, NAEP science content statements have been renamed “NAEP Disciplinary Concepts” and updated to reflect shifts in expectations evident from reviews of state and national standards, policy documents from leading professional organizations, and expectations for science achievement on U.S. and international assessments.

Topic	Change	Rationale
Three dimensions of science (continued)	NAEP Science and Engineering Practices describe the skills and knowledge necessary to develop scientific explanations of phenomena and to design engineering solutions to problems.	NAEP science practices have been renamed "Science and Engineering Practices" and updated to reflect shifts in expectations evident from reviews of state and national standards, policy documents from leading professional organizations, the social context of science, and expectations for science achievement on U.S. and international assessments.
	NAEP Crosscutting Concepts have been added to the NAEP science "content" and are defined as concepts used across all science disciplines that provide scientists and engineers, and thus also students, with tools for asking productive questions and organizing their thinking.	With the introduction of the NAEP Crosscutting Concepts, based on findings reported in research on science learning, the updated definition of science achievement now describes the need for an assessment that can provide evidence about what students know and are able to do with all three dimensions of science.
Technology and engineering	Technology concepts, engineering concepts, and measurement/observation tools that are relevant to science and engineering achievements have been integrated into the updated science and engineering practices.	The addition of technology and engineering concepts to NAEP Science reflect shifts in expectations evident from reviews of state and national standards, policy documents from leading professional organizations, and expectations for science achievement on U.S. and international assessments. The framework incorporates concepts that represent the overlap between the NRC Framework and the NAEP Framework for Technology and Engineering Literacy (TEL).
Assessment design	This framework calls for students to use the three dimensions of science. Assessment items should require students to bring the three dimensions of science together to engage with the item. Items, item sets, and scenario-based tasks should be three-dimensional whenever possible. No item will be one-dimensional.	With the updated definition of science achievement and the incorporation of the three dimensions of science, the assessment design should reflect the need for students to address all three dimensions in their demonstration of what they know and are able to do in science.

Topic	Change	Rationale
Assessment design (continued)	This framework provides expanded recommendations and guidance on the following: use of diverse tasks, phenomena, problems, and contexts for items; considerations for language complexity and cultural relevance. This framework also eliminates concept maps and hands-on tasks (HOTs).	NAEP assessment items should be reflective of students who have a wide range of knowledge, skills, and backgrounds. Feasibility is also a consideration, and HOTs are costly to administer and require additional personnel for implementation. Scenario-based tasks can address some of the same content as HOTs, but with easier administration and implementation.
	This framework calls for an even distribution of items across the three disciplines (Physical Science, Life Science, Earth and Space Sciences) across grades 4, 8, and 12.	Prior NAEP science assessment frameworks called for differing distribution levels (higher percentage of Earth and Space Sciences at grade 8, lower percentage of Earth and Space Sciences at grade 12) based on NAEP data regarding students' course-taking patterns. Recommended distributions reflect shifts in expectations evident from reviews of state and national standards, policy documents from leading professional organizations, and expectations for science achievement on U.S. and international assessments.
	This framework includes a complexity framework in Chapter 3. The purpose of the complexity framework is to inform item development so as to ensure that items are accessible to a wide range of learners.	The complexity framework will be applied to NAEP item development to reflect how complexity specifically scales within and across multidimensional science items. This, in part, guides the development of multidimensional items that assess the full range of student performance.
Reporting results	Subscale reporting categories in Physical Science, Life Science, and Earth and Space Sciences have "Sensemaking in" added to each.	With the updated definition of science achievement and the incorporation of the three dimensions of science, the reporting of results for NAEP Science should reflect the emphasis on student scientific sensemaking.

Topic	Change	Rationale
Reporting results (continued)	Recommendations for science-specific contextual variables have been updated and prioritized.	NAEP contextual variable survey items should be reflective of the changing nature of science instruction and opportunities for students to learn science.

Detailed lists of the NAEP Disciplinary Concepts, NAEP Science and Engineering Practices, and NAEP Crosscutting Concepts to be assessed in 2028 and beyond are the focus of Chapter 2. Explanations of how these dimensions are to be combined to create assessment items are included with examples in Chapter 3. The processes for scoring, analyzing, interpreting, and reporting on NAEP Science achievement and contextual variables are summarized in Chapter 4.

CHAPTER TWO: Dimensions of Science Achievement

The NAEP Science Steering and Development Panels defined the construct that the 2028 NAEP Science Framework will measure as follows:

> ***Science achievement*** *is the ability to use relevant disciplinary concepts, science and engineering practices, and crosscutting concepts to identify and address problems, make sense of phenomena, and evaluate information to make informed decisions.*

The dimensions of NAEP science achievement are the disciplinary concepts, science and engineering practices, and crosscutting concepts. While these dimensions largely reflect the corresponding dimensions in the NRC Framework, modifications have been made, resulting in the three dimensions of NAEP science achievement. These dimensions are defined as follows:

- **NAEP Disciplinary Concepts** (DCs) are well-tested theories and explanations developed by scientists and organized into three major disciplinary groupings: Physical Science, Life Science, and Earth and Space Sciences.
- **NAEP Science and Engineering Practices** (SEPs) are ways of working to develop scientific explanations of phenomena or design engineering solutions to problems.
- **NAEP Crosscutting Concepts** (CCCs) are ideas that are used across all science disciplines and provide scientists and engineers and thus also students with tools for applying their knowledge of science to new phenomena or problems.

Phenomena are observable real-world events that provide a setting for an item or set of items. Problems are challenges that arise from a human need or want. In the 2028 NAEP Science Framework, the term *problem* is used to describe a real-world issue that requires a designed solution; as such, it is an engineering problem. The use of multiple dimensions to make sense of phenomena and problems is the essence of authentic science achievement.

2A. NAEP Science Disciplinary Concepts

The core ideas assessed as the NAEP DCs are based on those recommended by the panel of distinguished scientists and educators who developed the NRC document *A Framework for K–12 Science Education*.

The NAEP DCs progress across the years of K–12 education. The exhibits below delineate the progressions across grades 4, 8, and 12. Similar NAEP DCs, presented at a growing level of sophistication, are grouped in rows of the tables. Some NAEP DCs have no entry at grade 4 because their development is expected to begin later in the sequences of learning used by most schools across the United States. Some NAEP DCs include a clarification and/or a boundary statement. These statements can be found in the NAEP Science Assessment and Item Specifications. **Clarification statements** enhance NAEP DCs by explaining the emphasis, giving examples, or providing a specific point of detail. **Boundary statements** tell the item writer what the item should not cover in relation to the NAEP DC. Since NAEP testing occurs midyear, the grade levels are defined by what a student would know and be able to do by the middle of that school year. Although the NRC Framework describes goals for the end of grade 5, some individual state's standards have grade-by-grade standards for elementary grades. The Development Panel therefore considered the grades at which these concepts appear in state standards across the country to ensure that the NAEP DCs reflect appropriate expectations for grade 4 rather than grade 5. The framework organizes the multiple disciplines of science into three major disciplines, several topics per discipline, and several subtopics per topic.

Exhibit 2.1. NAEP Science Disciplines, Topics, and Subtopics

Physical Science	**Life Science**	**Earth and Space Sciences**
Matter and Its Properties • Properties of Matter • Structure of Matter • Phases of Matter and Atomic Substructure • Chemical Processes • Nuclear Processes Motion and Forces • Forces on an Object • Forces between Objects • Types of Forces Energy • Energy Flow and Transfer • Kinetic and Potential Energy • Thermal and Radiant Energy • Energy Conservation Waves and Their Role as Carriers of Information • Wave Patterns • Sound Waves • Electromagnetic Waves	From Molecules to Organisms: Structures and Processes • Structure and Function of Living Things • Reproduction • Matter and Energy in Organisms Ecosystems: Interactions, Energy, and Dynamics • Interdependent Relationships • Cycles of Matter and Energy Transfer • Ecosystem Dynamics, Functioning, and Resilience Heredity: Inheritance and Variation of Traits • Inheritance • Variation Biological Evolution: Unity and Diversity • Evidence of Common Ancestry and Diversity • Mechanisms of Change	Universe, Solar System, and Earth • Patterns of Motion of Space Objects • Solar System • Formation of the Universe Earth's Systems • Plate Tectonics, Patterns on the Surface of the Earth • Earth's History • Water Cycling, Weathering, and Erosion • Weather and Climate Earth and Human Activity • Natural Resources • Natural Hazards • Human Impacts on Earth Systems • Climate Change

These groups should provide a coherent organization of the ideas to be tested.

The determination of which DCs to include in the NAEP Science Assessment Framework prioritized DCs that are

- useful in understanding the world and informing decisions in everyday life,
- central to the discipline,

- likely to endure after instruction,
- able to be measured meaningfully with items that engage students in scientific sensemaking about a variety of phenomena and finding solutions to problems,
- critical to measure and monitor to understand large-scale trends in students' science learning, and
- included in most state standards.

The focus in the selection process was on the central principles of each discipline. The selected big ideas represent foundational and pervasive knowledge, key points of scientific theories, and underpinnings upon which complex understanding is built. A primary consideration was the grade-level appropriateness and accuracy of the NAEP DCs based on the great majority of state standards.

As an organizational tool in the NAEP DC tables below, each NAEP DC is preceded by a specific code (e.g., L12.10). Within a code, the letter denotes broad content area ("P" for Physical Science, "L" for Life Science, and "E" for Earth and Space Sciences); the number before the period denotes grade level (grade 4, 8, or 12); and the number following the period denotes the concept's order of appearance within a given content area and grade. For example, L12.10 denotes that this is the tenth concept to appear in the grade 12 section of the Life Science DCs. Because the numbering within each content area and grade is sequential, code numbers do not necessarily indicate any relationships across grades.

Disciplinary Concepts in Physical Science

Matter and Its Properties

- Properties of Matter
- Structure of Matter
- Phases of Matter and Atomic Substructure
- Chemical Processes
- Nuclear Processes

Motion and Forces

- Forces on an Object
- Forces between Objects
- Types of Forces

Energy

- Energy Flow and Transfer
- Kinetic and Potential Energy
- Thermal and Radiant Energy
- Energy Conservation

Waves and Their Role as Carriers of Information

- Wave Patterns
- Sound Waves
- Electromagnetic Waves

Exhibit 2.2A. Topic: Matter and Its Properties

Overarching Question: *How can the great variety of substances and processes of change in matter be explained?*

Subtopics	Grade 4	Grade 8	Grade 12
Properties of Matter	P4.1: Different types of matter (materials) have different properties. Each material can be classified using a number of its properties. Materials with different properties are needed for different uses.	P8.1: Each pure substance can be identified by its characteristic properties.	
Structure of Matter		P8.2: All substances are made from atoms. There are over 100 different types of atoms, which combine with one another in various ways. Atoms form molecules or extended structures.	P12.1: All matter is made of atoms that contain protons that are positively charged, neutrons that have no electric charge in the nucleus, and electrons that have negative charge that surround the nucleus. Neutral atoms can lose electrons to become positively charged ions or gain electrons to become negatively charged ions. P12.2: Electrical attractions and repulsions between positively charged nuclei and negatively charged electrons explain both the structure of isolated atoms and the forces between two or more nearby atoms that cause them to form molecules, compounds, and extended materials (i.e., the formation of chemical bonds).

Subtopics	Grade 4	Grade 8	Grade 12
Phases of Matter and Atomic Substructure	P4.2: Many materials can be solid and liquid depending on temperature.	P8.3: In any state–gas, liquid, or solid–the temperature influences the motion of atoms and molecules. In solids the atoms are close together, held in place relative to each other by forces between them, and move only with small vibrations about those positions. In liquids, the atoms or molecules are close together but are moving around relative to one another. The atoms and molecules that make up gas are relatively far apart and move around freely.	
Chemical Processes		P8.4: In a chemical reaction, the atoms of the reacting substances are regrouped in characteristic ways into new substances with different properties. Atoms only rearrange. As such the amount of matter does not change.	P12.3: In gasses or liquids, the motion of atoms or molecules leads to collisions between them. Such collisions are necessary for chemical processes to occur. Higher rates of collisions occur at higher temperatures, because atoms are typically moving faster, and at higher pressure in a gas, because the atoms are closer together. P12.4: A stable molecule has less energy than the same set of atoms at rest far apart. Any process that results in a new set of molecules must start with some energy input that allows a break-up for the initial molecule or molecules to begin the process. Often this energy comes from the kinetic energy of colliding molecules.

Subtopics	Grade 4	Grade 8	Grade 12
Chemical Processes (continued)			P12.5: In some chemical reactions, energy is released as higher kinetic energy of motions of the products compared to that of the reactants. P12.6: The total number of atoms of each type does not change in any chemical process; that is, atoms are conserved in all such processes. Knowing that atoms are conserved during chemical processes, together with knowledge of the characteristic chemical properties of each element, allows individuals to describe and predict chemical reactions.
Nuclear Processes			P12.7: Nuclear processes, including fusion, fission, and radioactive decays of unstable nuclei, involve release or absorption of energy. The total number of neutrons plus protons does not change in any nuclear process.

Exhibit 2.2B. Topic: Motion and Forces

Overarching Questions: *How can motion be described? What makes the motion of an object change?*

Subtopics	Grade 4	Grade 8	Grade 12
Forces on an Object	P4.3: Unequal forces acting on an object can change its motion, or forces can balance against other forces to hold the object in place.	P8.5: The change in motion of an object is determined by the sum of the forces acting on it; if the net force on the object is zero, it will remain at rest or continue moving in a straight line with the same speed and direction as before. P8.6: The greater the mass of the object, the greater the force needed to achieve the same change in motion. For any given object, a larger net force causes a larger change in motion.	P12.8: The motion of an object changes if and only if the sum of the forces acting on it is non-zero.
Forces between Objects		P8.7: For any pair of interacting objects, the force exerted by the first object on the second object is equal in strength to the force that the second object exerts on the first but in the opposite direction.	P12.9: Momentum is always conserved, whether within a system or between two different systems. This is a consequence of the fact that the forces between any two interacting objects are equal and opposite and thus result in equal and opposite changes in momentum.

Subtopics	**Grade 4**	**Grade 8**	**Grade 12**
Types of Forces	P4.4: Objects exert forces on each other when they are touching or colliding with each other.	P8.8: Electric and magnetic forces between two objects can pull them together or push them apart. The magnitude depends on the magnitude of the charges, currents, or magnetic strengths involved and on the distances between the interacting objects. P8.9: The gravitational forces between any two objects with mass will pull them toward each other. The gravitational force between any two masses is very small except when one or both of the objects have large mass—e.g., Earth and the sun.	P12.10: Forces between objects at a distance are explained by fields (gravitational, electric, and magnetic) permeating space that can transfer energy and momentum through space. Any object with mass is a source of a gravitational field, which exerts an attractive force on any other mass. The strength of the pair of forces between any pair of masses is proportional to the product of their masses and depends on the distance between the two centers of mass. P12.11: Attraction, repulsion, and magnetic effects between electric charges (their electromagnetic interactions) at the atomic scale explain the structure, properties, and atomic-scale processes of matter and forces between surfaces in contact.

Exhibit 2.2C. Topic: Energy

Overarching Questions: *Why do we care about keeping track of energy? Why are so many different phenomena associated with energy?*

Subtopics	Grade 4	Grade 8	Grade 12
Energy Flow and Transfer	P4.5: Energy can move from place to place by the motion of objects or by sound, light, heat, or electricity. P4.6: When objects collide, the forces between them can transfer energy from one object to the other. Typically, a sound is produced, showing that some energy has been transferred to the air.	P8.10: When two objects interact, each one exerts a force on the other that can cause energy to be transferred from one object to the other. P8.11: Electric currents are generated in multiple ways using a variety of energy transfers to produce them. We use that energy to produce the movement of machines, heat, and/or light. All the energy so "used" is eventually transferred to the surrounding environment as thermal energy.	P12.12: When two objects interacting through a field change relative position, the energy stored in the field is changed.
Kinetic and Potential Energy	P4.7: Objects in motion have energy. The faster a given object is moving, the more energy it has.	P8.12: The energy of motion of particles or waves is called kinetic energy; for massive objects it is proportional to the mass of the moving object and grows with the square of its speed. P8.13: Any system of objects contains energy because of the gravitational, electric, and magnetic interactions between the objects. This energy is called potential energy. The amount depends on the relative positions of objects.	P12.13: Energy is a quantitative property of any system. The amount of energy available for processes in that system depends on the motion and interactions of matter and radiation within that system. The availability of energy limits what can occur in any system.

Subtopics	Grade 4	Grade 8	Grade 12
Thermal and Radiant Energy	P4.8: Heat and light from the sun are major sources of energy on Earth.	P8.14: The energy associated with random movements of atoms and molecules is called thermal energy. In all matter, the atoms are moving. The more thermal energy, the more the motion of atoms. The term heat is used only for energy transferred between two objects or systems at different temperatures. P8.15: Two systems at the same temperature could have different total energy; the relationship between the temperature and the total energy of a system depends on the types, states, and amounts of matter present.	P12.14: When sunlight is absorbed at Earth's surface, it is eventually reradiated as infrared radiation that transfers heat into the atmosphere. The average temperature of the atmosphere is determined by how long the energy stays in the system until it is reradiated into space from the top of the atmosphere.
Energy Conservation		P8.16: Any object absorbs energy from, or loses energy to, the air or other matter it is touching depending on whether it is colder or hotter than the surrounding matter. Energy is spontaneously transferred out of hotter regions or objects and into colder ones.	P12.15: Energy cannot be created or destroyed, but it can be transferred from one place to another and between systems. P12.16: Although energy cannot be destroyed, it can be converted to a less useful form, becoming thermal energy in the surrounding environment.

Exhibit 2.2D. Topic: Waves and Their Roles as Carriers of Information

Overarching Questions: *How can information be encoded, sent over long distances, and decoded? What physical phenomena do we use to do this?*

Subtopics	Grade 4	Grade 8	Grade 12
Wave Patterns	P4.9: Waves are regular patterns of motion in matter (e.g., waves can be made in water by disturbing the surface).	P8.17: Waves of the same type can differ in amplitude and wavelength, and multiple waves traveling together can add to give complex patterns that can be used to encode information. Waves of the same type traveling in different directions can pass through one another and emerge unchanged.	P12.17: The speed of a wave depends on the type of wave and on properties of the medium through which it is passing. P12.18: Information can be transmitted by continuous waves or as digital pulses and can be stored in digital form (e.g., a picture stored as the values of an array of pixels).
Sound Waves	P4.10: Sound can make matter vibrate, and vibrating matter can make a sound.	P8.18: A sound wave needs a medium through which it is transmitted. The medium can be solid, liquid, or gas.	
Electromagnetic Waves	P4.11: Some materials allow light to pass through them, others allow only some light through, and others reflect or absorb all the light that reaches them and cast a dark shadow on any surface beyond them, where the light cannot reach. An object can be seen only when light produced by it or reflected from its surfaces enters the eyes.	P8.19: When light shines on an object, it is reflected, absorbed, or transmitted through the object, depending on the object's material and the frequency (color) of the light. A wave model of light is useful for explaining brightness, color, and the frequency-dependent bending of light at a surface between media. However, because light can travel through space, it cannot be a matter wave, like sound or water waves.	P12.19: Many seemingly unrelated phenomena, from X-rays to radio waves, are electromagnetic waves like light but have very different wavelengths and frequencies. Electromagnetic waves are produced by patterns of motion of charges or magnets. The wave is a pattern of changing electric and magnetic fields.

Disciplinary Concepts in Life Science

From Molecules to Organisms: Structures and Processes

- Structure and Function of Living Things
- Reproduction
- Matter and Energy in Organisms

Ecosystems: Interactions, Energy, and Dynamics

- Interdependent Relationships
- Cycles of Matter and Energy Transfer
- Ecosystem Dynamics, Functioning, and Resilience

Heredity: Inheritance and Variation of Traits

- Inheritance
- Variation

Biological Evolution: Unity and Diversity

- Evidence of Common Ancestry and Diversity
- Mechanisms of Change

Exhibit 2.3A. Topic: From Molecules to Organisms: Structures and Processes

Overarching Question: *How do organisms live, grow, respond to their environment, and reproduce?*

Subtopics	Grade 4	Grade 8	Grade 12
Structure and Function of Living Things	L4.1: Plants and animals have both internal and external structures that serve central functions necessary for life—growth, survival, behavior, and reproduction.	L8.1: For both single cells and multiple cellular organisms, special structures within cells are responsible for particular functions. L8.2: In multicellular organisms, the body is a system of multiple interacting subsystems that are groups of cells that work together to form tissues and organs that are specialized for particular body functions.	L12.1: Systems of specialized cells within organisms help them perform the essential functions of life, which involve chemical reactions that take place between different types of molecules. L12.2: Multicellular organisms have a hierarchical organization, in which its systems support functions necessary for the organism's survival and reproduction. Each system is made up of numerous parts and is itself a component of the next level. L12.3: Feedback mechanisms maintain a living system's internal conditions within certain limits. Feedback mechanisms discourage change by means of negative feedback or proceed with changes through a system of positive feedback.
Reproduction	L4.2: Reproduction is essential to the continued existence of every kind of organism. Plants and animals have distinct and diverse life cycles.	L8.3: Organisms reproduce, using a variety of structures and processes (both sexual and asexual), and transfer their genetic information to their offspring.	L12.4: In most multicellular organisms, an organism begins as a single cell (a fertilized egg), and then divides successively to produce many cells. Mitosis is the process that allows all cells to divide after a period of growth.

Subtopics	Grade 4	Grade 8	Grade 12
Reproduction (continued)			This process starts with a parent cell copying its genetic material and passing identical genetic material to both cells that result from the division (the daughter cells).
Matter and Energy in Organisms	L4.3: All animals need food, water, and air in order to live and grow. They obtain their food from their surroundings—from plants or from other animals. Plants need air, water, minerals (in the soil), and light to live and grow.	L8.4: Photosynthesizers (i.e., plants, algae, and many microorganisms) use the energy from light to make sugars (food) from carbon dioxide from the atmosphere and water through the process of photosynthesis, which also releases oxygen into the atmosphere. L8.5: Within individual organisms, food moves through a series of chemical reactions in which it is broken down and rearranged to form new molecules, to support growth, or to release energy.	L12.5: The process of photosynthesis converts light energy to stored chemical energy by converting carbon dioxide plus water into sugars plus released oxygen. L12.6: The process of cellular respiration is a chemical process in which the bonds of food molecules and oxygen molecules are broken, and new compounds are formed that can transport energy. L12.7: As a result of photosynthesis and cellular respiration, energy is transferred from one system of interacting molecules to another. Matter and energy are conserved in each change. This is true of all biological systems, from individual cells to ecosystems.

Exhibit 2.3B. Topic: Ecosystems: Interactions, Energy, and Dynamics

Overarching Question: *How and why do organisms interact with their environment, and what are the effects of these interactions?*

Subtopics	Grade 4	Grade 8	Grade 12
Interdependent Relationships	L4.4: Most animals can move from place to place on their own, but plants cannot, and often rely on animals for pollination or to move their seeds around. Different plants survive better in different settings because they have varied needs for water, minerals, and sunlight.	L8.6: In any ecosystem, organisms and populations with similar requirements for food, water, oxygen, or other resources may compete with each other for limited resources, access to which consequently constrains their growth and reproduction. L8.7: Predatory interactions may reduce the number of organisms or eliminate whole populations of organisms. Mutually beneficial interactions, in contrast, may become so interdependent that each organism requires the other for survival. Although the species involved in these competitive, predatory, and mutually beneficial interactions vary across ecosystems, the patterns of interactions of organisms with their environments, both living and nonliving, are shared.	L12.8: Ecosystems have carrying capacities, which are limits to the numbers of organisms and populations they can support. Organisms would have the capacity to produce populations of great size were it not for the fact that environments and resources are finite. This fundamental tension affects the abundance (number of individuals) of species in any given ecosystem.
Cycles of Matter and Energy Transfer	L4.5: Much of the matter (materials) organisms need to grow and survive comes from other organisms and that same matter is used again later by other organisms.	L8.8: Food webs are models that demonstrate how matter and energy are transferred between producers, consumers, and decomposers as the three groups interact within an	L12.9: Photosynthesis and cellular respiration (including anaerobic processes) provide most of the energy for life processes.

Subtopics	Grade 4	Grade 8	Grade 12
Cycles of Matter and Energy Transfer (continued)		ecosystem. Transfers of matter into and out of the physical environment occur at every level. Decomposers recycle nutrients from dead plant or animal matter back to the soil in terrestrial environments or to the water in aquatic environments. The atoms that make up the organisms in an ecosystem are cycled repeatedly between the living and nonliving parts of the ecosystem.	L12.10: Plants or algae form the lowest level of the food web. At each link upward in a food web, only a small fraction of the matter consumed at the lower level is transferred upward to produce growth and release energy in cellular respiration at the higher level.
Ecosystem Dynamic, Functioning, and Resilience	L4.6: When the environment changes in ways that affect a place's physical characteristics (such as geography, effects of fire), temperature, precipitation, or availability of resources, some organisms survive and reproduce, some move to new locations, some move into the transformed environment, and some die.	L8.9: Ecosystems are dynamic in nature; their characteristics can vary over time. Disruptions to any physical or biological component of an ecosystem can lead to shifts in all its populations, therefore helping or hurting the health of the ecosystem, including its biodiversity. L8.10: Changes in biodiversity can influence the resources and ecosystem services that humans rely on.	L12.11: A complex set of interactions within an ecosystem can keep its numbers and types of organisms relatively constant over long periods of time under stable conditions. Extreme fluctuations in conditions or the size of any population, however, can challenge the functioning of ecosystems in terms of resources and habitat availability. L12.12: Changes induced by human activity in the environment—such as habitat destruction, pollution, introduction of invasive species, overexploitation, and climate change—can disrupt an ecosystem, reduce biodiversity, and threaten the survival of some species.

Subtopics	Grade 4	Grade 8	Grade 12
Ecosystem Dynamic, Functioning and Resilience (continued)			L12.13: Humans depend on the living world for the resources and other benefits provided by biodiversity. Changes in biodiversity can influence resources and ecosystem services that humans rely on.

Exhibit 2.3C. Topic: Heredity: Inheritance and Variation of Traits

Overarching Questions: *How are the characteristics of one generation passed to the next? How can individuals of the same species and even siblings have different characteristics?*

Subtopics	Grade 4	Grade 8	Grade 12
Inheritance	L4.7: Many characteristics of organisms are inherited from their parents. These inherited characteristics may result in variations in how they look and function. Other characteristics result from individuals' interactions with the environment. Many characteristics involve both.	L8.11: Genes are located in the chromosomes of cells, with each chromosome pair containing two variants of each of many distinct genes. Each distinct gene chiefly controls the production of specific proteins, which in turn affects the traits of the individual.	L12.14: Each chromosome consists of a single very long DNA molecule, and each gene on the chromosome is a particular region of that DNA. Genes contain the instructions to code for the formation of proteins that determine traits. Not all DNA codes for a protein; some segments of DNA are involved in regulatory or structural functions, and some have no currently known function.
Variation		L8.12: In sexually reproducing organisms, each parent contributes half of the genes acquired (at random) by the offspring. Individuals have two of each chromosome and hence two alleles of each gene, one acquired from each parent. These versions may be identical or may differ from each other. Variations of inherited traits between parent and offspring arise from the inherited subset of chromosomes (and therefore genes). L8.13: In addition to variations that arise from sexual reproduction, genetic information can be altered because of mutations. Although rare, mutations	L12.15: In sexual reproduction, chromosomes can sometimes swap sections during the process of meiosis, thereby creating new genetic combinations and thus more genetic variation. Although DNA replication is tightly regulated and remarkably accurate, errors do occur and result in mutations, which are a source of genetic variation. Environmental factors can also cause mutations in genes, and mutations can be inherited. L12.16: Environmental factors affect expression of heritable traits and hence affect the probability of occurrences of traits in a population.

Subtopics	Grade 4	Grade 8	Grade 12
Variation (continued)		may result in changes to the structure and function of proteins. Some changes are beneficial, others harmful, and some neutral to the organism.	

Exhibit 2.3D. Topic: Biological Evolution: Unity and Diversity

Overarching Questions: *How can there be so many similarities among organisms yet so many different kinds of plants, animals, and microorganisms? How does biodiversity affect humans?*

Subtopics	Grade 4	Grade 8	Grade 12
Evidence of Common Ancestry and Diversity	L4.8: Some kinds of plants and animals that once lived on Earth are no longer found anywhere. Fossils can provide evidence about these types of organisms that lived long ago and about the nature of their environments.	L8.14: The collection of fossils and their placement in chronological order (the fossil record) documents the existence, diversity, extinction, and change of many life-forms throughout the history of life on Earth. Similarities and differences in gross anatomical appearance or in embryological development, between organisms living today and between them and organisms in the fossil record, enable the reconstruction of evolutionary history and inference of lines of evolutionary descent.	L12.17: Genetic information provides evidence of evolution. DNA sequences vary among species, but there are many overlaps. Such information is also derivable from the similarities and differences in amino acid sequences and from anatomical and embryological evidence.
Mechanisms of Change	L4.9: Species change over time. Sometimes the differences in characteristics between individuals of the same species provide advantages in surviving, finding mates, and reproducing. This can be especially true when a habitat changes.	L8.15: Adaptation by natural selection acting over generations is one important process by which species change over time in response to changes in environmental conditions. Heritable traits that support successful survival and reproduction in the new environment become more common; those that do not become less common. Thus, the distribution of traits in a population changes. This can also	L12.18: Evolution by natural selection results from the interaction of four factors: (a) the potential for a species to increase in number, (b) the genetic variation of individuals in a species due to mutation and sexual reproduction, (c) competition for an environment’s limited supply of the resources that individuals need in order to survive and reproduce, and (d) the ensuing proliferation of those organisms that

Subtopics	Grade 4	Grade 8	Grade 12
Mechanisms of Change (continued)		be done artificially by humans selectively breeding for a desired trait in other organisms.	are better able to survive and reproduce in that environment, passing on those traits to offspring. Fitness, as measured by survival and reproduction rates, may be altered if changes in the physical environment, whether naturally occurring or human induced, take place.

Disciplinary Concepts in the Earth and Space Sciences

Universe, Solar System, and Earth

- Patterns of Motion of Space Objects
- Solar System
- Formation of the Universe

Earth's Systems

- Plate Tectonics, Patterns on the Surface of the Earth
- Earth's History
- Water Cycling, Weathering, and Erosion
- Weather and Climate

Earth and Human Activity

- Natural Resources
- Natural Hazards
- Human Impacts on Earth Systems
- Climate Change

Exhibit 2.4A. Topic: Universe, Solar System, and Earth

Overarching Question: *How do we explain Earth's relationship to objects in space?*

Subtopics	Grade 4	Grade 8	Grade 12
Patterns of Motion of Space Objects	E4.1: Many objects in the sky change position and are not always visible due to Earth's rotation. The patterns of motion of the sun and moon can be observed, measured, described, and predicted.	E8.1: The orbits of Earth around the sun and of the moon around Earth, together with the rotation of Earth on an axis that runs from its north pole to its south pole, cause observable and measurable patterns that can be used to predict apparent motions of the sun and moon and occurrence of tides and seasonal changes through models.	E12.1: Cyclical changes in the shape of Earth's orbit, together with changes in the orientation of the planet's axis of rotation—occurring from tens of years to hundreds of thousands of years—have altered the intensity and distribution of sunlight falling on Earth. This variation drives changes in Earth's climate patterns over time.
Solar System	E4.2: Some objects in the solar system can be seen with the naked eye, and some require tools that extend human perception.	E8.2: The solar system consists of the sun and a collection of objects, including planets, their moons, and asteroid belts in orbit around the sun. Gravitational interactions between the sun and planets in the solar system produce orbital patterns that can be observed and predicted.	E12.2: Orbiting objects can be described in terms of their elliptical paths around the sun, as described by Kepler's laws. These orbits can change slightly due to gravitational effects from, or collisions with, other objects in the solar system.
Formation of the Universe	E4.3: We can observe objects in the sky such as the moon, sun, other planets, and other stars. The sun is a star that appears larger and brighter than other stars because it is closer. E4.4: Unlike stars, the moon and other planets do not make their own light but reflect light from the sun so we can see them from Earth.	E8.3: The sun and its solar system are a small piece of a large group of stars called the Milky Way, which is only one of many such galaxies spread out in the universe. Scientific instruments collect and provide information about space objects to understand how they formed, became distributed, and evolved.	E12.3: The study of stars' light spectra and relative brightness is used to identify compositional elements of stars, their movements, and their distances from Earth. This is used to develop explanations of the formation, age, and change over time of the universe.

Exhibit 2.4B. Topic: Earth's Systems

Overarching Question: *What are the Earth's systems, and how do they change?*

Subtopics	Grade 4	Grade 8	Grade 12
Plate Tectonics, Patterns on the Surface of the Earth	E4.5: Locations of local, regional, and global surface features and phenomena reveal patterns on Earth's surface.	E8.4: The Earth consists of layers, including a solid, rigid outer layer divided into plates, which are always moving very slowly. Interactions between Earth's moving plates result in changes of physical features.	E12.4: The transfer of thermal energy from the Earth's interior, generated from radioactive decay, toward the surface, along with the gravitational movement of denser materials back toward the interior, drives the flow of matter inside the Earth. This convection cycle moves Earth's plates and causes the patterns of physical features.
Earth's History	E4.6: Earth and life on Earth have changed over time. The occurrence and location of certain fossil types provide evidence for changes in environmental conditions and the development of life over time.	E8.5: The geologic time scale interpreted from fossils and the sequence of rock strata provides a way to reconstruct how and when major events in Earth's history occurred in terms of relative time.	E12.5: The decay of radioactive isotopes in minerals and rocks provides a measurement for dating rock formations and for providing evidence for Earth's formation and early history.
Water Cycling, Weathering, and Erosion	E4.7: Water is found in oceans, rivers, lakes, and air. The downhill movement of water drives the flow of water toward the ocean. E4.8 Rocks on Earth's surface can be broken into pieces and moved by water, wind, and living organisms; this causes continual, observable changes to surface features.	E8.6: The movement of water within the water cycle is a function of phase changes—evaporation, condensation, freezing, and melting. E8.7: Water continually cycles within and among land, ocean, and atmosphere. Water's movements, both on the land and underground, are driven by gravity and change the land on and below Earth's surface.	E12.6: Interactions between the hydrosphere and the geosphere are influenced by water's unique properties, including its exceptional capacity to absorb, store, and release large amounts of thermal energy; expand upon freezing; dissolve and transport materials; separate different chemical elements; and change the properties of rocks.

Subtopics	Grade 4	Grade 8	Grade 12
Weather and Climate	E4.9: Patterns in when and where weather conditions occur can be used to make predictions about the kind of weather that can be expected in a region.	E8.8: Weather is influenced by interactions involving sunlight, the ocean, the atmosphere, ice, and landforms. Because the interactions are so complex, weather patterns in a given location can only be predicted through probabilities (likelihood to occur), and only for a short period of time into the future. E8.9: Influences on the climate at a given place include latitude, altitude, local and regional geography, and oceanic and atmospheric flow patterns.	E12.7: The absorption, reflection, storage, and redistribution of visible and infrared energy from the Sun among the atmosphere, hydrosphere, and geosphere, and the reradiation of infrared energy into space, lead to the geographic and temporal patterns in Earth's climate. E12.8: Geological and historical evidence indicates changes in past climates are linked to alterations in the composition of atmosphere and variations in solar output or Earth's orbit. The time scales of these changes vary from sudden—few tens of years (e.g., large volcanic eruptions or changes in ocean circulation), to gradual—millions of years (e.g., movement of Earth's plates).

Exhibit 2.4C. Topic: Earth and Human Activity

Overarching Question: *How do Earth's system processes and human activities affect each other?*

Subtopics	Grade 4	Grade 8	Grade 12
Natural Resources	E4.10: Humans depend on natural resources because all living things need water, air, and resources for food, transportation, and shelter, which influences where they live.	E8.10: Natural resources are distributed unevenly by biogeochemical processes around the planet as a result of Earth system processes. Humans depend on the Earth's geosphere, hydrosphere, atmosphere, and biosphere for resources, both renewable and nonrenewable, within human life spans.	E12.9: Resource availability guides the development of human societies. All forms of energy production and resource extraction have associated economic, social, and environmental cost-benefit factors.
Natural Hazards	E4.11: Natural hazards are caused by natural processes. Depending on where one lives, some kinds of natural hazards are more likely than others.	E8.11: Some natural hazards are typically preceded by observable phenomena, which provide a warning for their occurrence (e.g., volcanic eruptions and severe weather). Other hazards occur suddenly and often with very little or no advance warning (e.g., earthquakes and tornadoes). Data on the duration and frequency of the warning signs reveal patterns of natural hazards in a region, which can help forecast the locations and likelihoods of future events in order to minimize risks.	E12.10: Land use and city planning can affect the frequency and intensity of the impacts of some natural hazards; some have significantly altered the size and location of human populations.

Subtopics	**Grade 4**	**Grade 8**	**Grade 12**
Human Impacts on Earth's Systems	E4.12: Human activities cause changes to the local areas where they live. Human choices can increase or decrease the positive and negative impacts on the land, water, and air.	E8.12: Human activities have significantly altered the biosphere, atmosphere, and geosphere, sometimes damaging or destroying ecosystems and causing the extinction of organisms. Human choices can minimize harm to other organisms and risks to the health of the regional environment.	E12.11: When the sources of an environmental problem are understood, applying engineering and design solutions, new technology, and other creative ideas can mitigate negative impacts on Earth's resources and global environment, while inaction on the problem could magnify the negative impacts. When the sources of such problems are not well understood, some actions could magnify the problems.
Climate Change		E8.13: Human activities that release greenhouse gasses, such as production and combustion of fossil fuels, are major factors in the current rise in Earth's temperature. Monitoring the production and reducing the use of fossil fuels can slow the increase in global temperatures as well as the effects of climate change.	E12.12: Current models predict that, although future regional climate changes will be complex and varied, average global temperatures will continue to rise. Changing the outcomes predicted by global climate models strongly depends on reduction of the amounts of human-generated greenhouse gasses added to the atmosphere each year, but is also influenced by uncertainties about behavioral, economic, and political factors and how they will impact potential solutions and their success.

2B. NAEP Science and Engineering Practices

> *Engaging in the practices of science helps students understand how scientific knowledge develops; such direct involvement gives them an appreciation of the wide range of approaches that are used to investigate, model, and explain the world. Engaging in the practices of engineering likewise helps students understand the work of engineers, as well as the links between engineering and science. Participation in these practices also helps students form an understanding of the crosscutting concepts and disciplinary ideas of science and engineering; moreover, it makes students' knowledge more meaningful and embeds it more deeply into their worldview. (NRC 2012, p. 42)*

The ability to engage in the practices of science and engineering (along with the ability to use crosscutting concepts) allows students to apply their disciplinary science knowledge as they develop explanations of phenomena or design solutions to engineering problems. The 2028 NAEP Science Assessment will ask students to engage these abilities as part of achieving a successful response to multidimensional items.

Scientific explanations are explicit applications of theory to a specific situation or phenomenon. The goal for students is to construct coherent explanations of phenomena that incorporate their current understanding of science, or a model that represents it, and are consistent with the available evidence.

In engineering, the goal is a designed solution to a problem rather than an explanation. The term *engineering* applies to any such design, whether it is for an object, a system, or a process. The domain of the problem can be any area of applied science or technology. The problem can arise from individual, community, or global needs or wants. The process of developing a design is iterative and systematic, as is the process of developing an explanation or a theory in science. Engineers' activities have elements that are distinct from those of scientists. These elements include specifying constraints and criteria for desired qualities of the solution, developing a design plan, producing and testing models or prototypes, selecting among alternative design features to optimize the achievement of design criteria, and refining design ideas based on the performance of a prototype or simulation. Engineering design focuses on the appropriate use of technology. Appropriate use of technology refers to using the simplest level of technology that can achieve the intended purpose in each location, using fewer natural resources, emitting less pollution, and costing less.

The NAEP Science and Engineering Practices (SEPs) occur in social contexts and can be used in ways that can benefit or harm individuals, communities, or the environment. As students use these practices, they should do so ethically and recognize the risks and harms that can be caused and have been caused by negligent use, as well as the benefits.

The NAEP SEPs listed in Exhibits 2.6–2.13 focus on aspects of engaging in the practices that can be assessed in large-scale science and engineering-oriented assessments and do not include all aspects needed in instruction. The grade-specific sub-statements in the charts under each

Science and Engineering Practice are individual elements that are organized by aspects of each NAEP SEP. Not every sub-statement needs to be assessed by NAEP. When the same wording in a sub-statement is used for a NAEP SEP at different grade levels, the sophistication of student performances is expected to change based on the NAEP DC. When the NAEP DC is more complex, the use of the NAEP SEP becomes more complex.

The NAEP SEPs are as follows:

- **Asking Questions and Defining Problems**
- **Planning and Carrying Out Investigations**
- **Analyzing and Interpreting Data**
- **Using Mathematics and Computational Thinking**
- **Developing and Using Models**
- **Constructing Explanations and Designing Solutions**
- **Engaging in Argument from Evidence**
- **Obtaining, Evaluating, and Communicating Information**

Scientific work involves all these practices used in an iterative and recursive process to achieve an eventual explanation based on a well-tested model. It further requires honest reporting and critical review to be effective. Engineering work likewise involves an iterative and recursive process using all these practices to design, test, and redesign to achieve a successful problem solution. Individuals and organizations working within science and engineering also consider how their work contributes to ecological and social matters and how to optimize their work, products, and applications to benefit society and minimize harms, including consideration of unintended negative effects. Scientific theories and explanations are empirically based and subject to revision based on new or evolving evidence. The same applies to student-developed models, explanations, and engineering design solutions. Students should be able to reflect on what needs to be revised and whether additional evidence is required to improve the outcome or strengthen the claim.

For NAEP Science Assessment purposes, the NAEP SEPs have been organized into four categories: Investigating, Analyzing, Explaining, and Evaluating (Exhibit 2.5). On any given assessment, at least 10 percent of the items must fall into each of the four categories.

Exhibit 2.5. NAEP Science and Engineering Practices

Investigating	Asking Questions and Defining Problems
	Planning and Carrying Out Investigations
Analyzing	Analyzing and Interpreting Data
	Using Mathematics and Computational Thinking
Explaining	Developing and Using Models
	Constructing Explanations and Designing Solutions
Evaluating	Engaging in Argument from Evidence
	Obtaining, Evaluating, and Communicating Information

Asking Questions and Defining Problems

Scientific questions arise in a variety of ways. They can be driven by curiosity about the world; inspired by the predictions of a model, theory, or findings from previous investigations; or driven by the need to solve a problem. Scientific questions are distinguished from other types of questions in that the answers lie in explanations supported by empirical evidence. Engineering design work also begins with asking questions to help define a problem to solve.

Many aspects of asking questions do not lend themselves to assessment. The aspects of questioning listed in Exhibit 2.6 are those that can reasonably be the practice element of a science or engineering assessment item.

Exhibit 2.6. Asking Questions and Defining Problems

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Asking questions to inform an investigation or develop an explanation or model of phenomena	S4.1: Ask questions to help refine observations, develop interpretations of data, develop and/or evaluate models, or define an engineering problem. S4.2: Ask “what if” questions about a system or phenomenon being observed that	S8.1: Ask questions to clarify and/or refine an observation, model, or explanation of phenomena; or to clarify and/or refine an engineering problem. S8.2: Ask questions that can be answered with empirical evidence to investigate relationships between	S12.1: Ask questions that arise from examining a model, an explanation, or a design plan to clarify and/or identify additional needed information or tests. S12.2: Ask investigable questions to determine relationships, including quantitative

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Asking questions to inform an investigation or develop an explanation or model of phenomena (continued)	could be investigated empirically.	variables in a system model or in phenomena.	relationships, between independent and dependent variables in a model, and when appropriate frame a hypothesis about potential findings.
Asking questions as part of understanding, evaluating, and/or challenging the work of others	S4.3: Ask questions to clarify an argument or interpretation of a data set.	S8.3: Ask questions to clarify or respectfully challenge the evidence and/or the premise(s) of an argument or interpretation of a data set.	S12.3: Ask and/or evaluate questions that challenge the premise(s) of an argument, the interpretation of a data set, or the suitability of design considerations.
Defining a design problem that addresses a need	S4.4: Define a design problem to provide a solution for a situation people want to change that can be solved through the development of a new or improved object or tool.	S8.4: Define a design problem that considers relevant scientific principles and potential impacts on people and the natural environment that may limit possible solutions and can be solved through the development of an object, tool, process, or system that includes multiple criteria and constraints.	S12.4: Define a design problem that involves the development of a process or system with interacting components and criteria and constraints that may include social, technical, ethical, and/or environmental considerations.

Planning and Carrying Out Investigations

Scientific investigations may be undertaken to describe a phenomenon and to test a theory or model for how the world works. The purpose of engineering investigations might be to determine conditions under which the design solution needs to function, to find out how to fix or improve the functioning of a technological system, or to compare different solutions to see which best solves a problem. Whether students are doing science or engineering, it is always important for them to state the goal of an investigation, predict outcomes, and plan a course of action that will provide the best evidence to support their conclusions or design solutions. Students should design investigations that generate data to provide evidence to support claims they make about phenomena. Students should build engineering investigations that address the criteria and constraints.

Over time, students are expected to become more systematic and careful in their designing methods, including the selection of instruments and tools for collecting data. To plan for laboratory experiments, students are expected to decide which variables should be treated as results or outputs, which should be treated as inputs and intentionally varied from trial to trial, and which should be controlled, or kept the same across trials. Planning for field observations involves deciding how to collect different samples of data under different conditions, even though not all conditions are under the direct control of the investigator. In planning for engineering investigations to test design solutions, students select tools, materials, and processes relative to constraints and criteria.

NAEP Science Assessment items should provide students with tools or instrument-specific information that is needed for successful item completion. Students will not be required to carry out experiments with physical equipment, but simulations or virtual laboratories could be made available for some items.

Exhibit 2.7. Planning and Carrying Out Investigations

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Developing or revising an investigation plan	S4.5: Plan an investigation to explore a scientific question or design problem taking into consideration appropriate variables and tests.	S8.5: Evaluate and/or revise an experimental design that can produce data to serve as the basis for evidence that meets the goals of the investigation or design problem.	S12.5: Plan an investigation that will produce data to serve as the basis for evidence as part of building and revising models, supporting explanations for phenomena, or testing solutions to problems. Consider possible confounding variables or effects and evaluate the investigation's design to ensure appropriate variables are controlled.
Selecting and evaluating appropriate tools for an investigation	S4.6: Select methods and/or tools for collecting data.	S8.6: Select and evaluate tools to collect, record, and analyze data.	S12.6: Select and evaluate appropriate tools to collect, record, analyze, synthesize, and evaluate data.

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Predicting expected outcomes	S4.7: Make predictions about what would happen if a variable changes. S4.8: Predict the outcome of an experiment or a design solution based on a model, a phenomenon, or a design plan.	S8.7: Predict the change in a dependent variable when a change in an independent variable occurs in an investigation or test of a design plan.	S12.7: Predict the direction and magnitude of change of a dependent variable for a change in the independent variable and provide rationale to support the prediction. S12.8: Predict the outcome of an investigation or test of a design plan and support that prediction with an argument including evidence from models, evidence from prior experiments, and/or the application of science knowledge to support the prediction.

Analyzing and Interpreting Data

Data must be organized, analyzed, and interpreted to serve as the evidence to support claims. In the data-rich world of today, this work has become a discipline called data science. Students, like scientists and engineers, use a range of tools to display and analyze data and to identify patterns, sources of error, and degrees of certainty in sets of data. They organize and analyze data to test model-based predictions, to infer relationships and trends in a system, to provide evidence for claims and arguments, to support or refute hypotheses or explanations, or to compare different solutions to specific design criteria and determine which design best solves the problem within given constraints.

Exhibit 2.8. Analyzing and Interpreting Data

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Displaying data to observe patterns and relationships	S4.9: Represent data in tables and/or various graphical displays (e.g., bar graphs and pictographs) to provide information or visualize relationships that can help to explain phenomena or solve design problems.	S8.8: Construct, analyze, and/or interpret graphical displays of data and/or large data sets from an investigation (e.g., maps, charts, graphs, and/or tables) to identify relationships between variables (linear vs. nonlinear relationships, causal vs. correlational relationships, and temporal and spatial relationships).	S12.9: Construct, analyze and/or interpret representations of small and large data sets from an investigation using tools, technologies, and/or models (e.g., computational, mathematical), including statistical analysis (descriptive statistics) and probability.
Analyzing data to support or reject claims about phenomena or improve design solutions	S4.10: Analyze data to determine whether it supports or refutes a claim about a phenomenon or design solution. S4.11: Analyze data from tests of two solutions to the same problem to compare the strengths and weaknesses of how each performs.	S8.9: Analyze data to provide evidence to support or reject a model or explanation or to use to improve a design solution.	S12.10: Analyze data to provide evidence to support or reject a model or explanation or to use to optimize a design solution relative to criteria for success.
Evaluating the quality and adequacy of data		S8.10: Evaluate the limitations of the data for the intended use, considering factors such as quantity and quality of the data, the tools used to obtain it, and its presentation.	S12.11: Evaluate whether the data are sufficient in quantity, accuracy, and reliability for the purpose intended and suggest needed improvements.

Using Mathematics and Computational Thinking

Both science and engineering require mathematics and information technology. Students apply their understanding of mathematics in science and engineering contexts. It is also in these contexts that they are expected to manipulate quantities with physical units, not just pure numbers.

This practice links to student assessment of mathematics and ability to use computational tools, and the progression of expectations across grade levels is therefore closely aligned with the mathematics expected at each grade level. The item demands for students using this practice will be at or below what is expected on the NAEP Mathematics Assessment. Items should not be purely a mathematical or computational item that can readily be completed without demonstrating any understanding of the disciplinary content of the item. Additional guidance about mathematics is provided in the NAEP Science Assessment and Item Specifications.

Exhibit 2.9. Using Mathematics and Computational Thinking

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Using mathematics	S4.12: Apply simple mathematical concepts and/or processes (such as simple computation, measurement) to a scientific question or a design problem.	S8.11: Apply mathematical concepts and/or processes (such as ratio, rate, percent, basic operations, and simple computations) to scientific questions and/or design problems. S8.12: Interpret and use quantities involving ratios based on two different types of units of measure (e.g., speed, density, and population density).	S12.12: Apply mathematical techniques (such as functions, statistical reasoning, and computational algorithms) to represent and solve scientific questions and/or design problems. S12.13: Interpret and apply ratios, rates, percentages, and unit conversions in the context of complicated measurement problems involving quantities with derived or compound units (such as mg/mL, kg/m^3, etc.).

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Computational thinking	S4.13: Break a process into a series of steps.	S8.13: Use algorithms (a series of ordered steps) to solve a design problem. S8.14: Apply digital tools and/or mathematical concepts and arguments to test and compare proposed solutions to design problems.	S12.14: Apply or revise algorithms when analyzing data or designing, programming, testing, and revising scientific models, explanations, and design solutions. S12.15: Apply mathematical expressions, computer programs, algorithms, or simulations of a process or system to evaluate the model by comparing the outcomes with what is known about the phenomena or design problem.

Developing and Using Models

A practice of both science and engineering is to use and construct models as helpful tools for representing ideas and explanations. These tools include diagrams, drawings, physical replicas, mathematical representations, analogies, and computer simulations. Scientists use the term *model* for all these, whereas engineers may talk of a *design plan* for a diagrammatic representation of a system or a *prototype* for a scaled physical replica. In science, models are used to develop questions and predictions and are repeatedly tested and revised until they can provide successful predictions for tests. They then form the basis of an explanation of the phenomenon of interest. They are likewise a key part of the process of engineering design and of troubleshooting to analyze and identify flaws in designed systems.

Students are expected to develop, test, critique, and apply models as a core feature of their science and engineering assessment. They use models to express, examine, and refine their thinking and support their arguments for a claimed explanation.

While the full cycle of developing a model takes too much time to be included as an assessment item, the phrase “develop a model” is included in the elements described below to cover inclusion of items that ask students to carry out some part of the work of model development.

Exhibit 2.10. Developing and Using Models

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Developing and using models to explain phenomena or design a solution	S4.14: Develop, use, and/or revise a model to describe and explain a phenomenon or describe a design proposal. S4.15: Identify and describe how the parts of a model and the relationships between them represent a phenomenon.	S8.15: Develop, use, and/or revise a model to describe, explain, and/or predict phenomena by identifying relationships among parts and/or quantities in a system, including both visible and invisible quantities. S8.16: Use a model to test ideas about phenomena in natural systems or proposed design solutions.	S12.16: Develop, use, and/or revise a model that includes mathematical relationships (including both visible and invisible quantities) to describe, explain, and/or predict phenomena or to test a proposed design solution.
Identifying and addressing limitations of models	S4.16: Identify limitations of a model for a phenomenon in terms of what the model can or cannot yet explain.	S8.17: Evaluate limitations of a model for a phenomenon and propose revisions to address what the model cannot yet explain.	S12.17: Evaluate merits and limitations of two different models of the same proposed tool, process, mechanism, or system to select or revise a model that best fits the evidence or design criteria.

Constructing Explanations and Designing Solutions

Students are expected to apply scientific knowledge to explain phenomena or to develop designs that offer a solution to a problem. Explanations must be supported with an argument based on evidence (see the following practice). In science, the argument is most often model-based, and the evidence enters in the process of testing and revising the model. Designed solutions must be supported by tests of the design through prototypes or simulations.

Exhibit 2.11. Constructing Explanations and Designing Solutions

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Data-based explanations	S4.17: Develop an evidence-based description or explanation supported by evidence and reasoning of a phenomenon or the action of a designed solution.	S8.18: Construct or revise an explanation that uses a chain of cause and effect or evidence-based associations between factors to account for the qualitative or quantitative relationships between variables in a phenomenon.	S12.18: Construct or revise an explanation that uses a chain of cause and effect or evidence-based associations between factors to account for the qualitative or quantitative relationships between variables in a phenomenon.
Model-based explanations	S4.18: Relate an explanation of a phenomenon to a model.	S8.19: Evaluate whether a model provides sufficient explanation of the phenomenon and how the model could be revised to better explain the observations.	S12.19: Evaluate a model-based explanation or a design proposal using empirical evidence and the application of disciplinary concepts.
Designing and comparing solutions	S4.19: Compare multiple possible solutions to a design problem based on how well each is likely to meet the criteria and constraints of the problem.	S8.20: Apply scientific ideas or principles to propose tests or trade-offs needed to optimize a design.	S12.20: Evaluate and/or refine a solution for a design problem, based on scientific knowledge, evidence, prioritized criteria, and trade-off considerations.

Engaging in Argument From Evidence

Evidence in science and engineering is based on the analysis of empirical data and its comparison with the predictions of a model or the goals and constraints of a design plan.

Scientists argue to critique or defend a model or explanation; engineers likewise argue to support the merits or critique flaws of a design. Students are expected to argue or critique proposed models, explanations, and designs—both their own and those of others—using evidence from multiple sources as part of the cycle of testing and improving them. The evidence that the students are expected to use in supporting or refuting an argument in an assessment context should be provided to them, possibly also with evidence that is not to be used.

Exhibit 2.12. Engaging in Argument From Evidence

Aspects of the NAEP SEP	Grade 4	Grade 8	Grade 12
Constructing an argument to support or refute a model, explanation, or design solution	S4.20: Construct and/or support an argument with evidence to support or reject a claim about a phenomenon or a design solution. S4.21: Make a claim about the merits of a design solution by citing relevant evidence about how it meets the criteria and constraints of the problem.	S8.21: Construct an argument with evidence and scientific reasoning to support or reject a proposed model, explanation, or design solution for a problem. S8.22: Identify evidence that could be used to refute a claim about a phenomenon.	S12.21: Construct an argument with evidence and scientific reasoning to support or reject a proposed model, explanation, or design solution for a problem.
Evaluating and/or improving an argument for an explanation, model, or design solution	S4.22: Evaluate an argument based on the evidence or reasoning it includes.	S8.23: Revise an argument that supports or rejects a model, explanation, or design solution for a problem to address new evidence. S8.24: Compare and critique two arguments on the same question to analyze their fit with the evidence and/or whether they emphasize similar or different evidence and/or interpretations.	S12.22: Revise an argument to support or reject a model, explanation, or design solution for a problem to address new evidence or to address a counterclaim. S12.23: Compare and evaluate the arguments for two competing design solutions, based on design criteria, empirical evidence, and/or relevant factors such as economic, societal, environmental, or ethical considerations.

Obtaining, Evaluating, and Communicating Information

Reading, interpreting, evaluating, and producing scientific and technical texts, which can include both written and visual information along with data presentation and mathematical relationships, are fundamental practices of science and engineering, as is communicating clearly and persuasively using both verbal and visual resources.

Being a critical consumer of information about science and engineering requires the ability to read or view reports of scientific or technological advances or applications (whether found in the press, the internet, or social media) and to recognize the salient ideas; identify sources of error and methodological flaws; and distinguish observations from inferences, arguments from explanations, and claims from evidence. Scientists and engineers employ multiple sources to obtain information used to evaluate the merit and validity of claims, methods, and designs.

Evaluating information is a critical skill in the world today, where both information and misinformation (even deliberate disinformation) are widely available through digital sources.

Students need to know how to compare information from multiple sources and, where contradictions exist, to use reasonable criteria to determine the most reliable sources and to argue for the merits or unreliability of a source of information.

Communicating information, evidence, and ideas can be done in multiple ways: using tables, diagrams, graphs, models, interactive displays, and equations; speaking; writing; and discussing.

NAEP Science items will require students to use their skills in reading and interpreting text, combining that with graphic information to understand the item context and to communicate their conclusions, so these aspects of this practice are not stressed in the table of elements of the practice to be specifically assessed. In addition, reading comprehension is not intended to be explicitly measured by NAEP Science items.

Exhibit 2.13. Obtaining, Evaluating, and Communicating Information

Grade 4	Grade 8	Grade 12
S4.23: Evaluate whether the information presented is evidence, an opinion, or a fictional story. S4.24: Evaluate whether the information presented in a text summarizing a graph or table of data accurately reflects the claim that could be made from the data.	S8.25: Assess the credibility, accuracy, and possible bias of an article on a science topic (e.g., based on where it is found, the qualifications of the source, and/or the evidence given to make the claim). S8.26: Evaluate information from two different sources to determine whether there are conflicts between them. S8.27: Identify and critique standard flaws in science-related arguments (e.g., poor assumptions, cause vs. correlation, faulty explanations, or overgeneralizations from limited data).	S12.24: Assess the credibility, accuracy, and possible bias of an article on a science topic (e.g., based on where it is found, the qualifications of the source, and/or the evidence given to make the claim). S12.25: Evaluate scientific and/or technical information from multiple sources, assessing the evidence used by and the information on qualifications and expertise of each source. S12.26: Identify and critique standard flaws in science-related arguments (e.g., poor assumptions, cause vs. correlation, faulty explanations, or overgeneralizations from limited data).

2C. NAEP Science Crosscutting Concepts

> *Some important themes pervade science, mathematics, and technology and appear over and over again, whether we are looking at an ancient civilization, the human body, or a comet. They are ideas that transcend disciplinary boundaries and prove fruitful in explanation, in theory, in observation and in design. (American Association for the Advancement of Science, 1990, p. 123)*

> *These crosscutting concepts were selected for their value across the sciences and in engineering. These concepts help provide students with an organizational framework for connecting knowledge from the various disciplines into a coherent and scientifically based view of the world. (NRC Framework, 2012, p. 83)*

The idea that broad concepts common to nearly all fields of science and engineering should be included as an essential part of science education for all students was initially proposed in the seminal work *Science for All Americans* (AAAS, 1990), in which they were referred to as "themes." Later, the list of these concepts was refined and renamed "unifying concepts" in the *National Science Education Standards* (NRC, 1996), and further refined as "crosscutting concepts" in *A Framework for K–12 Science Educatio*n (NRC, 2012).

The NAEP Crosscutting Concepts (CCCs) are based on those defined in the NRC Framework. Students can use these concepts in many contexts as tools for scientific sensemaking. NAEP CCCs are deeply linked to NAEP SEPs and are conceptual tools that guide effective and reflective practice.

Three-dimensional items require students to use a NAEP DC, SEP, and CCC to answer them. A three-dimensional item should elicit evidence that the student demonstrated using the NAEP CCC to solve it. Some aspects of the NAEP CCCs overlap with NAEP DCs and SEPs; combinations that include these overlaps would not be considered sufficient for developing three-dimensional items. For example, an item that requires students to develop a model to explain the relationships in an ecosystem using energy transfers and matter flows between organisms (L8.8, S8.15) might overlap aspects of C8.7, interactions between components of a system and C8.12, track and model the energy transfers and matter flows. To write three-dimensional items, the developer should either emphasize different aspects of these NAEP CCCs or select different NAEP CCCs to assess.

The NAEP CCCs are as follows:

- **Patterns**
- **Mechanisms and Explanation: Cause and Effect**
- **Scale, Proportion, and Quantity**
- **Systems and System Models/Systems Thinking**
- **Conservation, Flows, and Cycles: Tracking Energy and Matter**
- **Relationships Between Structure and Function**
- **Conditions for Stability and Change in Systems**

Exhibits 2.14–2.20 describe each of the NAEP CCCs in detail. The sub-statements in the charts are individual elements that pull out aspects of each NAEP CCC that might be assessed at this grade level, but not every sub-statement needs to be assessed by NAEP.

Patterns

Patterns exist everywhere—in regularly occurring shapes or structures and in repeating events and relationships. Patterns are discernible in the symmetry of flowers and snowflakes, the cycling of the seasons, and the repeated base pairs of DNA. Noticing patterns is often a first step to organizing and asking scientific questions about why and how the patterns occur.

One major use of pattern recognition is in classification, which depends on careful observation of similarities and differences; objects can be classified into groups on the basis of similarities of visible or microscopic features or on the basis of similarities of function. Such classification is useful in codifying relationships and organizing a multitude of objects or processes into a limited number of groups. Patterns of similarity and difference and the resulting classifications may change, depending on the scale at which a phenomenon is being observed. For example, isotopes of a given element are different—they contain different numbers of neutrons—but from the perspective of chemistry they can be classified as equivalent because they have identical patterns of chemical interaction. Once patterns and variations have been noted, they lead to questions; scientists seek explanations for observed patterns and for the similarity and diversity within them. Engineers often look for and analyze patterns, too. For example, they may diagnose patterns of failure of a designed system under test in order to improve the design, or they may analyze patterns of daily and seasonal use of power to design a system that can meet the fluctuating needs.

The ways in which data are represented can facilitate pattern recognition and lead to the development of a mathematical representation, which can then be used as a tool in seeking an underlying explanation for what causes the pattern to occur. Biologists studying changes in population abundance of several different species in an ecosystem can notice the correlations between increases and decreases for different species by plotting all of them on the same graph and can eventually find a mathematical expression of the interdependencies and food web relationships that cause these patterns.

Exhibit 2.14. Patterns

Patterns: Observed patterns in nature guide organization and classification and prompt questions about relationships and causes underlying them.

Grade 4	Grade 8	Grade 12
C4.1: Similarities and differences in patterns can be used to sort, classify, communicate, predict, and explain, with various representations (such as physical graphs or diagrams) to describe and analyze features of simple natural phenomena and designed products.	C8.1: Patterns in data can be identified and represented using graphs, charts, and tables. Analyzing patterns can help identify cause-and-effect relationships and estimate probabilities of events.	C12.1: Patterns in data can be identified and represented using graphs, mathematical relationships, and statistical quantities. Analyzing correlated patterns can help identify cause-and-effect relationships and estimate probabilities of events, but correlation alone is not sufficient information to infer a causal relationship.

Mechanisms and Explanation: Cause and Effect

Cause and effect involves the search for the underlying cause of a phenomenon. Any tentative answer, or hypothesis, that A causes B requires a model or mechanism for the chain of interactions that connects A and B. For example, the notion that diseases can be transmitted by a person’s touch was initially treated with skepticism by the medical profession for lack of a plausible mechanism. Today infectious diseases are well understood as being transmitted by the passing of microscopic organisms (bacteria or viruses) between an infected person and another. A major activity of science is to uncover such causal connections, often with the hope that understanding the mechanisms will enable predictions and, in the case of infectious diseases, the design of preventive measures, treatments, and cures.

In engineering, the goal is to design a system to cause a desired effect, so cause-and-effect relationships are as much a part of engineering as of science. The process of design is a good place to help students begin to think in terms of cause and effect, because they must understand the underlying causal relationships in order to devise and explain a design that can achieve a specified objective.

When students perform the practice of Planning and Carrying Out Investigations, they often use ideas related to cause and effect. At early ages, this involves doing something to the system of study and then watching to see what happens. At later ages, experiments are set up to test the sensitivity of the parameters involved, and this is accomplished by making a change (cause) to a single component of a system and examining, and often quantifying, the result (effect). The NAEP CCC of Mechanisms and Explanation: Cause and Effect is also closely associated with the NAEP SEP of Engaging in Argument from Evidence. In scientific practice, deducing the cause of an effect is often difficult, so multiple hypotheses may coexist. For example, though the occurrence (effect) of historical mass extinctions of organisms, such as the dinosaurs, is well established, the reason or reasons for the extinctions (cause) are still debated, and scientists develop and debate their arguments based on different forms of evidence. When students engage in scientific argumentation, it is often centered on identifying the causes of an effect.

Exhibit 2.15. Mechanisms and Explanation: Cause and Effect

Mechanisms and Explanation: Cause and Effect: Events have causes, sometimes simple, sometimes multifaceted. Deciphering causal relationships, and the mechanisms by which they are mediated, is a major activity of science and engineering.

Grade 4	Grade 8	Grade 12
C4.2: Cause-and-effect relationships are routinely identified, tested, and used to explain changes. C4.3: Events that occur together with regularity might have a cause-and-effect relationship or might have some other shared explanation.	C8.2: Relationships can be classified as causal or correlational, and correlation does not necessarily imply causation. C8.3: Cause-and-effect relationships may be used to predict phenomena in natural or designed systems. C8.4: Phenomena may have more than one cause, and some cause-and-effect relationships in systems can only be described using probability.	C12.2: Empirical evidence is required to differentiate between cause and correlation and make claims about specific causes and effects. C12.3: Cause-and-effect relationships can explain and predict complex natural and human-designed systems. Such explanations may require examining and modeling small-scale mechanisms within the system.

Scale, Proportion, and Quantity

Scale, proportion, and quantity are fundamental assessments of dimension that form the foundation of observations about nature. Before an analysis of function or process can be made (the how or why), it is necessary to identify the what. These concepts are the starting point for scientific understanding, whether it is of a total system or its individual components.

An understanding of scale involves not only understanding that systems and processes vary in size, time span, and energy, but also that different mechanisms operate at different scales. In engineering, no structure could be conceived, much less constructed, without a precise sense of scale. At a basic level, in order to identify something as bigger or smaller than something else—and how much bigger or smaller—a student must appreciate the units used to measure it and develop a feel for quantity. Metric units of measure are used for grades 4, 8, and 12.

The ideas of ratio and proportionality as used in science can extend and challenge students' mathematical understanding of these concepts. To appreciate the relative magnitude of some properties or processes, it may be necessary to grasp the relationships among different types of quantities—for example, speed as the ratio of distance traveled to time taken, density as a ratio of mass to volume. This use of ratio is quite different from a ratio of numbers describing fractions of a pie. Recognition of such relationships among different quantities is a key step in forming mathematical models that interpret scientific data.

The NAEP CCC of Scale, Proportion, and Quantity figures prominently in the NAEP SEPs of Using Mathematics and Computational Thinking and of Analyzing and Interpreting Data. This concept addresses taking measurements of structures and phenomena, and these fundamental observations are usually obtained, analyzed, and interpreted quantitatively. This NAEP CCC also figures prominently in the NAEP SEP of Developing and Using Models.

Scale and proportion are often best understood using models. For example, the relative scales of objects in the solar system or of the components of an atom are difficult to comprehend mathematically (because the numbers involved are either so large or so small), but visual or conceptual models make them much more understandable (e.g., if the solar system were the size of a penny, the Milky Way galaxy would be the size of Texas).

Exhibit 2.16. Scale, Proportion, and Quantity

Scale, Proportion, and Quantity: In considering phenomena, it is critical to recognize what is relevant at different size, time, and energy scales and to recognize proportional relationships between different quantities as scales change.

Grade 4	Grade 8	Grade 12
C4.4: Natural objects and/or observable phenomena exist from the very small to the immensely large or from very short to very long time periods.	C8.5: The observed function of natural and designed systems may change with scale. Phenomena that can be observed at one scale may not be observable at another scale. C8.6: Time, space, and energy phenomena can be observed at various scales using models to study systems. Proportional relationships (e.g., speed as the ratio of distance traveled to time taken) among different types of quantities provide information about the magnitude of properties and processes.	C12.4: Explanations of phenomena observable at one scale may require models of the system or of processes at many-orders-of-magnitude-smaller scale (e.g., macroscale processes in matter require atomic-level understanding of forces between and among atoms). C12.5: Algebraic thinking is used to examine models and scientific data and predict the effect of a change in one variable on another (e.g., linear growth vs. exponential growth).

Systems and System Models / Systems Thinking

Systems thinking and system models are useful in science and engineering because the world is complex, so it is helpful to isolate a single system and construct a simplified model of it. To do so, scientists and engineers imagine an artificial boundary between the system in question and everything else. They then examine the system in detail while treating the effects of things outside the boundary as either forces acting on the system or flows of matter and energy across it. Consideration of flows into and out of the system is a crucial element of system design. In the laboratory or even in field research, the extent to which a system under study can be physically isolated or external conditions controlled is an important element of the design of an investigation and interpretation of results. The properties and behavior of the whole system can be very different from those of any of its parts, and large systems may have emergent properties, such as the shape of a tree, that cannot be predicted in detail from knowledge about the components and their interactions.

Models can be valuable in predicting a system's behaviors or in diagnosing problems or failures in its functioning, regardless of what type of system is being examined. In a simple mechanical system, interactions among the parts are describable in terms of forces among them that cause changes in motion or physical stresses. In more complex systems, it is not always possible or useful to consider interactions at this detailed mechanical level, yet it is equally important to ask what interactions are occurring (e.g., predator-prey relationships in an ecosystem) and to recognize that they all involve transfers of energy, matter, and (in some cases) information among parts of the system. Any model of a system incorporates assumptions and approximations; the key is to be aware of what they are and how they affect the model's reliability and precision.

Exhibit 2.17. Systems and System Models / Systems Thinking

Systems and System Models / Systems Thinking: A system is an organized group of related objects or components; system models can be used for understanding and predicting the behavior of systems.

Grade 4	Grade 8	Grade 12
C4.5: To explain or make predictions about a phenomenon, it often helps to develop a model of a system of related parts, each of which plays some role in the phenomenon.	C8.7: A system model specifies the essential components and quantities involved in a phenomenon and the relationships or interactions between them. The model includes both material and conceptual aspects of the system, such as forces between objects or relationships between species. System models can help to analyze and explain a phenomenon, and, after testing, to make predictions about the phenomenon. C8.8: Systems may interact with other systems; they may have	C12.6: A system model is used to explain or simulate and predict phenomena that occur in the system. A system model defines a boundary for each system or subsystem and delineates and, where relevant, quantifies all necessary parts of the system. The parts include both invisible features such as forces or flows and transfers of energy or information. Such models may include equations that describe relationships between relevant quantities in the system. C12.7: Engineered systems are designed to achieve particular

Grade 4	Grade 8	Grade 12
	subsystems and be a part of larger, more complex systems. C8.9: Engineers design systems to achieve particular functions or do specific items. An engineering design plan includes a system model. Engineers also use system models to troubleshoot system failures.	functions. Such systems may be specific objects (e.g., a satellite) or involve large-scale networks of objects (e.g., a transportation system).

Conservation, Flows, and Cycles: Tracking Energy and Matter

Energy and matter are essential concepts in all disciplines of science and engineering, often in connection with systems. The supply of energy and of each needed chemical element restricts a system's operation. For example, without inputs of energy (sunlight) and matter (carbon dioxide and water), a plant cannot grow. It is informative to track the transfers of matter and energy within, into, or out of a system.

In many systems, there also are cycles of various type—for example, water going back and forth between Earth's atmosphere and its surface and subsurface reservoirs. Any such cycle of matter also involves associated energy transfers at each stage, so to fully understand the water cycle, one must model not only how water moves between parts of the system but also the energy transfer mechanisms that are critical for that motion.

Consideration of energy and matter inputs, outputs, and flows or transfers within a system or process are equally important for engineering. A major goal in design is to maximize certain types of energy output while minimizing others, in order to minimize the energy inputs needed to achieve a desired item.

Exhibit 2.18. Conservation, Flows, and Cycles: Tracking Energy and Matter

Conservation, Flows, and Cycles: Tracking Energy and Matter: Tracking energy transfers and matter flows into, out of, and within systems helps one understand their system's behavior.

Grade 4	Grade 8	Grade 12
C4.6: To understand the function of a system, it is often useful to keep track of the flows and cycles of matter into, out of, and within the system. The only way that the total weight of matter in a system can	C8.10: Matter is conserved because atoms are conserved in physical and chemical processes. C8.11: Energy manifests itself to our observation in multiple different ways, including in mechanical, thermal, electrical,	C12.8: Flows of matter and transfers of energy into, out of, and within a system are analyzed and described using a system model. The amount of matter or energy in any system changes only by flow of matter

Grade 4	Grade 8	Grade 12
change is by flow of matter into or out of the system.	and chemical processes. Energy can transfer between these different observed effects and between objects or systems. C8.12: To analyze the function or behavior of a system, it is often useful to track and model the energy transfers and matter flows. Within any natural or designed system, transfers of energy are needed to drive any motion or cycling of matter.	or transfer of energy into or out of the system. C12.9: Tracking of matter flows and energy transfers is useful because the availability of matter and/or energy within a system limits what can occur and regulates how the system functions.

Relationships Between Structure and Function

Structure and function are complementary properties. The shape and stability of structures of natural and designed objects are related to their function(s). The functioning of natural and built systems depends on the shapes and relationships of key parts as well as on the properties of the materials from which they are made. The selection of an appropriate scale depends on the question being asked. For example, the substructures of molecules are not particularly important in understanding the phenomenon of pressure, but they are relevant to understanding why the ratio between temperature and pressure at constant volume is different for different substances.

Understanding how a bicycle works involves examining the structures and their functions at the scale of the frame, wheels, pedals, and so on. However, building a bicycle may require knowledge of the properties (such as rigidity and hardness) of the materials needed for specific parts of the bicycle. In that way, the builder can change the heaviness of the bicycle by using less dense materials with appropriate properties. This pursuit may lead in turn to an examination of the atomic-scale structure of candidate materials. As a result, new parts with the desired properties can be designed and fabricated.

Exhibit 2.19. Relationships Between Structure and Function

Relationships Between Structure and Function: The way an object is shaped or structured determines many of its properties and functions.

Grade 4	Grade 8	Grade 12
C4.7: Different materials have different substructures, which can influence how they behave (function).	C8.13: Complex macroscopic and microscopic structures within systems can be visualized and modeled. These structures and their relationships influence	C12.10: The functions and properties of natural and designed objects and systems can be inferred from their overall structure, the way their components are shaped and

Grade 4	Grade 8	Grade 12
C4.8: Within any system, natural or designed, the structures of objects, their composition, influences the overall function of the system and its subsystems.	how the system and its subsystems behave. C8.14: Structures can be designed to serve particular functions by taking into account properties of different materials and how materials can be shaped and used.	interconnected, and the molecular substructures of various component materials. C12.11: Designing new systems or structures requires a detailed examination of the properties of different materials and intentional design of the shapes and structures of different components and of connections between and among components.

Conditions for Stability and Change in Systems

Stability and change are the primary concerns of many, if not most, scientific and engineering endeavors. Stability denotes a condition in which some aspects of a system are unchanging, at least at the scale of observation. Such stability can take different forms, with the simplest being a static equilibrium, such as a ladder leaning on a wall. By contrast, a system with steady inflows and outflows (i.e., constant conditions) is said to be in dynamic equilibrium. A dam may be at a constant level with steady quantities of water coming in and out. A repeating pattern of cyclic change (e.g., the moon orbiting Earth) can also be seen as a stable situation, even though it is clearly not static.

An understanding of dynamic equilibrium is crucial to understanding the major issues in any complex system—for example, population dynamics in an ecosystem or the relationship between the level of atmospheric carbon dioxide and Earth's average temperature. Dynamic equilibrium is an equally important concept for understanding the physical forces in matter. Stable matter is a system of atoms in dynamic equilibrium.

In designing systems for stable operation, the mechanisms of external controls and internal feedback loops are important design elements; feedback is important to understanding natural systems as well. A feedback loop is any mechanism in which a condition triggers some action that causes a change in that same condition, such as the temperature of a room triggering the thermostatic control that turns the room's heater on or off.

A system can be stable on a small-time scale, but on a larger time scale it may be seen to be changing. For example, when looking at a living organism over the course of an hour or a day, it may maintain stability; over longer periods, the organism grows, ages, and eventually dies. For the development of larger systems, such as the variety of living species inhabiting Earth or the formation of a galaxy, the relevant time scales may be very long indeed; such processes occur over millions or even billions of years. Example systems that are appropriate for each grade can be found in the NAEP DCs in Chapter 2, the sample items in Chapter 3, and the NAEP Science Assessment and Item Specifications.

Exhibit 2.20. Conditions for Stability and Change in Systems

Conditions for Stability and Change in Systems: For both designed and natural systems, conditions that affect stability and factors that control rates of change are critical elements to consider and understand.

Grade 4	Grade 8	Grade 12
C4.9: Change in conditions can be described or predicted for a stable or ongoing situation (e.g., a growing plant, a healthy body).	C8.15: Stability or change over time in a system depends on external conditions as well as on relationships and conditions within the system. C8.16: Systems can appear stable on one time scale but viewed on a longer time scale are seen to be changing.	C12.12: Rates of change are quantifiable and are important quantities to consider in modeling any system. C12.13: Feedback mechanisms within a system are important elements for explaining or designing for either the stability or instability of the system. C12.14: Changes in a system can be caused by changes in other systems or in conditions affecting the system as well as by prior changes within the system. The scale of the effect is not always comparable to that of the cause but may be much larger or smaller.

Chapter 2 describes what students should be able to know and do with respect to the three dimensions of science achievement. Chapter 3 provides guidance on how these dimensions can be combined to create assessment items.

CHAPTER THREE: Assessment Design

This chapter provides an overview of the major components of the science assessment design, which includes the types of assessment items and how they can be used to expand the ways in which students are asked to demonstrate what they know and are able to do in science. In addition, this chapter describes how the assessment is distributed across the disciplines and practices described in Chapter 2. Chapter 3 is organized into the following sections:

- 3A. Types of Items
- 3B. Distribution of Items
- 3C. Scientific Sensemaking in NAEP Science
- 3D. Features of Phenomena and Problems Used in Item Contexts
- 3E. Features of Multidimensional Items
- 3F. Assessing the Full Range of Student Performance
- 3G. Reflecting a Wide Range of Students
- 3H. Science Achievement Expectations
- 3I. Digital Tools

Several sections in this chapter include sample items as illustrative examples of the major components of assessment design. Appendix A includes additional details on each of the sample items, including grade; discipline; item type; alignment to DCs, SEPs, and CCCs; complexity; and scoring information (where applicable). Additional sample items and description of important aspects of assessment design are included in the NAEP Science Assessment and Item Specifications.

A Note About Sample Items: The sample items included in this framework have been created by framework panelists or staff or modified from items originally designed for a different purpose. The sample items were developed to illustrate important aspects and features of this framework. Items that are included on operational NAEP assessments go through rigorous item development procedures in a multiyear process, which includes reviews by multiple stakeholders, pretesting, piloting, and iterative revisions. The sample items included in this framework have not gone through the item development process for NAEP that will be used for actual items created for the operational NAEP science assessments based on this framework.

3A. Types of Items

The essential element of any test is an ***item***, the basic scorable part of assessment. The NAEP Science Framework provides recommendations and guidelines for developing items for the 2028 NAEP Science Assessment for a broad audience. A technical specifications document that accompanies this framework will describe in greater detail how items are to be developed and used in the overall design of the 2028 NAEP Science Assessment. In brief, items will be constructed according to the following guidelines:

- The assessment will include a variety of item types, including selected response and constructed response formats, discrete and multipart standalone items, item sets, and scenario-based tasks.
- Each item will assess science achievement in the context of a phenomenon or problem.
- The performance required by each item will involve sensemaking about the phenomenon or problem. No item will assess only rote content or procedural knowledge.
- A two-dimensional item will include a NAEP DC and SEP.
- A three-dimensional item will include all three dimensions: a NAEP DC, SEP, and CCC.
- Items will be constructed with different levels of complexity to assess students with a wide range of knowledge and skills. See section 3F for more details and sample items.
- The assessment as a whole will be responsive to learners with rich and diverse cultural and linguistic backgrounds, identities, and learning environments. That is, as much as possible, the assessment will give students opportunities to leverage references to their own cultures in responding to items (Stembridge, 2019).

Items will be either *selected response items,* in which students choose a response from provided options, or *constructed response items*, in which students respond by generating an original response. These are further divided into subcategories as follows. It is important to note the need for NAEP Science to be consistent with NAEP program guidelines and principles to ensure that the science assessment is consistent with other NAEP assessments and the program more broadly.

Selected Response Items require a student to select one or more response options from a given, limited set of choices. Different types of selected response items that may be used on the 2028 NAEP Science Assessment include the following:

- **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., an image) into target fields (e.g., a table).
- **Zones**: Students respond by selecting one or more regions on a graphic stimulus.
- **Grid**: Students evaluate statements, such as claims or explanations, or classify components of a system based on their properties or interactions. The answer is entered by selecting cells in a table.
- **Inline choice**: Students respond by selecting one option from one or more drop-down menus that may appear in various sections of an item.

Constructed Response Items are generally more challenging than selected response items because the correct and alternative answers are not provided as part of the item. Constructed response item types that may be used on the 2028 NAEP Science Assessment are listed below.

- **Short constructed response**: Students respond by giving a short response, from a single word or number to a few sentences. For example, students may label a model or system, classify data, or describe a pattern in a system or dataset.
- **Extended constructed response**: Students respond by giving a description or explanation that requires more than a few words. For example, students may explain a system model with supporting evidence, synthesize information from multiple sources, or describe a process with multiple components or interactions.

Item scoring is straightforward for selected response items, which can be scored by machine. However, the large number of constructed response items require interpretation of open-ended responses. Every constructed response item has a scoring guide that defines the criteria used to evaluate students' responses. Some short constructed response items can be scored according to guides that permit partial credit, while others are scored as either correct or incorrect. All constructed response scoring guides are refined from work with a sample of actual student responses gathered during item pilot testing.

NAEP assessments use a variety of item types to fully assess students' knowledge and skills. Varying the item types students engage with on the assessment is essential to balance complexity, time on task, and validity and reliability considerations. Different types of items or groups of items are used for different purposes. Discrete items, multipart items, item sets, and scenario-based tasks may all use any combination of the item types described above. Test questions are arranged either as short, separate items or in groups. These arrangements include the following:

- A ***discrete item*** **(DI)** is a single, standalone item. Students need to be able to read the stimulus/prompt and answer the question in no more than a few minutes. Compared with other item types, discrete items allow for a large number of items to be included on the assessment, increasing the reliability of the assessment. Examples of discrete items are included below in section 3E.
- A ***multipart item*** **(MPI)** has a few parts that are dependent on each other. For example, a multipart item might ask students to make a choice or decision and follow up with another question to explain their reasoning. Multipart items take somewhat more time than discrete items, but they can probe for deeper understanding than discrete items. Since multipart items are aimed at different aspects of a single performance, they generally receive a single score that may consist of multiple points. Examples of multipart items are included below in section 3E.
- An ***item set*** **(IS)** uses common stimulus material to ask a group of independent questions. Item sets make it possible to take advantage of efficiency by presenting rich and engaging stimulus material, then asking several questions to collect evidence through a number of different items. Since the items do not depend on each other, questions in an item set each receive a separate score. If an item is rejected during pilot testing because it is found to not be functioning as intended, the other items in the set may be preserved. Although not

a strict requirement, it is expected that item sets will play a prominent role in the implementation of this framework. Groups of independent items that make use of some common phenomena and problems may provide the best balance of breadth and depth by creating opportunities to measure related but distinct content with independent items. Item sets should include at least one item that is three-dimensional.

- A ***scenario-based task*** **(SBT)** includes a sequence of items presented through an unfolding context, often with rich and engaging stimulus material such as images and video. SBTs are often interactive, asking students to respond to several short tasks and questions. However, the task does not have to be interactive to be a scenario-based task. SBTs typically present meaningful and compelling phenomena and problems, including those that require a large amount of background information. Scenario-based tasks should include at least one item that is three-dimensional.

A note about SBTs: While scenario-based tasks can be rich opportunities for student sensemaking, they are often more resource-intensive to develop. SBTs should be used judiciously, with a particular focus on those performances that are difficult to assess in other contexts.

These may include:

- **Scenarios that require considerable contextual information to fairly surface the intended targets**. Assessing some NAEP SEPs (e.g., Using Mathematics and Computational Thinking; Obtaining, Evaluating, and Communicating Information), DCs, and CCCs requires a large amount of contextual information for students to engage in the task. This is often because the assessment target itself does not specify specific contexts, methods, models, or experiences students should focus on in instruction; as a result, this information needs to be provided in the item, with enough context so that any student can understand the information. While it is feasible to assess these NAEP SEPs with discrete items, the amount of information students need makes it difficult to justify for a single item. In some cases, this can be addressed through an item set; in others, SBTs will be particularly helpful.
- **Phenomena/problems and assessment targets that require iteration**. Particularly in grades 8 and 12, students are expected to iteratively engage with information, updating their sensemaking with new/multiple sources of information. SBTs can be particularly useful in these contexts.
- **High-complexity and increasingly complete performances.** Many high-complexity performances involve cascades of NAEP SEPs, CCCs, and DCs. These often draw on NAEP DCs from multiple domains, SEPs that connect with each other (e.g., Constructing Explanations and Designing Solutions based on multiple sources of information of varying credibility), and complex CCCs; student performances on items early in the SBT may include fewer components of these DCs, SEPs, and CCCs than items later in the SBT and, over the course of the entire SBT, student performances may provide evidence

that addresses these dimensions in a more complete way. These performances are often particularly necessary for addressing authentic phenomena and problems using the specified dimensions in grades 8 and 12.

Different item types can be used in combination to create an assessment that appropriately balances breadth and depth of content coverage, while also accounting for measuring a construct that requires time for students to process the phenomenon or context necessary for sensemaking.

3B. Distribution of Items

Balance by Disciplinary Concept Grouping

The distribution of items by discipline should be approximately equal across Physical Science, Life Science, and Earth and Space Sciences at all grades. With respect to NAEP CCCs and SEPs, the emphasis should be on meaningful representation rather than a strictly equal distribution. When an authentic query requires only an application of a NAEP SEP to a NAEP DC, a two-dimensional item is acceptable.

Exhibit 3.1. Approximate Distribution of Items by NAEP DC Grouping and Grade

Percentage of items	Grade 4	Grade 8	Grade 12
Physical Science	33.3%	33.3%	33.3%
Life Science	33.3%	33.3%	33.3%
Earth and Space Sciences	33.3%	33.3%	33.3%

Balance by Response Type

The assessment will consist of about 65 percent selected response items and 35 percent constructed response items. Since items requiring a constructed response take a longer time to answer, it is anticipated that the amount of time students spend answering selected response items and constructed response items will be approximately equal.

Exhibit 3.2. Approximate Distribution of Items by Response Type

Type of Response	Distribution
Selected response	65%
Constructed response	35%

Distribution of NAEP SEPs and CCCs

In doing science or engineering, the eight NAEP SEPs are used in an iterative and recursive cycle that often blurs the boundaries between them. For NAEP assessment purposes, the NAEP SEPs will be grouped into four categories, labeled Investigating, Analyzing, Explaining, and

Evaluating, as shown in Exhibit 3.3 below. These groupings put together practices most often used with a common purpose.

Exhibit 3.3. Grouping of NAEP SEPs

Investigating	Asking Questions and Defining Problems
	Planning and Carrying Out Investigations
Analyzing	Analyzing and Interpreting Data
	Using Mathematics and Computational Thinking
Explaining	Developing and Using Models
	Constructing Explanations and Designing Solutions
Evaluating	Engaging in Argument from Evidence
	Obtaining, Evaluating, and Communicating Information

To ensure that a variety of the practices are used throughout the assessment, item developers should use a minimum of 10 percent of the items at each grade level from each of the four categories of NAEP SEPs but may otherwise choose SEPs that work well within other item design considerations. Similarly, all seven NAEP CCCs should be used in items where appropriate for the item and grade level. More guidance about using NAEP CCCs with DCs and SEPs can be found in the NAEP Science Assessment and Item Specifications.

3C. Scientific Sensemaking in NAEP Science

An essential aspect of all test items is that they will surface *sensemaking* (Exhibit 3.4). In contrast to items that measure a student's ability to recall rote knowledge, items that measure sensemaking require students to actively apply NAEP DCs, SEPs, and CCCs together to figure out a phenomenon or address a real-world problem. Sensemaking can be considered as the binding agent that connects the three dimensions of science to the definition of science achievement and the NAEP Science construct described in this framework.

Items that require sensemaking enable students to demonstrate that they deeply understand and can apply the NAEP DCs to explain a phenomenon or address a problem. Doing so requires that all items present either a phenomenon or a problem that invites sensemaking. The role of phenomena and problems in sensemaking is illustrated in Exhibit 3.4.

Exhibit 3.4. Visualizing the Sensemaking Process[3]

This graphic is fully described in Appendix D.

---

[3] Adapted from Achieve. (2019b). Task Annotation Project in Science: Sense-making. Retrieved from https://issuu.com/achieveinc/docs/sense-making_02142019__7_

3D. Features of Phenomena and Problems Used in Item Contexts

In this framework, an assessment designed to measure science achievement requires students to demonstrate scientific knowledge while engaging in the practices of science and engineering—that is, scientific sensemaking using NAEP DCs, SEPs, and CCCs. To do so, all two-dimensional and three-dimensional items are designed to focus on phenomena and/or problems. Without a phenomenon or problem at the center of an assessment item, there is nothing for students to make sense of, problem-solve about, or apply their knowledge to.

Compelling phenomena and problem-based contexts present authentic uncertainty of a situation in ways that give students something to make sense of. Phenomena and problems that are considered compelling often have some of the following features:

- a specific instance
- authentic uncertainty
- relevance to particular communities

Note that compelling phenomena and problems in assessment contexts focus on explicit relevance, not on individual student interest or a specified degree of impact. A phenomenon may be compelling without each student being deeply invested in the outcome or without the phenomenon having a huge global impact.

Phenomena provide a setting for an item or set of items. They should be chosen to engage student attention and sensemaking that requires the targeted NAEP DCs, SEPs, and CCCs for a satisfying explanation or effective solution. Problems are meaningful challenges that present a situation requiring new or improved technologies or processes. Where appropriate, phenomenon and problem descriptions should include the impact, such as effects on people, animals, or the environment. Phenomena may result from both human-designed and natural processes and systems. To serve as the context of an item, phenomena and problems must involve a NAEP DC identified in Chapter 2.

From the perspective of the student taking the assessment, they are answering questions about what, why, or how something occurs or what to do about a problem. Compelling phenomena and problems promote student engagement in items and demonstration of their knowledge and skills. They do so by providing contexts that make the authentic uncertainty of a situation clear to students and by giving students something puzzling to solve.

Criteria for Selecting High-Quality Science Phenomena and Problems

Phenomena and problems provide the context for all NAEP Science items. Some contexts will be short and simple; for example, they will have one or two sentences and one or two images. Other contexts will present more complex phenomena and problems or support a broader range of items. High-quality phenomena and problems are important for science assessments because they provide access points for students, ensuring that all students can make their thinking visible and that assessments are accessible, and providing opportunities for all students to show what they know and are able to do. Following are criteria and guidelines for choosing high-quality phenomena and problems.

High-quality items based on phenomena and/or problems (a) position items to be compelling and motivating to students, (b) cue students toward the targeted dimensions they need to apply, (c) help students with different and diverse prior learning and lived experiences understand what they are being asked to do, and (d) provide scaffolds for students to engage and demonstrate their understanding. In this way, high-quality phenomenon- or problem-based items are essential to truly surface what all students know and are able to do and to ensure that scores are trustworthy representations of students' knowledge and skills in science.

While the exact nature of contexts will depend on what NAEP DCs and SEPs are intended to be elicited, some common features of high-quality contexts for scientific sensemaking include the following:

- Focus on a specific, observable, and/or measurable event(s) that is relatable and motivating to students.
- Use an authentic question or other prompt that leads the student to use the targeted NAEP DC and SEP (and CCC when appropriate) to explain the phenomenon or figure out a solution to the problem.
- Provide just the right amount of information about the phenomenon or problem that enables the student to engage their thinking, but not too much to be distracting.
- The context should be accurate and presented in an engaging way through text, images, video, or other means to engage student interest.
- The length of a phenomenon or problem description should scale with the scope of the assessment item. The context for a discrete item will be shorter than that for an item set or scenario-based task. The most important consideration is that the context is appropriate to measure the item-level targets.
- Require the appropriate level of conceptual understanding as described in Chapter 2, but not highly specific or technical levels of understanding beyond what students are expected to bring to the assessment.
- Avoid an additional cognitive burden by not asking students to hold a lot of contextual information in working memory or determine which pieces are relevant for each item.
- Do not give away the punchline. Avoid including information that students should have been expected to bring to the table. Leave space for students to demonstrate their understanding and not only their reading and logical reasoning skills.

Creating Contexts for Different Types of Items

The context for discrete selected response items should provide just enough information for the student to select the response that answers the question. For example, if the item is about data analysis, the context will need to provide data to analyze; if it is about making a claim from evidence, the context will need to provide evidence. In multiple-choice questions, the answer choices themselves are also part of the information students use to understand and engage with the item and should be designed accordingly.

Discrete constructed response items may ask students to engage more comprehensively in practices such as modeling, explaining, or arguing from evidence. Such items elicit a wide range of performances that allow for more expansive sensemaking than selected response items. Therefore, contexts for these items may provide more information. Like all items, the information should be only what is needed to engage with the item.

Multipart items, item sets, and scenario-based tasks will typically require more expansive contextual information to support a wider range of performances and to compel student sensemaking throughout the set of items. This may begin with an observation of a phenomenon, such as a volcanic eruption, or a meaningful problem, such as preventing a pandemic. Such contexts will often be richer; involve more text, images, and data than contexts for discrete items; and include multiple uncertainties that can be leveraged across many items. For such complex items, the context can be revealed one step at a time, providing just the amount of information needed to answer the next question or complete the next part of the item, so as not to burden the student with too much information to retain as they deploy their sensemaking abilities.

Language Considerations in Contexts

Assessments that present phenomena and problems to enable sensemaking often require more language use (reading, writing) than do traditional assessments focused on recall and memorization. While this is necessary both to better engage learners and to elicit student sensemaking, attending to some specific considerations for language use can ensure that all learners can successfully engage with the assessment item. For example:

- Use only as many words as needed to convey a compelling and necessary context.
- Use narrative, expository, and scientific types of language and vary it to make it appropriate for the context.
- Use everyday language and active voice where possible.
- Analyze the reading level to ensure it is accessible to the vast majority of students.
- Use a variety of modalities to convey information, such as text, images, and video.
- Avoid using words that have different scientific and colloquial meanings.
- Use similar language conventions within and across disciplines.
- Research about how students learn should be taken into consideration in scoring guidance.

It is important to note that NAEP Science is an assessment of science in English. The considerations listed above not only benefit all students but have the potential to reduce the cognitive load for students who may be simultaneously learning English and science in schools (e.g., designated English Learners, emergent bilingual learners, etc.).

It should also be noted that reading and writing (and speaking and listening) in science are central elements of many of the NAEP SEPs, and therefore are part of expectations for science learning and achievement. NAEP SEPs such as Obtaining, Evaluating, and Communicating Information; Engaging in Argument from Evidence; and Constructing Explanations and Designing Solutions require students to develop and demonstrate disciplinary literacy in science. While items should take care to limit any gratuitous language use or complexity, it is equally important that NAEP not overly limit language engagement so that the construct cannot be meaningfully assessed.

Deciding What to Include and Exclude in the Context

The 2028 NAEP Science Framework is designed to enable students to demonstrate their conceptual understanding of NAEP DCs and CCCs and to use NAEP SEPs. However, students are generally ***not*** expected to know or recall specifics of a given phenomenon or a specialized topic. For example, students may be expected to understand that bodies have systems that work together, but not the specific parts and functioning of the human digestive system. This means that in any given item, students will often need to be provided with additional contextual information for them to fully understand and access the question or perform the required item and apply their conceptual understanding appropriately. Chapter 2 of this framework identifies the information that students should bring to the table. All other details required for satisfactory responses would need to be provided in the context. Details provided to students in item contexts should be carefully constructed in developmentally appropriate ways, accounting for reading and cognitive load, as well as assumptions about prior knowledge and schema development of the targeted students for the test. The following three illustrative contexts demonstrate ways that context can support students in applying their conceptual understanding. The amount and complexity of the context varies across these three examples to demonstrate what is appropriate for different grade levels, different targeted complexities of items, and different numbers of items (discrete or item set).

Illustrative Contexts

- **Park Flooding** includes a simple phenomenon that can be used for a discrete or multipart item.
- **Human Migration to Appalachia** includes an example of a richer phenomenon with authentic, compelling phenomena that can support an item set.
- **Locusts** includes an example of a robust phenomenon that can be used not only for multiple items (item set) but across multiple disciplines.

The first of these illustrative contexts is included in Exhibit 3.5, an item that requires students to make sense of a simple phenomenon: a local park has flooded. This phenomenon is an example of an everyday phenomenon that many students may have directly experienced or have sufficient experiences to understand. This item elicits simple sensemaking because students must understand that flooding is most likely to occur on the day with the most rain, rather than just any day with rain. This requires interpreting the context as well as the data, albeit in a very straightforward way. In grade 4, students are still developing the concept of flooding and its relationship to rain; while this item elicits sensemaking based on developmentally appropriate and expected schema for 4th graders, this same performance would not be considered sensemaking in grades 8 and 12.

Exhibit 3.5. Park Flooding, Version 1

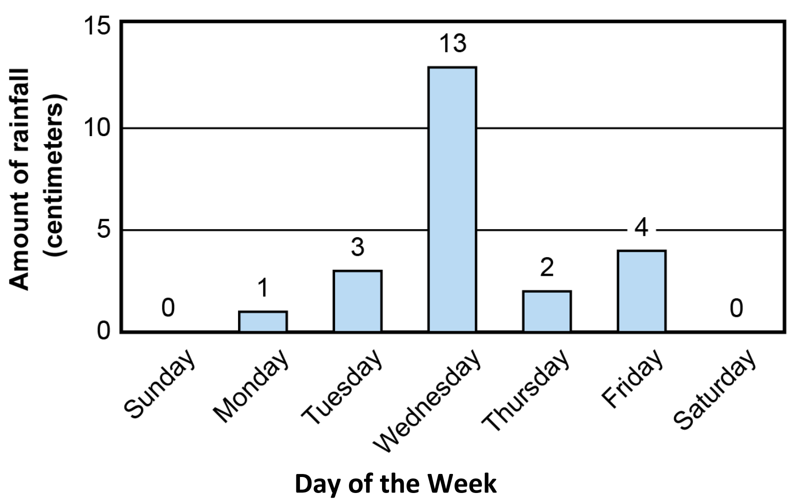
Item ID: Park Flooding (adapted from Formative Assessment Bundling Literacy and Elementary Science)
Phenomenon: A park flooded when it was raining one day but not other days.

People visiting a local park noticed that the park was flooded and was closed for the day. The picture shows the flooded park.

Flooded Park[4]

The park was closed only on the day the flooding happened. The bar graph shows the rainfall for each day of that week.

Amount of Rainfall at the Park



[4] Betty Longbottom / Flooded Playground! - Cliffe Avenue / CC BY-SA 2.0

Based on the data, on which day was the park **most likely** flooded?

A. Sunday
B. Monday
C. Tuesday
D. Wednesday

Key: D

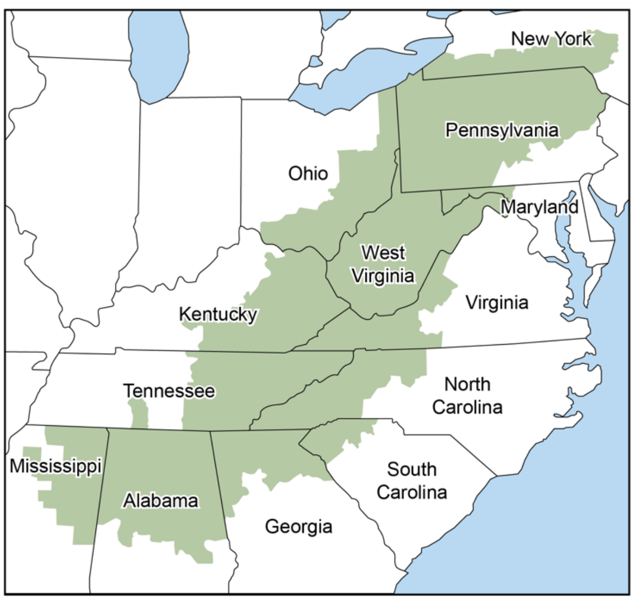
The second illustrative context, Exhibit 3.6, presents a stimulus with a meaningful phenomenon and problem context that deeply matters to people around the world, poses considerable challenges, and presents authentic and compelling problems. Sample items that use this stimulus are provided in Exhibit 3.12.

Exhibit 3.6. Human Migration to Appalachia

Phenomenon/Problem: Human migration to Appalachia has been greater than predicted by computer models. The model used to make the prediction needs to be revised to better reflect the factors that influence migration into different regions of Appalachia.

Appalachia is considered "climate resilient." This means that the area can successfully handle the impacts of changes to climate and can prevent those impacts from growing worse. The green areas in Figure 1 show where Appalachia is located in the United States.

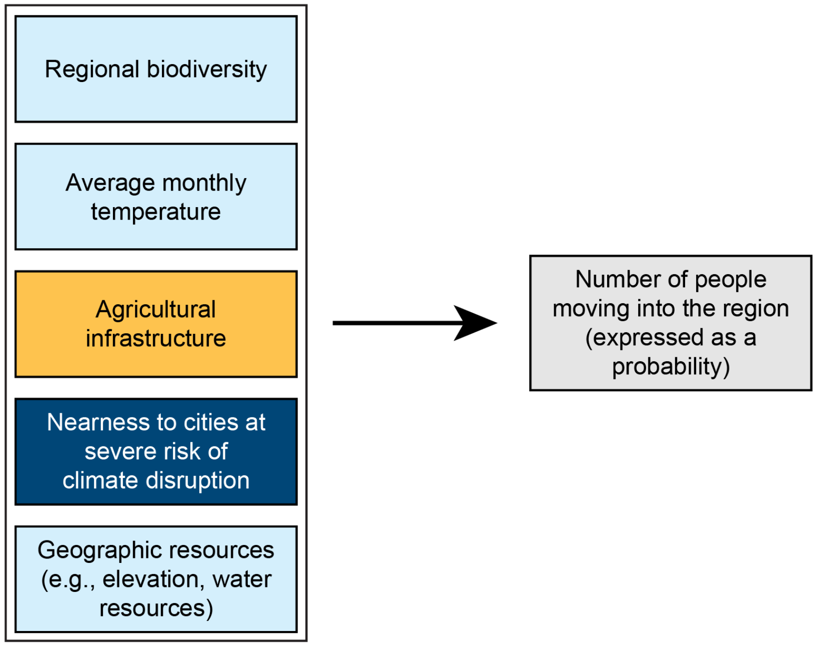
Figure 1. Map of Appalachia



Computational models predict that many people will move into the Appalachian region over the next 20 years as they seek to find places to live that are safer and more stable.

Figure 2 shows one model local leaders are using to predict how many people will move into Appalachia. Blue indicates factors that are expected to increase migration, and orange indicates factors expected to decrease migration. Darker colors indicate more weight on that factor in the model. Agricultural infrastructure includes farms, markets and businesses that support farms, and the transportation and communication systems in the area.

Figure 2. Computational Model for Predicting Human Migration



When this model was tested against recent population growth due to migration into two locations in Appalachia, leaders noticed some differences between what the model predicted and what the data showed. The table shows these differences for the two locations, along with information about how high or low each location is rated on several factors.

Predicted and Actual Population Growth in Pittsburgh and the Shenandoah Valley

Location	Pittsburgh	Shenandoah Valley
Predicted population growth	high	low
Actual population growth	low	high
Relative biodiversity	low	high
Average monthly temperature range	29–73° F	32–74° F
Relative agricultural infrastructure	low	medium
Nearness to climate-impacted urban centers	high	high
Access to usable water	medium	high

The third illustrative example, Exhibit 3.7, is a stimulus that presents a meaningful phenomenon and problem context that deeply matters to many people around the world and is posing considerable challenges right now. It was selected to show how a wider range of NAEP SEPs and CCCs can be engaged in items across an item set than the examples above, including items across multiple disciplines (e.g., Life Science and Earth and Space Sciences), and some that are often difficult to assess. An item set based on this stimulus could also be used to address the underlying biology (e.g., genetics, specialized subsystems) connected to the physiological changes locusts undergo; research on potential solutions; impacts on biodiversity in regions with swarming; consideration of patterns of locust swarming going back thousands of years (stability and change); consideration of whether current upticks are significant or not (more sophisticated data analysis); and so on. This context could support items for either grades 8 or 12 in both Life Science and Earth and Space Sciences. Possible NAEP DCs, SEPs, and CCCs that could be included with the stimulus are listed in Appendix A.

Exhibit 3.7. Locusts

Phenomenon and engineering design problem: Locust swarms can cause a lot of damage.

A desert locust is an insect that undergoes changes to its body in certain environmental conditions. Figure 1 describes some differences between two modes of a desert locust.

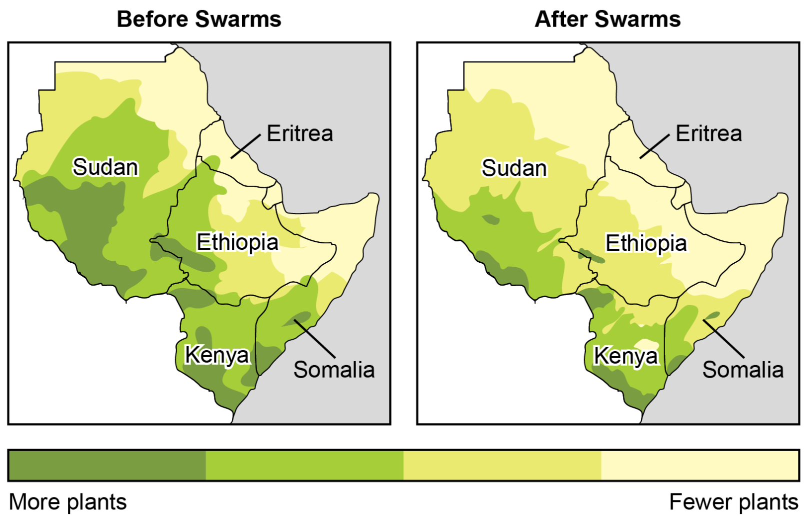
Figure 1. Desert Locust[5]

Mode 1: Grasshopper (Dry, warm or cool weather)	Mode 2: Locust (Wet/rainy, warm or hot weather)
• Behave independently • Stay away from other desert locusts • Mostly walk slowly and jump • Limited diet • Small, scattered populations that stay in one place • Very stable population; females lay eggs but most don’t hatch until the environment is wet and hot.	• Behave as a united group (swarms) • Gather together with other desert locusts • Walk quickly and fly long distances • Broad diet, including crops • Tens of billions of locusts in a swarm that can travel up to 100 miles per day • Population can increase 400x in six months.

[5] Bernard DuPont Bird Locust Attribution-ShareAlike (CC BY-SA 2.0); Magnus Ullman, CC BY-SA 3.0, via Wikimedia Commons

When these insects are in Mode 2, they are able to swarm. A single swarm of locusts can cover an area of up to 100 square miles, with 40 to 80 million locusts in each square mile. Swarms can travel up to 100 miles a day. Figure 2 shows the effect of three months of locust swarms on available vegetation in an area of Africa.

Figure 2. Available Vegetation Before and After Locust Swarms



3E. Features of Multidimensional Items

As described previously, measuring the construct described in this framework requires that each item requires students to bring together NAEP DCs, SEPs, and when possible, CCCs to successfully address a question or accomplish a task. Following are some of the questions that should guide development of multidimensional items:

- Is there an appropriate phenomenon or problem driving student thinking and responses?
- Does the item require students to demonstrate an understanding of at least one NAEP DC?
- Does the item require students to demonstrate their understanding of the NAEP DC through application of a NAEP SEP and a CCC?
- Does the student need to engage in sensemaking to explain a phenomenon or solve a problem?
- Is the understanding appropriate to the grade level being assessed?

Each discrete item and each multipart item should be at least two-dimensional and three-dimensional if appropriate. Item sets and scenario-based tasks should provide evidence of students' ability to use the three dimensions together to explain a phenomenon or address a problem by including at least one item that is three-dimensional. Each item (discrete or individual items within item sets and scenario-based tasks) will receive one score representing the integration of the dimensions measured by the item.

The following four illustrative examples demonstrate how items require students to use two or three dimensions together in different types of discrete items and how an item set can require students to use different combinations of dimensions for different items. The first two items involve multiple parts to illustrate how to capture more evidence of student understanding through discrete items; however, these items could be limited to the first part to serve as a single part discrete item if test developers determine that this would provide sufficient evidence of student understanding.

Illustrative Examples

- **Plant Growth** is a two-dimensional discrete, multipart item.
- **Park Flooding** is a two-dimensional discrete, multipart item.
- **Making Soap** is a three-dimensional discrete item.
- **Human Migration to Appalachia** is an item set with items across multiple disciplines, using the three dimensions together.

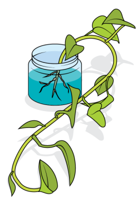
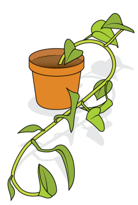
In Exhibit 3.8, an example of a two-dimensional discrete item, students have to apply their understanding of what plants need to grow to make a prediction. Students do not need to understand a specific NAEP CCC element to respond to the item—the NAEP SEP and DC are sufficient to respond to this question. While the item is an implicit example of the NAEP CCC Mechanisms and Explanation: Cause and Effect, students do not need to explicitly bring an understanding of cause-and-effect relationships to respond and, therefore, the item is not considered three-dimensional. This item provides an example of lower-level sensemaking with the NAEP DC and SEP—while students do need to understand (a) that plants need water and air to grow and (b) need to be able to use this information to evaluate a phenomenon across multiple (experimental) conditions, they are very closely applying simple grade-appropriate NAEP DCs and SEPs. The NAEP SEP is engaged in service of surfacing NAEP DC understanding, rather than expanding the nature of how students explain the phenomenon. This level of sensemaking would be appropriate to surface understanding from students who have had the opportunity to begin developing an understanding of the grade 4 NAEP DC and SEP.

Exhibit 3.8. Plant Growth

Item ID: Plant Growth (adapted from the Next Generation Science Assessment Project)

The plants shown were placed in a classroom on the same day. They are all the same kind of plant. The plants were placed on the same side of the room near a window so they receive the same amount of light each day. Students in the class want to find out what the plants need the most in order to grow. They grow the plants using the conditions shown in the table.

Conditions for Growing Plants

Plant	Planted in Soil	Water
 Plant A	No	Water added regularly for one month
 Plant B	Yes	Water added regularly for one month
 Plant C	Yes	No water added

Part A

Which plant will likely grow the **least** over the next month?

A. Plant A
B. Plant B
C. Plant C

Key: C

Part B

Provide one reason the plant you chose in Part A will grow the least over the next month.

Exhibit 3.9. Plant Growth Part B Constructed Response Scoring Notes

- Reasons students provide should leverage understanding of what plants need to grow (water, air, minerals from soil).
- Note that while a complete answer might include comparisons among plants (e.g., Plant A and B have X, but Plant C does not), this is not a requirement.
- Possible reasons include:
    - Plant C does not get water.
    - Plant C does not get minerals.
    - Plant C does not get water or minerals.
- Students should receive credit as long as their reason supports their choice, with an accurate understanding of plant needs for growth.

Exhibit 3.10 presents version 2 of the item in Exhibit 3.5. This version shows a modification that provides additional evidence of student understanding of both the NAEP DC and SEP in service of sensemaking. The additional component of the second discrete item does not change the alignment or complexity of the item, but does add time to complete and some additional reading load. This trade-off between more comprehensive evidence and the additional time needed to complete the items may be valuable at some times over the range of the assessment.

Exhibit 3.10. Park Flooding, Version 2

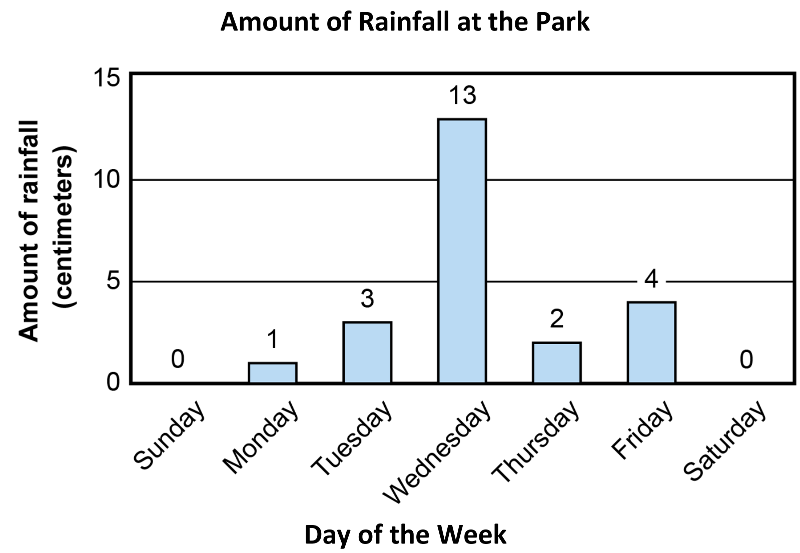
Item ID: Park Flooding (adapted from Formative Assessment Bundling Literacy and Elementary Science)

People visiting a local park noticed that the park was flooded and was closed for the day. The picture shows the flooded park.

Flooded Park[6]

The park was closed only on the day the flooding happened. The bar graph shows the rainfall for each day of that week.

[6] Betty Longbottom / Flooded Playground! - Cliffe Avenue / CC BY-SA 2.0



Part A

Based on the data, on which day was the park **most likely** flooded?

A. Sunday
B. Monday
C. Tuesday
D. Wednesday

Key: D

Part B

Based on your understanding of weather, which piece of evidence **best** supports your answer in Part A?

A. This day was rainy.
B. The rain started falling on this day.
C. This day had more rainfall than any other day in the week did.
D. The amount of rainfall on this day was lower than on any other day.

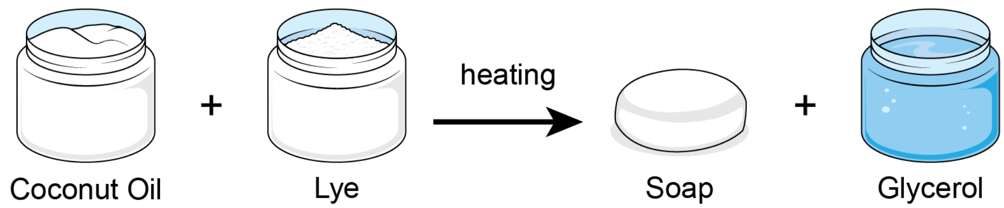
Key: C

In Exhibit 3.11, students must apply their understanding of chemical reactions to analyze data, while looking for patterns among the specific characteristic properties that will indicate that a chemical reaction has occurred. This performance requires that students use their understanding of the NAEP DC, use the NAEP SEP to analyze data, and use the NAEP CCC to look for patterns, to figure out whether a chemical reaction has occurred. This is a three-dimensional performance that requires relatively simple sensemaking with multiple dimensions.

Exhibit 3.11. Making Soap

Item ID: Making Soap (adapted from the Next Generation Science Assessment Project)

One way to make soap is to heat a combination of coconut oil and lye. The diagram shows a simplified model of the soapmaking process.



The data table shows properties of each substance in the model of the soapmaking process.

Properties of Soapmaking Substances

Substance	Mass (g)	Odor	Density (g/cm^3)	Melting point (°C)
Coconut oil	100	Coconut	0.93	27
Lye	20	Odorless	2.13	318
Soap	115	Coconut	0.95	48
Glycerol	5	Odorless	1.26	17.8

Which data provide evidence that making soap involves a chemical reaction?

A. Coconut oil and soap both smell like coconut.
B. The density of soap is different from the density of glycerol.
C. The total mass of soap and glycerol is the same as the total mass of coconut oil and lye.
D. The melting points of soap and glycerol are different from the melting points of coconut oil and lye.

Key: D

Exhibit 3.12 is an item set for grade 12 that demonstrates how NAEP DCs, SEPs, and CCCs can be used in the service of sensemaking. The items leverage simple uses of NAEP SEPs and CCCs, allowing a wider range of students to access and engage with the rich context.

It should be noted that this item set could be expanded to more deeply and comprehensively assess (a) related Earth and Space Sciences DCs (e.g., how the geography of the Appalachian region has contributed to climate resilience over time), (b) the Life Science DCs (e.g., how ecosystem dynamics contribute to resilient biodiversity), or (c) their integrated use (e.g., how geographic features that have evolved over time have led to adaptations and speciation, contributing to the rich biological systems in the area and impacts on human civilizations). These could be explored in further independent items within an item set or in related items within a scenario-based task.

Exhibit 3.12. Human Migration to Appalachia

Item ID: Human Migration to Appalachia

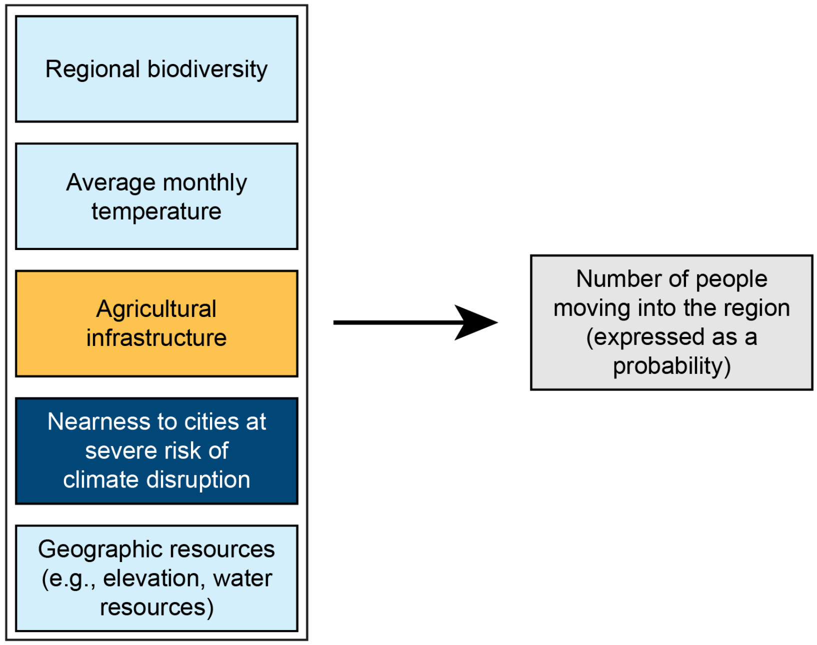
Appalachia is considered “climate resilient.” This means that the area can successfully handle the impacts of changes to climate and can prevent those impacts from growing worse. The green areas in Figure 1 show where Appalachia is located in the United States.

Figure 1. Map of Appalachia

Computational models predict that many people will move into the Appalachian region over the next 20 years as they seek to find places to live that are safer and more stable.

Figure 2 shows one model local leaders are using to predict how many people will move into Appalachia. Blue indicates factors that are expected to increase migration, and orange indicates factors expected to decrease migration. Darker colors indicate more weight on that factor in the model. Agricultural infrastructure includes farms, markets and businesses that support farms, and the transportation and communication systems in the area.

Figure 2. Computational Model for Predicting Human Migration



When this model was tested against recent population growth due to migration into two locations in Appalachia, leaders noticed some differences between what the model predicted and what the data showed. The table shows these differences for the two locations, along with information about how high or low each location is rated on several factors.

Predicted and Actual Population Growth in Pittsburgh and the Shenandoah Valley

Location	Pittsburgh	Shenandoah Valley
Predicted population growth	high	low
Actual population growth	low	high
Relative biodiversity	low	high
Average monthly temperature range	29–73° F	32–74° F
Relative agricultural infrastructure	low	medium
Nearness to climate-impacted urban centers	high	high
Access to usable water	medium	high

Item 1

Based on your understanding of what human societies need to be successful, which idea **best** explains why the model prediction was different from the observed data?

A. Pittsburgh has highly limited access to water. The model did not account for how this shortage would limit the growth of local businesses.
B. Human societies require reasonable annual temperature ranges. The model incorrectly assumed that average monthly temperature was a less important factor.
C. Human societies require access to food and water. The model did not account for how important access to food and water would be for human migration to Appalachia.
D. The Shenandoah Valley has very little access to fresh fruits and vegetables. The model incorrectly assumed that agricultural infrastructure would not promote migration.

Key: C

Item 2

Human migration to areas like Appalachia can result in rural gentrification. People who are currently living in low-cost, natural resource–rich areas are forced to leave, resulting in the loss of access to the resources and communities they have actively contributed to developing and maintaining. As a result, they can no longer enjoy the benefits of these regions.

Describe **one** way you could revise or build on the computational model to better understand how migration into Appalachia could impact current residents' access to natural resources. Be sure to explain how the change you describe will provide a better understanding of how migration into Appalachia could impact current residents' access to natural resources.

Exhibit 3.13. Human Migration to Appalachia Item 2 Constructed Response Scoring Notes

Emphasis here is on an understanding of how to revise the model parameters to better understand more nuanced population-environment dynamics. Student ideas can be wide-reaching but should be justifiable as at least one of the following: (a) updating the model to better understand who moves into and out of the region, and/or (b) determining the feedback impacts of population growth on biodiversity, and natural resources such as water. This item specifically focuses on the computational reasoning aspect of this SEP and can include either quantitative or qualitative reasoning from students. Appropriate lines of reasoning here can include, but are not limited to, the following:

- updating the model to account for demographic subgroups
- considering housing costs/other metrics for socioeconomic status as part of the inputs and/or outputs, as a mechanism to better understand the relative wealth/characteristics of who lives in the area.
- calculations of migration out of Appalachian regions
- feedback mechanisms that influence biodiversity and natural resource availability (Note: This could be specific to resources, or general at the level of the categories included in the model.)
- relative factors for scaling variables (quantitative or qualitative)

3F. Assessing the Full Range of Student Performance

It is important that the NAEP Science Assessment provide a complete picture of student performance. Although there have been concerns that creating an assessment consisting largely of multidimensional items, item sets, and scenario-based items might prove too difficult for students who have not been provided the opportunity to develop proficiency in science, research from the learning sciences, including research on how students learn and develop three-dimensional science understanding, suggests otherwise (NRC, 2005, pp. 407-411; NRC, 2007; NASEM, 2017, pp. 5-14; NASEM, 2018, pp. 145-146). While traditional approaches to assessment often assume that rote understanding or simple procedural skills (e.g., definitions, facts, lab skills) are less complex and therefore more likely to be doable by students who are still developing their science understanding, this is not borne out in practice. Students do not learn by mastering one dimension at a time before integrating the dimensions, nor by memorizing content before applying it—they learn by using the dimensions together in increasingly sophisticated ways. For the purpose of this framework, increasing sophistication refers to a student expression of understanding that is more thorough, more precise, more accurate, and more coherent throughout. Likewise, assessments intended to surface what students who have not yet mastered grade-level expectations know and are able to do may do so more effectively by varying the complexity of multidimensional performances, rather than focusing on one-dimensional items.

Students at all grade levels and all performance levels can and do find success with multidimensional performances if students are presented with items that (a) use appropriately complex contexts, (b) sufficiently scaffold and support learners in engaging with the item, and (c) use the dimensions in appropriate combinations to right size the complexity. These considerations are particularly important for multilingual learners and other students who may have conceptual understanding without having yet mastered vocabulary or rote facts and procedures. By focusing on multidimensional items that range in complexity, NAEP can better capture student thinking along progressions that mirror how student thinking develops.

The complexity framework that will be applied to NAEP item development will reflect how complexity specifically scales within and across multidimensional science items, including:

- the nature of the phenomenon or problem context;
- the sophistication of language, graphics, or mathematical elements, and their presentation together as appropriate;
- the complexity of the item stem, response mode, and response choices;
- the extent of sensemaking that is required of the student;
- the degree and nature of scaffolding and guidance provided; and
- the nature of the intersections of dimensions within items, including how each dimension contributes to the complexity of sensemaking in the item.

Complexity Framework

The proposed 2028 NAEP Science Framework is informed by the item complexity frameworks proposed by Achieve (2019a); Tekkumru-Kisa, Stein, and Schunn (2015); and WestEd, Center on Standards and Assessment Implementation, and Delaware Department of Education (2019). The purpose of the complexity framework is to inform item development so as to ensure that items are accessible to a wide range of learners. The complexity framework considers two underlying contributors to complexity:

- The degree and nature of guidance provided to students. That is, how much direction or cueing are students given for what to consider and how to approach the item?
- The degree and nature of sensemaking required by students. That is, how sophisticated is the sensemaking required by students, and how does each dimension contribute to that sophistication in each item?

The complexity framework intentionally goes deeper than some traditional approaches to complexity (e.g., cognitive demand or content complexity approaches, such as Webb's Depth of Knowledge). By considering not only the overall complexity of each item, but also how each dimension contributes to sensemaking, items can be designed more intentionally. For example, some items provide substantial scaffolding for engaging in the NAEP SEP, with limited cueing for the NAEP DC, while other items engage in a lower level of sensemaking with NAEP DCs while providing students the opportunity to engage with the NAEP SEP and CCC more deeply. In some items, the NAEP CCC can be used to reduce item complexity (e.g., by asking students to identify a pattern as a step toward figuring out the phenomenon) while in other items, the NAEP CCC expands complexity by asking students to consider a nonroutine lens on a phenomenon or problem (e.g., asking students to examine a seemingly causal relationship that is correlational). These are important considerations for developing a balanced assessment that can intentionally surface a range of student thinking.

Exhibit 3.14. Complexity of Multidimensional Items

	How does the NAEP DC contribute to the sophistication of sensemaking?	**How does the NAEP SEP contribute to the sophistication of sensemaking?**	**How does the NAEP CCC contribute to the sophistication of sensemaking?**	**Overall**
High	Students are given limited prompting about which DC to use. Students may leverage ideas from multiple DCs that are not closely related (within or	Students are given limited prompting about which SEP to use, and how to engage in it. Students may use a series of SEP	Students make decisions about which CCC to use to organize their approach to / reasoning within an item.	Two or three dimensions are used to engage in a high degree of sensemaking. Students are given limited prompting about how to

	How does the NAEP DC contribute to the sophistication of sensemaking?	How does the NAEP SEP contribute to the sophistication of sensemaking?	How does the NAEP CCC contribute to the sophistication of sensemaking?	Overall
High (cont.)	across multiple disciplines). Students use DCs to address a significant uncertainty, with many possible alternative accounts.	elements in a sequence of sophisticated thinking that expands the nature of sensemaking. Students use SEPs to navigate complex interactions among multiple components of phenomena and problems.	Students explicitly use the CCCs to expand sensemaking. With limited prompting, students use CCCs to navigate phenomena and problems with significant uncertainty and many possible alternative accounts.	approach the item, requiring them to decide what understandings and practices to apply. Students address a high degree of authentic uncertainty in the phenomenon or problem, navigating many possible (and valid) accounts.
Medium	Students are cued to use specific DCs to address the item. Students may leverage multiple components of a given DC together or demonstrate a sophisticated use of a single DC component. Students use DCs to address a moderate uncertainty, with limited alternative accounts.	Students are cued to use specific SEPs and components of SEPs to address the item. Students use a single SEP component in support of authentic sensemaking. Students use SEPs to navigate simple interactions among components of phenomena and problems.	Students are cued to use a specific CCC component. CCCs serve to focus student thinking within the item. With guidance, students use the CCCs to navigate simple interactions among components of phenomena and problems with moderate uncertainty.	Students are provided substantial cues for addressing the phenomenon or problem. They are prompted with specific DCs, SEPs, and CCCs, and provided guidance on how to use them. One dimension may be more heavily cued than others. Students address a moderate degree of uncertainty with limited possible accounts.
Low	Students are directed to use specific components of a DC to address the item.	Students are directed to use specific components of the SEP, using a well-defined set of actions or procedures.	Students use given CCCs in service of lower-level sensemaking, addressing phenomena and problems with	Students use a well-defined set of actions to engage in the item and address the phenomenon or problem. They

	How does the NAEP DC contribute to the sophistication of sensemaking?	How does the NAEP SEP contribute to the sophistication of sensemaking?	How does the NAEP CCC contribute to the sophistication of sensemaking?	Overall
Low (cont.)	Students use limited DC components in routine or highly specific ways. Students engage in a simple application of the DC component to a phenomenon with a low degree of uncertainty.	Students use the SEP as structure to make an idea visible, without using the SEP in service of significant sensemaking.	limited uncertainty and limited alternative accounts.	engage in applications of DCs, SEPS, and CCCs, often involving one or two scaffolded steps. Students address a low degree of uncertainty with a single possible account.

The following three illustrative examples demonstrate ways that items may vary in complexity, including a low-complexity item, a series of different versions of an item with different complexities, and an item set with items that vary in complexity.

Illustrative Examples for Varying Levels of Complexity

- **Park Flooding** presents a low-complexity item.
- **Permafrost** presents ways to vary complexity across different versions of an item.
- **Human Migration to Appalachia** presents ways to vary complexity across an item set that is medium complexity overall.

For example, Exhibit 3.15 illustrates a low-complexity grade 4 item assessing Earth and Space Sciences with low DC complexity and low SEP complexity. The phenomenon is presented through simple text, an image, and a simple graph—this provides students with enough information to demonstrate the targeted NAEP DC and SEP in service of sensemaking, without unnecessary reading or cognitive load.

Exhibit 3.15. Park Flooding, Version 1

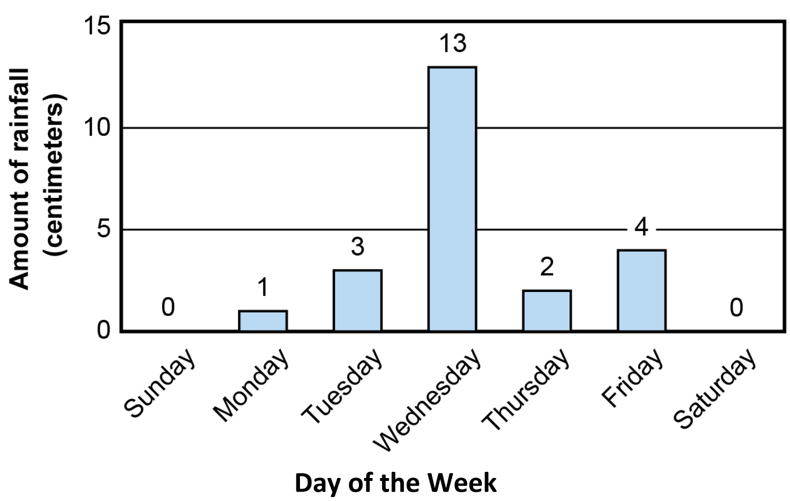
Item ID: Park Flooding (adapted from Formative Assessment Bundling Literacy and Elementary Science)

People visiting a local park noticed that the park was flooded and was closed for the day. The picture shows the flooded park.

Flooded Park[7]

The park was closed only on the day the flooding happened. The bar graph shows the rainfall for each day of that week.

Amount of Rainfall at the Park



Part A

Based on the data, on which day was the park **most likely** flooded?

A. Sunday
B. Monday
C. Tuesday
D. Wednesday

Key: D

[7] Betty Longbottom / Flooded Playground! - Cliffe Avenue / CC BY-SA 2.0

Exhibit 3.16 illustrates a low complexity grade 12 item with low DC complexity, low SEP complexity, and low CCC complexity. By comparison to Exhibit 3.15, this shows how adding a dimension does not necessarily increase the complexity. While students cannot respond to this item without bringing some understanding of the DC, SEP, and CCC, it is heavily supported, thus limiting the amount of sensemaking students engage in with any dimensions. Students are given the relationship they need to map, only relevant statements to move, and a very structured and nonquantitative model illustrating feedback loops. It should be noted that this item leans into the intentional progressions built into the dimensions: the SEP and CCC are appropriate for grade 12 at a low level, and they are related to expectations and performances at lower grade levels. This allows a way to reduce the complexity of the item, remain consistent with the grade-level targets, and account for the fact that lower-performing students may have a less sophisticated, less well-developed understanding of the targeted dimensions that more closely approximates sophisticated performances at lower grade bands.

Exhibit 3.16. Permafrost Melting, Version 1

Item ID: Permafrost Melting (adapted from OpenSciEd)

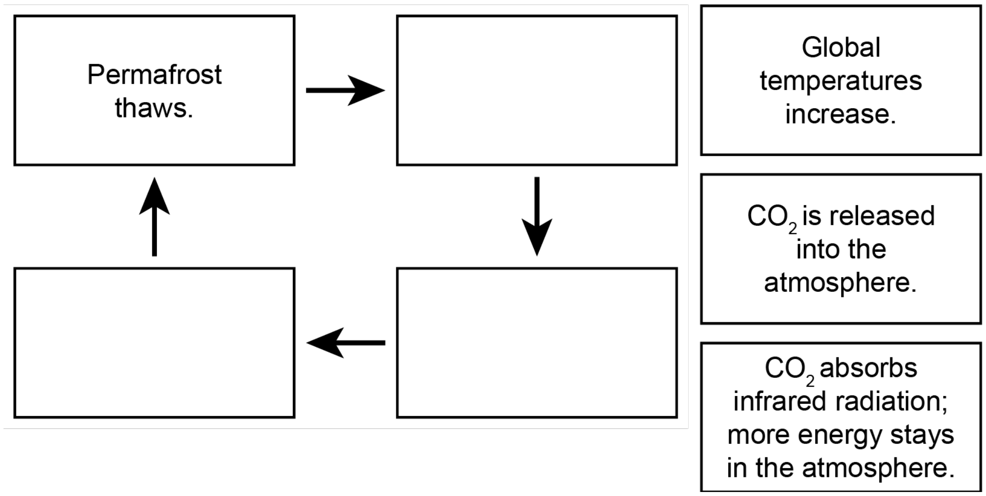
Permafrost is a layer of soil and ice that is just below the surface in the Arctic. An example of permafrost is shown in the picture.

Historically, permafrost stayed frozen for many years at a time. However, in some areas, permafrost is now melting, which can cause many changes to Earth's surface and living things. One major concern is that permafrost contains carbon dioxide.

Permafrost[8]

[8] Permafrost from OpenSciEd, CC BY-SA 4.0

Use the statements to develop a model that shows the relationships between thawing permafrost and rising global temperatures. Drag the statements into the correct boxes to complete the model. Each statement will be used once.



The following two versions of the item are modified to be higher complexity by (a) requiring students to develop the model with significantly less support (Exhibit 3.17), and/or (b) asking students to consider implication and limitations of the model (Exhibit 3.18). By making these modifications, students are more independently responsible for figuring out why permafrost melting contributes to rising global temperatures by using their understanding of modeling, feedback loops, and energy in systems. In these variations, the complexity has been increased, but there is still substantial cueing in the form of the provided statements. A more complex item might ask students to develop an original model, without any prompting statements.

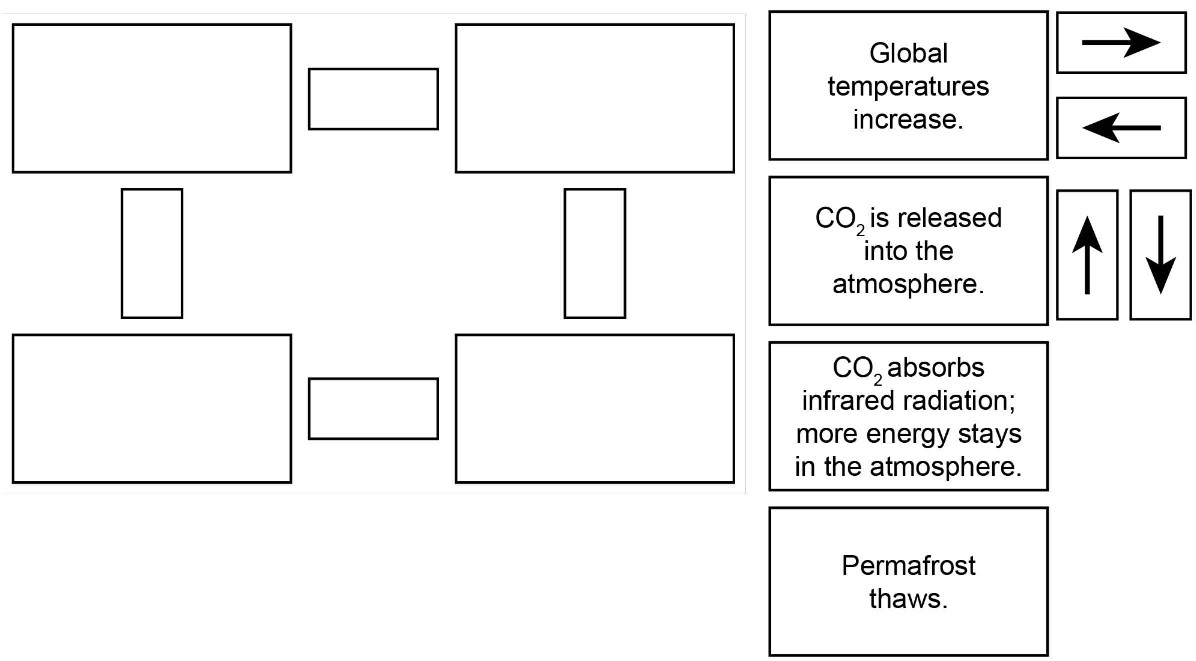
Exhibit 3.17. Permafrost Melting, Version 2

Permafrost is a layer of soil and ice that is just below the surface in the Arctic. An example of permafrost is shown in the picture.

Historically, permafrost stayed frozen for many years at a time. However, in some areas, permafrost is now melting, which can cause many changes to Earth's surface and living things. One major concern is that permafrost contains carbon dioxide.

Permafrost[9]

Use the statements and arrows to develop a model that shows the relationships between thawing permafrost and rising global temperatures. Drag the statements and arrows into the correct boxes to develop the model. Each statement and each arrow will be used once.



[9] Permafrost from OpenSciEd, CC BY-SA 4.0

Exhibit 3.18. Permafrost Melting, Version 3

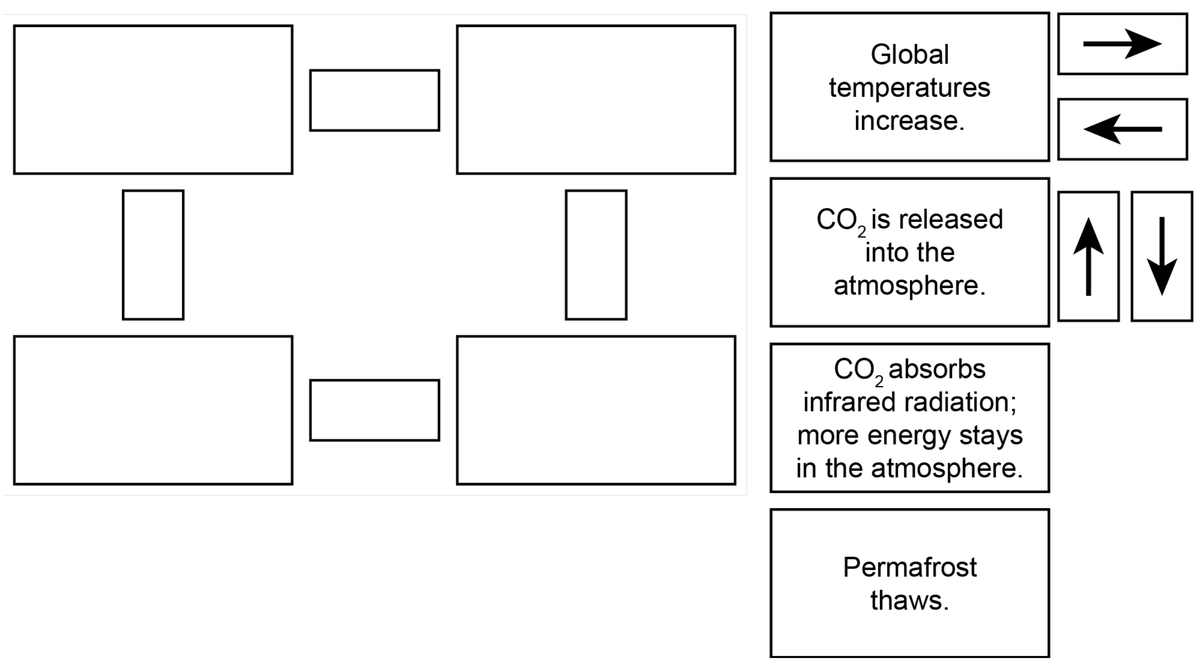
Permafrost is a layer of soil and ice that is just below the surface in the Arctic. An example of permafrost is shown in the picture.

Historically, permafrost stayed frozen for many years at a time. However, in some areas, permafrost is now melting, which can cause many changes to Earth's surface and living things. One major concern is that permafrost contains carbon dioxide.

Permafrost[10]

Part A

Use the statements to develop a model that shows the relationships between thawing permafrost and rising global temperatures. Drag the statements and arrows into the correct boxes to complete the model. Each statement and arrow will be used once.



[10] Permafrost from OpenSciEd, CC BY-SA 4.0

Part B

Complete the sentence by choosing the correct answer from the drop-down menu.

The model predicts that the rate of melting of the permafrost will likely (increase / stay the same / decrease) over the next 50 years.

Key: increase

Part C

Describe a limitation of the model you developed in Part A.

Exhibit 3.19. Permafrost Melting Part C Constructed Response Scoring Notes

- Students provide one limitation of the model, for example:
  - The model is missing factors such as interactions with other systems and feedback loops.
  - Thawing may occur at different rates in different locations around the world.
  - There may be a time lag between temperature increases and thawing.

Exhibit 3.20 presents ways to vary complexity across an item set that is medium complexity overall.

Exhibit 3.20. Human Migration to Appalachia

Item ID: Human Migration to Appalachia

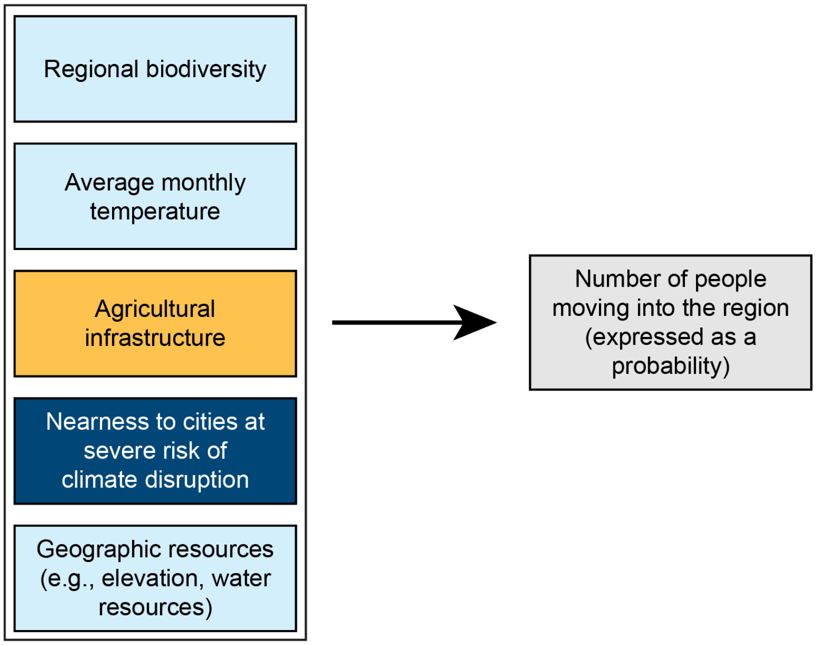
Appalachia is considered “climate resilient.” This means that the area can successfully handle the impacts of changes to climate and can prevent those impacts from growing worse. The green areas in Figure 1 show where Appalachia is located in the United States.

Figure 1. Map of Appalachia

Computational models predict that many people will move into the Appalachian region over the next 20 years as they seek to find places to live that are safer and more stable.

Figure 2 shows one model local leaders are using to predict how many people will move into Appalachia. Blue indicates factors that are expected to increase migration, and orange indicates factors expected to decrease migration. Darker colors indicate more weight on that factor in the model. Agricultural infrastructure includes farms, markets and businesses that support farms, and the transportation and communication systems in the area.

Figure 2. Computational Model for Predicting Human Migration



When this model was tested against recent population growth due to migration into two locations in Appalachia, leaders noticed some differences between what the model predicted and what the data showed. The table shows these differences for the two locations, along with information about how high or low each location is rated on several factors.

Predicted and Actual Population Growth in Pittsburgh and the Shenandoah Valley

Location	Pittsburgh	Shenandoah Valley
Predicted population growth	high	low
Actual population growth	low	high
Relative biodiversity	low	high
Average monthly temperature range	29–73° F	32–74° F
Relative agricultural infrastructure	low	medium
Nearness to climate-impacted urban centers	high	high
Access to usable water	medium	high

Item 1

Based on your understanding of what human societies need to be successful, which idea **best** explains why the model prediction was different from the observed data?

A. Pittsburgh has highly limited access to water. The model did not account for how this shortage would limit the growth of local businesses.
B. Human societies require reasonable annual temperature ranges. The model incorrectly assumed that average monthly temperature was a less important factor.
C. Human societies require access to food and water. The model did not account for how important access to food and water would be for human migration to Appalachia.
D. The Shenandoah Valley has very little access to fresh fruits and vegetables. The model incorrectly assumed that agricultural infrastructure would not promote migration.

Key: C

Item 2

Human migration to areas like Appalachia can result in rural gentrification. People who are currently living in low-cost, natural resource–rich areas are forced to leave, resulting in the loss of access to the resources and communities they have actively contributed to developing and maintaining. As a result, they can no longer enjoy the benefits of these regions.

Describe **one** way you could revise or build on the computational model to better understand how migration into Appalachia could impact current residents' access to natural resources. Be sure to explain how the change you describe will provide a better understanding of how migration into Appalachia could impact current residents' access to natural resources.

Exhibit 3.21. Human Migration to Appalachia Item 2 Constructed Response Scoring Notes

Emphasis here is on an understanding of how to revise the model parameters to better understand more nuanced population-environment dynamics. Student ideas can be wide-reaching but should be justifiable as at least one of the following: (a) updating the model to better understand who moves into and out of the region, and/or (b) determining the feedback impacts of population growth on biodiversity, and natural resources such as water. This item specifically focuses on the computational reasoning aspect of this SEP and can include either quantitative or qualitative reasoning from students. Appropriate lines of reasoning here can include, but are not limited to, the following:

- updating the model to account for demographic subgroups
- considering housing costs/other metrics for socioeconomic status as part of the inputs and/or outputs, as a mechanism to better understand the relative wealth/characteristics of who lives in the area
- calculations of migration out of Appalachian regions
- feedback mechanisms that influence biodiversity and natural resource availability (Note: This could be specific to resources, or general at the level of the categories included in the model.)
- relative factors for scaling variables (quantitative or qualitative)

The NAEP Science Assessment and Item Specifications include additional examples of items at a range of complexity levels, including how a given item can be modified.

Complexity, Opportunity to Learn, and Surfacing the Full Range of Student Performance

One major goal of the complexity framework is to provide a way to vary the nature of items so that all students who have had the opportunity to learn some science can make their understanding visible on the NAEP Science Assessment. The complexity framework assumes that every item still needs to require grade-level NAEP DCs and SEPs. What varies are the degree and sophistication of the sensemaking required to respond to the item, which we expect will scale with increasing proficiency (e.g., students who have a robust grasp of science will be able to successfully engage with increasingly complex items, across a range of NAEP DCs, SEPs, and CCCs). Some students may not demonstrate robust science understanding simply because they did not have sufficient opportunity to learn the science being assessed by NAEP. The complexity framework attends to this to some degree by leveraging the intentional NAEP DC, SEP, and CCC progressions across grades 4, 8, and 12. The grade 8 NAEP SEP expectation may reflect a lower level of sophistication of a similar target at grade 12, so that an item that represents performance at the *NAEP Advanced* level for grade 8 may be similar to an item that represents performance at the *NAEP Basic* level for grade 12. If a student has not had the opportunity to develop their understanding beyond the grade 8 expectation, some information about their science understanding can still be captured. However, if students are performing at the lower end of the scale because they have not had opportunity to learn the targeted science, even low-complexity items may not surface their understanding.

3G. Reflecting a Wide Range of Students

It is essential for the 2028 NAEP Science Assessment to be responsive and relevant to a wide diversity of students. Specifically, students taking the assessment should see themselves and their communities represented in the items across the assessment as a whole, and the range of assets and "funds of knowledge" diverse learners bring to the table should be acknowledged as important elements of science achievement. Funds of knowledge are the historical accumulation of abilities, bodies of knowledge, assets, and cultural ways of interacting that a student might possess. Below are definitions and general principles for culturally relevant contexts for NAEP science followed by a list of particular features of these contexts.

General Principles and Definitions

- All students have culture, and when we think about diverse cultural representation, we mean to be inclusive of cultural and linguistic experiences across a range of geographies, cultural practices, disabilities, languages, and gender.
- Although some phenomena will be more relevant to some students than others, all students should be able to see themselves and their peers represented in some

phenomena/problems included across the assessment. This framework does not suggest that every student be matched with particular items, but rather that all learners should see a range of phenomena, geographies, and people represented so that the assessment is culturally relevant.

- By varying the range of who the phenomena/problems are relevant to, we ensure that there is authentic relevance to multiple student groups.
- When contexts focus on legitimate interests of communities, it is more likely that all students will be engaged with the items. A culturally relevant lens asks whether the item elicits a productive affective response.
- Providing sufficient background information, including multiple modalities for conveying contexts and any additional information about a context for a phenomenon that is needed, will help reduce inadvertent issues of bias by ensuring that all students have an opportunity to become familiar with a context. It should be noted that this kind of appropriate background information is essential in all items and can help ensure that student performance on the assessment is a trustworthy indicator of what they know and are able to do, not whether they were able to understand the task or were motivated to complete it.

Specific Features of Culturally Relevant Contexts and Assessment Design

- Item contexts consider geographic, demographic, and time-related factors to create enough distance between groups of students intended to be taking the assessment and the phenomenon to limit any negative affective responses.
- Contexts include diverse representations of who is considered a scientist and/or engineer.
- Contexts position people of color as (a) more than a stereotyped experience and (b) powerful doers and contributors to science and the broader world.
- Contexts do not include (or limit) gratuitous or superficial representation of diverse races, ethnicities, and so on.

The following two illustrative examples provide culturally relevant contexts and items. Each example draws upon different cultures and provides access to cultural information in different ways. In both cases, the culturally relevant information is necessary information for students to respond to the items.

Illustrative Examples

- **Limu Kohu**
- **Human Migration to Appalachia**

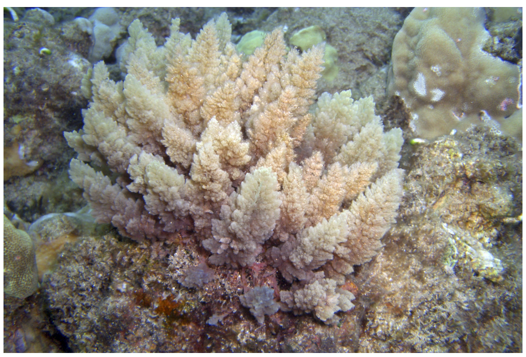
The first of these two examples, Exhibit 3.22, is an example of a two-dimensional grade 8 item. Several features of culturally responsive items are included, such as (a) the use of native/home language in the item (i.e., limu kohu is the Hawaiian language term for this seaweed species), (b) use of non-traditional evidence sources that have been useful in university-based science endeavors (i.e., the use of multi-generational/elder accounts as evidence, as used by Stanford botanist Dr. Isabella Aiona Abbott), and (c) explicitly addressing a problem that is meaningful to specific communities (loss of limu kohu is very important to Hawaiian communities, and is representative of a broader conversation about the loss of indigenous foodways currently happening).

Exhibit 3.22. Limu Kohu

Item ID: Limu Kohu (adapted from the State Performance Assessment Learning Community)

Limu kohu is a type of seaweed that is native to the waters around Honolulu, Hawaiʻi. It is an important part of food systems as well as cultural and religious practices. Although limu kohu was easy to find for hundreds of years, limu kohu populations around Honolulu have been rapidly declining over the past 60 years. An example of limu kohu seaweed is shown in the picture. The table describes observations of limu kohu.

Limu Kohu[11]



Observations from Generations of Hawaiian Elders about Limu Kohu Growth and Harvesting
• Limu kohu needs warm water and high salinity to grow. • Limu kohu grows and reproduces well on the edges of coral reefs. • When limu kohu is trimmed, it regrows. • When the base of the limu kohu is harvested, it cannot regrow.

Part A

Which human activity is **least** likely to cause harm to limu kohu populations?

A. Companies using industrial methods of harvesting limu kohu remove the whole limu kohu plant.
B. Restaurants using traditional methods of harvesting limu kohu remove the top of the limu kohu.
C. Industrial runoff changes the temperature and salinity of the water in coastal regions where limu kohu live.
D. Ships visiting Hawaiʻi introduce invasive seaweed species that use the same resources as limu kohu into coastal regions where limu kohu live.

Key: B

Part B

Use the information provided and your understanding of the impacts of human activities on the environment to support your answer to Part A.

[11] MDC Seamarc Maldives, CC BY-SA 4.0, via Wikimedia Commons

Exhibit 3.23. Limu Kohu Part B Constructed Response Scoring Notes

Students provide one statement that is based on the information provided, shows understanding of the impacts of human activities on the environment, and supports the answer to Part A. For example:

- Traditional harvesting practices that focus on harvesting only the top portions of limu kohu are less destructive compared to other methods that involve uprooting the entire plant. When only the top is harvested, the base of the plant remains intact, allowing the limu kohu to potentially regrow.
- The information provided indicates that limu kohu has the ability to regrow when it is trimmed. This suggests that harvesting the upper parts of the seaweed allows it to regenerate, ensuring the sustainability of the population over time.

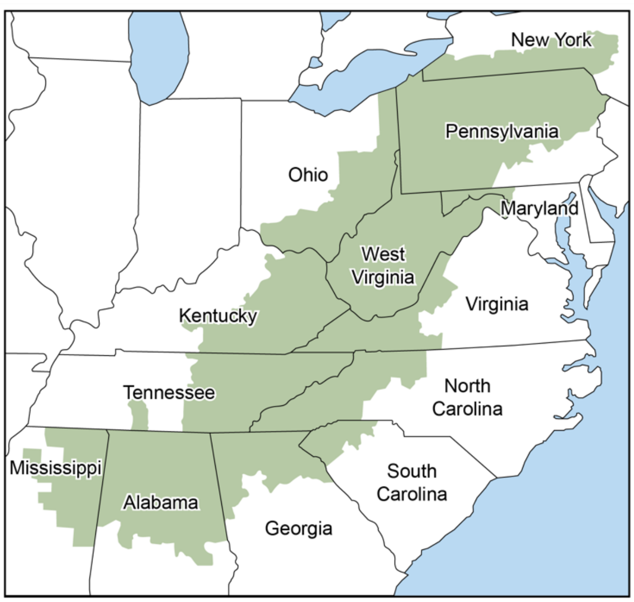
The second example, Exhibit 3.24, highlights one way to attend to cultural relevance by highlighting a specific community (in this case, the many largely rural communities that make up Appalachia) through an asset-based lens. This is particularly important because these specific rural communities are often portrayed and understood through deficit-oriented lenses (e.g., focusing on poverty, lack of educational resources and college degrees, economically less advantageous career options, etc.).

Exhibit 3.24. Human Migration to Appalachia

Item ID: Human Migration to Appalachia

Appalachia is considered “climate resilient.” This means that the area can successfully handle the impacts of changes to climate and can prevent those impacts from growing worse. The green areas in Figure 1 show where Appalachia is located in the United States.

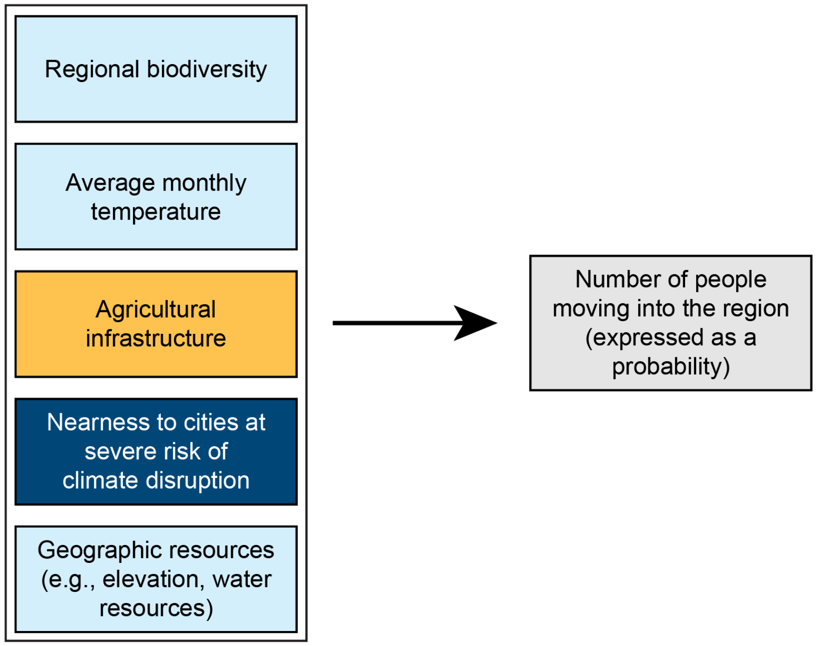
Figure 1. Map of Appalachia



Computational models predict that many people will move into the Appalachian region over the next 20 years as they seek to find places to live that are safer and more stable.

Figure 2 shows one model local leaders are using to predict how many people will move into Appalachia. Blue indicates factors that are expected to increase migration, and orange indicates factors expected to decrease migration. Darker colors indicate more weight on that factor in the model. Agricultural infrastructure includes farms, markets and businesses that support farms, and the transportation and communication systems in the area.

Figure 2. Computational Model for Predicting Human Migration



When this model was tested against recent population growth due to migration into two locations in Appalachia, leaders noticed some differences between what the model predicted and what the data showed. The table shows these differences for the two locations, along with information about how high or low each location is rated on several factors.

Predicted and Actual Population Growth in Pittsburgh and the Shenandoah Valley

Location	Pittsburgh	Shenandoah Valley
Predicted population growth	high	low
Actual population growth	low	high
Relative biodiversity	low	high
Average monthly temperature range	29–73° F	32–74° F
Relative agricultural infrastructure	low	medium
Nearness to climate-impacted urban centers	high	high
Access to usable water	medium	high

Item 1

Based on your understanding of what human societies need to be successful, which idea **best** explains why the model prediction was different from the observed data?

A. Pittsburgh has highly limited access to water. The model did not account for how this shortage would limit the growth of local businesses.
B. Human societies require reasonable annual temperature ranges. The model incorrectly assumed that average monthly temperature was a less important factor.
C. Human societies require access to food and water. The model did not account for how important access to food and water would be for human migration to Appalachia.
D. The Shenandoah Valley has very little access to fresh fruits and vegetables. The model incorrectly assumed that agricultural infrastructure would not promote migration.

Key: C

Item 2

Human migration to areas such as Appalachia can result in rural gentrification. People who are currently living in low-cost, natural resource–rich areas are forced to leave, resulting in the loss of access to the resources and communities they have actively contributed to developing and maintaining. As a result, they can no longer enjoy the benefits of these regions.

Describe **one** way you could revise or build on the computational model to better understand how migration into Appalachia could impact current residents' access to natural resources. Be sure to explain how the change you describe will provide a better understanding of how migration into Appalachia could impact current residents' access to natural resources.

For additional examples, please see Appendix A of the NAEP Science Assessment and Item Specifications.

3H. Science Achievement Expectations

Although each student will answer only a subset of items, the full NAEP Science Assessment will measure student sensemaking in each of the NAEP DCs in Chapter 2. The following guidance is provided to support item development. It is not intended to be prescriptive or limiting to item development.

An essential part of the item development process is to write a multidimensional science achievement expectation expressed as a *performance*—something that the student can be expected to do to indicate they understand the targeted NAEP DC and can apply it via the cued associated NAEP SEP (and CCC, when possible). Following are examples that can be used to build items for grades 4, 8, and 12. Additional guidance for creating science achievement expectations is provided in the NAEP Science Assessment and Item Specifications.

Exhibit 3.25. Examples of Science Achievement Expectations

Science Achievement Expectation	NAEP DC	NAEP CCC	NAEP SEP	Rationale
Grade 4 Earth and Space Sciences: Interpret patterns in sunrise/sunset data for a given location to explain seasonal differences in day length.	E4.1: Many objects in the sky change position and are not always visible due to Earth's rotation. The patterns of motion of the sun and moon can be observed, measured, described, and predicted.	C4.1: Similarities and differences in patterns can be used to sort, classify, communicate, predict, and explain, with various representations (such as physical graphs or diagrams) to describe and analyze features of simple natural phenomena and designed products.	S4.10: Analyze data to determine whether it supports or refutes a claim about a phenomenon or design solution.	One of the first age appropriate CCCs for younger students to engage with is patterns. Sunrise/sunset times have seasonal patterns to them that are caused by motion in the sun/Earth system over the course of a year. This smaller idea (day length) is an important component to many larger ideas (seasonal temperature differences, light/temperature cues for plant life cycles, etc.).

Science Achievement Expectation	NAEP DC	NAEP CCC	NAEP SEP	Rationale
Grade 8 Physical Science: Ask questions about the interactions between systems of objects to determine how changes in their motions are determined by the sum of the forces acting on each object.	P8.5: The change in motion of an object is determined by the sum of the forces acting on it; if the net force on the object is zero, it will remain at rest or continue moving in a straight line with the same speed and direction as before.	C8.8: Systems may interact with other systems; they may have subsystems and be a part of larger, more complex systems.	S8.2: Ask questions that can be answered with empirical evidence to investigate relationships between variables in a system model or in phenomena.	Students can begin to ask questions to develop a qualitative understanding of forces at entry points to making sense of phenomena related to interactions between objects. The sophistication of their questions grows as students progress toward mastery of complex material, providing opportunities to write items at all levels of difficulty and complexity.

Science Achievement Expectation	NAEP DC	NAEP CCC	NAEP SEP	Rationale
Grade 12 Life Science: Examine data on different types of grass that can be used in a design for a new public park. Take into account several factors when deciding on the type of grass that will have the smallest negative effect on the environment.	L12.12: Changes induced by human activity (anthropogenic change) in the environment—such as habitat destruction, pollution, introduction of invasive species, overexploitation, and climate change—can disrupt an ecosystem and threaten the survival of some species.	C12.14: Changes in a system can be caused by changes in other systems or in conditions affecting the system as well as by prior changes within the system. The scale of the effect is not always comparable to that of the change but may be much larger or smaller.	S12.20: Evaluate and/or refine a solution for a design problem based on scientific knowledge, evidence, prioritized criteria, and trade-off considerations.	By grade 12, students are able to prioritize criteria and take into account information from several sources to decide how to solve an engineering problem in a way that minimizes the disruption of an ecosystem.

3I. Digital Tools

The NAEP Science Assessment based on this framework will be administered via computer. A digital environment provides opportunities to include a number of digital tools—and, at times, science-specific tools—students can use to respond to the items. The 2028 NAEP Science Assessment will include digital tools to support NAEP DCs, SEPs, and CCCs.

The testing environment will need to provide a computational tool equivalent to a four-function calculator. Continuing a practice that has been in place for recent NAEP administrations, before the assessment, students complete a brief interactive tutorial designed to orient them to the digital tools they will use during the assessment. The tutorials for each grade level can be found on the National Center for Education Statistics website.

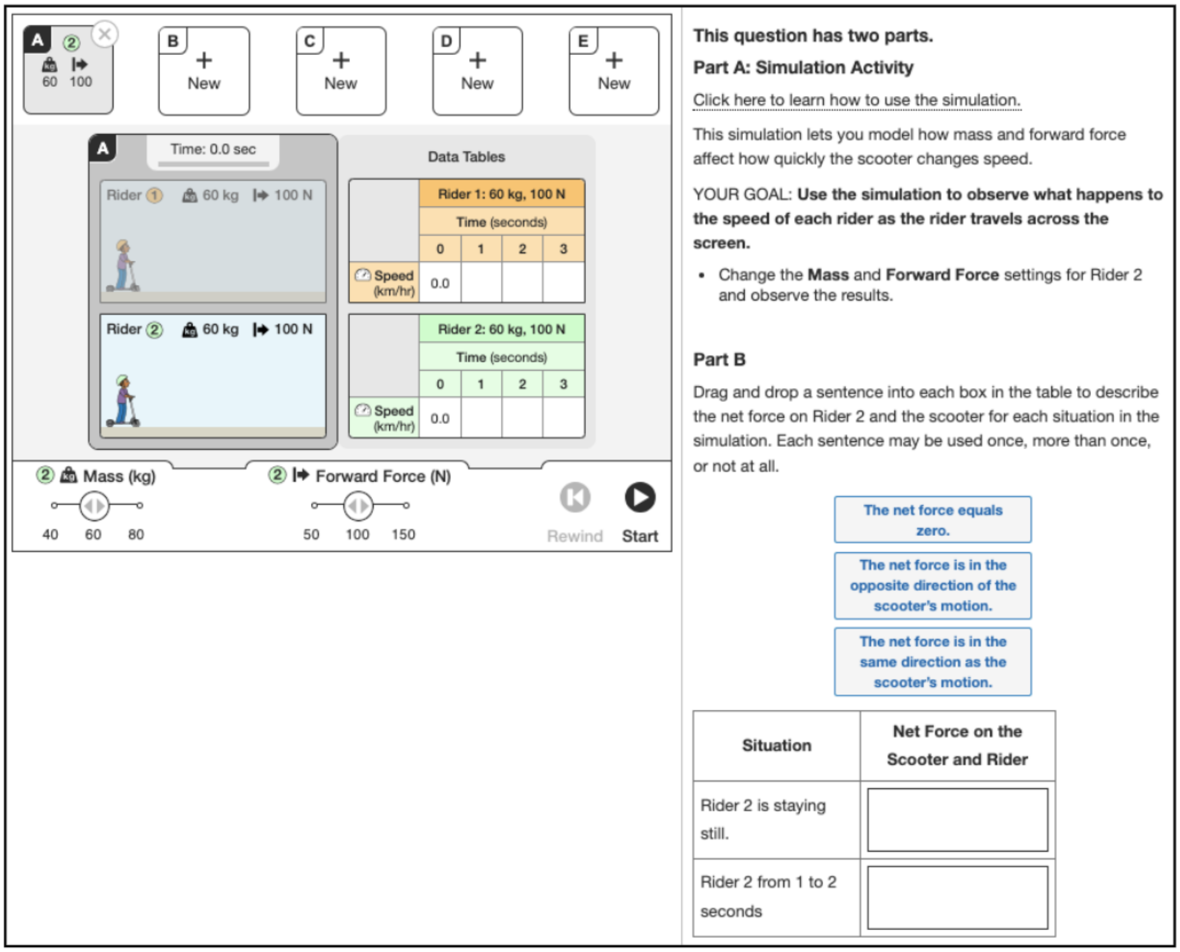
All digital NAEP assessments include system tools, which are always available and common across all NAEP assessments. There are also science and mathematics tools, which are specific to and available only for certain items on NAEP science assessments. The materials and accompanying items are carefully chosen to cause minimal disruption of the administration process and are typically provided only when relevant to solving the item.

The illustrations in this framework are static screenshots to illustrate examples of these digital tools; however, the screenshots represent only a small subset of the many images, videos, and simulations students encounter during the assessment. Digital tools should be used when the item format offers advantages over other assessment modes. Examples include (but are not limited to) testing student scientific sensemaking related to the following situations:

- using simulations and modeling tools for scientific phenomena that cannot easily be observed in real time, such as seeing things in slow motion (e.g., the motion of a wave) or at a higher speed (e.g., erosion caused by a river)
- modeling scientific phenomena that are invisible to the naked eye (e.g., the movement of molecules in a gas)
- working safely in lab simulations to collect and analyze data that would otherwise be disorderly in an assessment situation or hazardous (e.g., using dangerous chemicals)
- situations that require several repetitions of an experiment while the student varies the parameters (e.g., rolling a ball down a slope while varying the mass, the angle of inclination, or the coefficient of friction of the surface)

The following example (Exhibit 3.26) highlights how simulations might be used within NAEP Science. In this example, students use the simulation to better understand the forces acting within a context. Please note that this simulation is intended to illustrate the use of a simulation within a set of items. These examples were taken from a larger item set or scenario-based task and, as standalone simulations, do not feature the phenomenon described earlier in the task. This item therefore is not intended to be an example of a possible NAEP item.

Exhibit 3.26. Sample Simulation Taken from a Multidimensional Item Set[12]



Similarly, the example below (Exhibit 3.27) shows how digital tools might be used to allow students to construct dynamic models that they independently develop. This example leverages SageModeler, a free, open-source, web-based systems dynamics modeling tool commonly used in science education. This tool allows students to define variables, relationships, and degree of influence and to run models and collect data.

The exhibit shows a model of the interactions between two populations and the resulting changes to the populations over time. This example could be used with items that assess ecosystems NAEP DCs; analyzing, explaining, and evaluating SEPs; and relevant CCCs, including but not limited to Patterns, Cause and Effect: Mechanism and Explanation, Systems and System Models, Energy and Matter: Flows, Cycles, and Conservation, and Stability and Change. The complexity of the model could be tailored to the targeted grade and complexity.

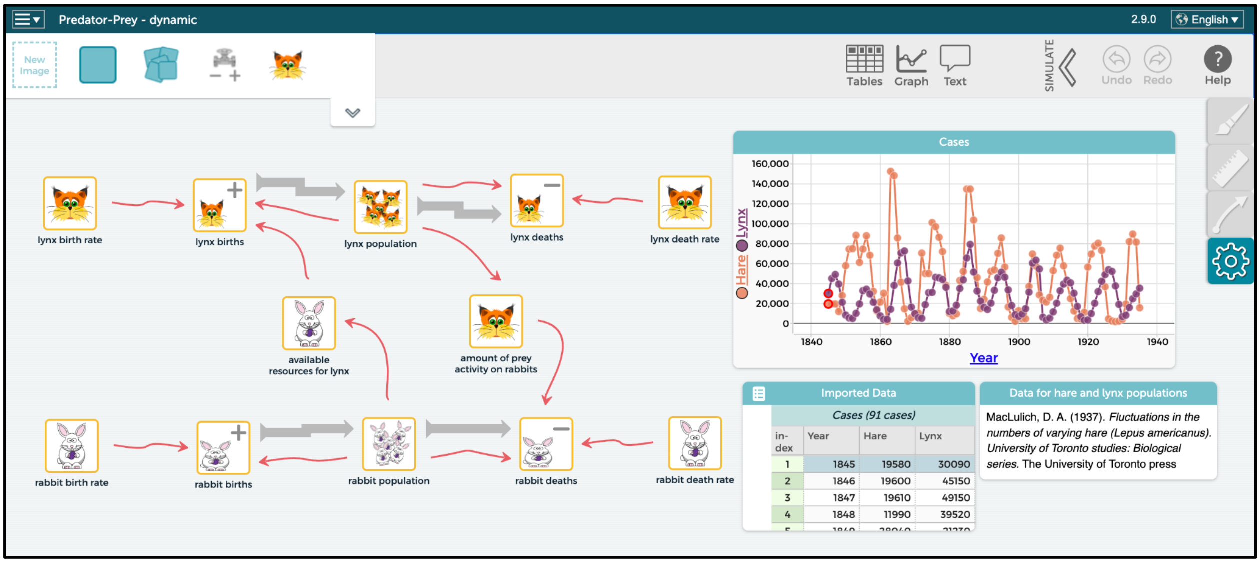
A modeling tool such as this could be used to model any system in which understanding how the components interact and how those interactions lead to changes in the overall system is

---

12 Image used with permission from Massachusetts Department of Elementary and Secondary Education

appropriate for assessing the NAEP DCs, SEPs, and CCCs. Please note that this item is not intended to be an example of a possible NAEP item.

Exhibit 3.27. Sample Modeling Tool (SageModeler)[13]



13 Image generated using SageModeler (https://sagemodeler.concord.org/), which was developed at the Concord Consortium and Michigan State University.

CHAPTER FOUR: Reporting Results of the NAEP Science Assessment

4A. NAEP Assessments and the Nation's Report Card

The NAEP Science Assessment provides the nation with a snapshot of what U.S. students know and are able to do in science. Results of the NAEP Science Assessment administrations are typically reported in terms of average scores for groups of students on the NAEP 0–300 scale and as percentages of students who attain each of the three achievement levels (*NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*). This is an assessment of overall achievement, not a tool for diagnosing the needs of individuals or groups of students. Reported scores are always at the aggregate level; by law, scores are not produced for individual schools or students. Results are reported for the nation as a whole, for regions of the nation, and sometimes for states and large districts that volunteer to participate. The NAEP results are published in an interactive report online as The Nation's Report Card.

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 science achievement, such as performance by disciplinary area 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 science activities, confidence in science knowledge and skills, teachers' satisfaction, and views of school resources).

In 2002, NAEP initiated the Trial Urban District Assessment (TUDA) program 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, and Los Angeles Unified School Districts and New York City Public Schools). In 2003, additional large urban districts began to participate in these assessments, growing to a total of 27 districts by 2017. Sampled students in TUDA districts are assessed in the same subjects and use the same NAEP field materials as students selected as part of national or state samples. TUDA results are reported separately from the state in which the TUDA is located, but results are not reported for individual students or schools. With student performance results reported by district, participating TUDA districts can use results for evaluating their achievement trends and for comparative purposes.

4B. Reporting Scale Scores and Achievement Levels

NAEP typically reports average results on a scale of 0–300 in science. In the past, the average scores have also been reported on three disciplinary groups: Life Science, Physical Science, and Earth and Space Sciences. Reports from the new assessment will include average scores on the same three disciplinary groups, with an updated title for each to reflect the emphasis on student scientific reasoning and problem solving on the assessment. These reporting categories harken back to the claims described in Section 1D (pages 4-5) with their focus on sensemaking and multidimensionality. Scale scores for the disciplinary groups will be reported using the following definitions of each reporting category:

- **Sensemaking in Physical Science**: The student reasons scientifically using NAEP DCs in **physical science**, in combination with NAEP SEPs and CCCs.
- **Sensemaking in Life Science**: The student reasons scientifically using NAEP DCs in **life science**, in combination with NAEP SEPs and CCCs.
- **Sensemaking in Earth and Space Sciences**: The student reasons scientifically using NAEP DCs in **Earth and space sciences**, in combination with NAEP SEPs and CCCs.

Despite the focus on multiple dimensions, NAEP will not report on any of the three dimensions separately, as they work together. That is, there will be no separate scores for students' knowledge of NAEP DCs, SEPs, or CCCs. Given the goal to report on sensemaking in the three disciplinary groupings, all three dimensions are essential in surfacing and measuring students' abilities to apply their understanding of the NAEP DCs to real-world contexts—the phenomena and problems that frame each item and group of items.

These definitions are intended to emphasize that a score for each disciplinary group reflects students' abilities to integrate the three dimensions of science—NAEP DCs, SEPs, and CCCs—and does not prioritize knowledge of the NAEP DCs.

Since 1990, the Governing Board has used 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 very general terms what students at each grade level should know and be able to do on the assessment (see Exhibit 4.1). Achievement level descriptions specific to the 2028 NAEP Science Framework are included in Appendix B. These will be used to guide item development and initial stages of standard setting for the 2028 NAEP Science Assessment (if it is necessary to conduct a new standard setting).

Reporting on achievement levels is one way the Nation's Report Card helps the general public and policymakers interpret NAEP results. Results are reported as percentages of students within each achievement level range as well as the percentage of students at or above *NAEP Basic* and at or above *NAEP Proficient*. Students performing at or above the *NAEP Proficient* level on NAEP assessments demonstrate solid academic performance and competency over challenging subject matter. Following the first administration of the science assessment based on the updated framework, new Reporting Achievement Level Descriptions (ALDs) will be created to specify certain skills in which students are likely to have demonstrated competency at each achievement level. Results for students not reaching the *NAEP Basic* achievement level are reported as below *NAEP Basic*. As noted, individual student performance cannot be reported based on NAEP results.

Note that the *NAEP Proficient* achievement level does not represent grade-level proficiency as determined by other assessment standards (e.g., state or district assessments), and there are significant differences between achievement in the context of NAEP as compared to the context of state-level annual tests. For one, teachers and students are not expected to have studied the NAEP framework or systematically aligned state standards or local curricula with it, nor are students expected to study for the assessment. Furthermore, the NAEP assessment is broader than a typical state grade-level test, for NAEP covers multiple years of study and does not focus on specific instructional units and school years. In addition, there is not a uniform definition of grade-level proficiency across states.

All achievement level setting activities for NAEP are performed in accordance with current best practices in standard setting and the Governing Board's Developing Student Achievement Levels for the National Assessment of Educational Progress Policy Statement (2018). The Governing Board policy does not extend to creating achievement level descriptions for performance below the *NAEP Basic* level.

Achievement level descriptions specific to the NAEP Science Framework were developed to elaborate on the generic definitions. Exhibit 4.1 presents the generic policy definitions. See Appendix B for the achievement level descriptions that illustrate how the policy definitions apply to NAEP Science for grades 4, 8, and 12.

Exhibit 4.1. Generic Achievement Level Policy Definitions for NAEP

Achievement level	Definition
NAEP Advanced	This level signifies superior performance beyond *NAEP Proficient*.
NAEP Proficient	This level represents solid academic performance for each NAEP assessment. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.
NAEP Basic	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for performance at the *NAEP Proficient* level.

4C. Contextual Variables

NAEP legislation[14] requires reporting according to various student populations (see section 303[b][2][G]), including:

- gender,
- race/ethnicity,
- eligibility for free/reduced-price lunch,
- students with disabilities, and
- English Learners.

NAEP users mistakenly may presume that the categories used to report data are related to causal explanations for observed differences (e.g., that gender predicts or explains performance differences or “achievement gaps”). However, scholars find that these differences reflect gaps in students’ opportunities to learn. When results are interpreted in ways that emphasize achievement gaps without attending to opportunity gaps, score differences across subgroups of students can be misinterpreted as differences in student ability rather than as differences due to unequal educational opportunities.

The *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014) recommend that reports of group differences in assessment performance be accompanied by relevant contextual information, where possible, to both discourage erroneous interpretation and enable meaningful analysis of the differences. That standard reads as follows:

Reports of group differences in test performance should be accompanied by relevant contextual information, where possible, to enable meaningful interpretation of the differences. If appropriate contextual information is not available, users should be cautioned against misinterpretation. (Standard 13.6)

Contextual data about students, teachers, and schools are needed to fulfill the statutory requirement that NAEP include information, whenever feasible, that promotes meaningful interpretation of NAEP results. Contextual variables are selected to be of topical interest, timely, and directly related to academic achievement and current trends and issues in science. In the past, a range of information has been collected as part of NAEP.

4D. Science-Specific Contextual Variables

As noted in Chapter 1, research has informed an expanded view of the factors that shape opportunities to learn (Anderson et al., 2018; NRC, 2015; Penuel et al., 2015) including time, content and practices, instructional strategies (e.g., how students are grouped for learning; the scientific tasks they engage in; the opportunities students have to reason, model, and debate

[14] National Assessment of Educational Progress Authorization Act, Pub. L. 107-279, 3 U.S.C. §§ 301–305 (2002)

ideas), and instructional resources—human, material, and social resources that shape student access to science (Brown, 2019; NRC, 2015).

Studies have demonstrated that what students learn is shaped by the availability of various science programs, curricula, extracurricular activities geared toward science, proximity to a science museum or a science and technology center, the percentage of teachers certified in science subjects, teacher years of experience, percentage of science teachers on an emergency license or vacancies / substitute teachers in the school, and number of teachers with science degrees, among other factors (Anderson et al., 2018; NRC, 2015; Penuel et al., 2015). Teachers' and administrators' beliefs about what science is, how one learns science, and who can learn science also affect student learning (Anderson et al., 2018; Brown, 2019). What students learn is shaped by their sense of identity and agency. Students who see themselves, and who are seen by others, as capable scientific thinkers are more likely to participate in ways that further their learning; students who do not see themselves, and are not seen by others, as capable scientific thinkers are likely to be disengaged (Brown, 2019). Steele, Spencer, and Aronson (2002) found that even passing reminders that a student is a member of one group or another—often, in this case, a group that is stereotyped as intellectually or academically inferior—can undermine student performance.

There are countless factors that shape what and when students learn. The NAEP Science student, teacher, and administrator surveys cannot possibly cover all such factors. Even though it would be helpful to ask students and teachers the same questions, this is also not possible given time constraints. Student questionnaires have a strict time limit of 15 minutes. There are also limitations on the content of the questions that can be included on the student questionnaire. Questions about some factors may not be appropriate in the NAEP context. The NAEP legislation prohibits the collection of information on personal or family beliefs and attitudes and specifies that only information directly related to the appraisal of academic achievement may be collected. Given the constraints, not all topics can be addressed.

To support prioritization and ensure that NAEP results have appropriate context for interpretation, this framework sets the following topics to receive the greatest emphasis in the 2028 NAEP Science Assessment's contextual questionnaires (in order of priority).

- *Science content*. The 2028 NAEP Science Framework conceptualizes science content as disciplinary concepts, science and engineering practices, and crosscutting concepts. Therefore, contextual variables related to science content are expanded to include reference to NAEP SEPs and NAEP CCCs as well. Interpreting students' achievement requires a basic understanding of what NAEP DCs, SEPs, and CCCs students have engaged with. Given variation across states in standards and frameworks, this information is crucial.
- *Teacher factors*. Research demonstrates that teacher quality is the most important in-school factor in predicting student achievement. This framework prioritizes the collection of data on teacher preparation and professional development.

- *Student science identity*. Research demonstrates that students' perceptions of their science identity directly relate to their learning. This framework prioritizes gathering information about students' science identities through questions that address student participation in activities such as discussion of phenomena, science ideas, or evaluation of how a science problem or investigation is framed.
- *Instructional resources*. A range of resources influences instruction, including instructional leadership, additional instructional personnel, time, technology, curriculum, and materials. This framework prioritizes gathering information about school resources that can inform the interpretation of results, including the time devoted to science teaching and learning in school, across current and prior grade levels, and the curricular and instructional materials at teachers' and students' disposal to support learning. In terms of technology, questionnaires will capture what technology is available to support science and engineering teaching and learning and how it is used.
- *Instructional organization and strategies*. Interpreting student achievement levels will also depend on understanding the instructional strategies used in science class, including collaborating in small-group work, engaging in science discussions, working hands-on and using grade-appropriate measurement and data analysis tools, and using a range of methods and tools to represent and model science phenomena and engineering design problems. This framework prioritizes gathering information both on the organization of classrooms and on the instructional routines and approaches that teachers use. It also includes what technologies and assessment approaches are used in instruction.

4E. Conclusion

As the Nation's Report Card, NAEP reports on student performance over time, presenting an analysis of national trends in students' science achievement. The 2028 NAEP Science Assessment is designed to assess the achievement of groups of students through robust and challenging assessments that are well aligned with current understanding of the three dimensions of science to be learned and that use technology in ways that maximize both student engagement and accessibility. The results of the assessment are informed by data on contextual variables that illuminate potential differences in opportunities to learn for students.

The ultimate goal of our nation's schools is to ensure that every student can realize their full academic potential through access to high-quality science instruction, curriculum, and resources. NAEP plays an important role in providing to the nation a broad picture of students' knowledge and skills in science. NAEP scores, illuminated by relevant contextual information, can provide the public, families, students, and schools useful data on student performance that complements information provided by state tests that are more tightly aligned with specific state standards. As a view of present trends, it provides invaluable data to inform policy and practice in the future.

APPENDIX A: Sample Item Metadata

Exhibit A.1. Park Flooding, Version 1

Item ID: Park Flooding (adapted from Formative Assessment Bundling Literacy and Elementary Science)
Grade and discipline: Grade 4, Earth and Space Sciences
Item type: Single part, single-select multiple choice
Alignment: This item is a **two-dimensional item**, measuring parts of the following:

- DC: E4.9: Patterns in when and where weather conditions occur can be used to make predictions about the kind of weather that can be expected in a region.
- SEP: S4.10: Analyze data to determine whether it supports or refutes a claim about a phenomenon or design solution.

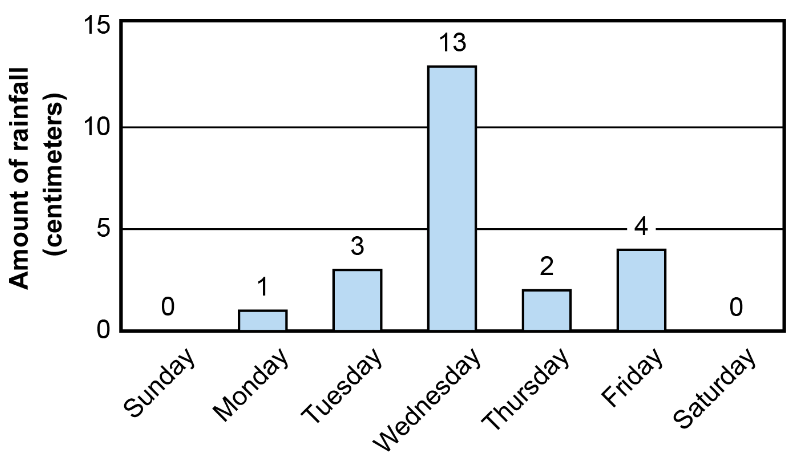
Phenomenon: A park flooded when it was raining one day but not other days.

People visiting a local park noticed that the park was flooded and was closed for the day. The picture shows the flooded park.

Flooded Park[15]

The park was closed only on the day the flooding happened. The bar graph shows the rainfall for each day of that week.

Amount of Rainfall at the Park



Day of the Week

[15] Betty Longbottom / Flooded Playground! - Cliffe Avenue / CC BY-SA 2.0

Based on the data, on which day was the park **most likely** flooded?

A. Sunday
B. Monday
C. Tuesday
D. Wednesday

Key: D

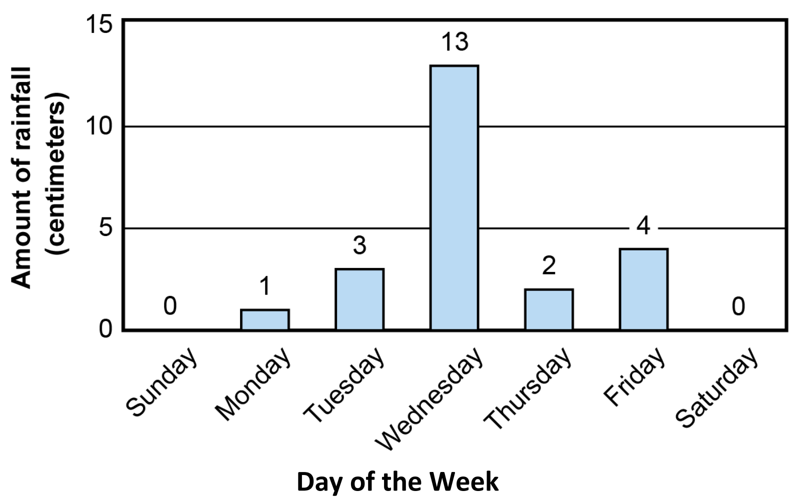
Exhibit A.2. Park Flooding, Version 2

People visiting a local park noticed that the park was flooded and was closed for the day. The picture shows the flooded park.

Flooded Park[16]

The park was closed only on the day the flooding happened. The bar graph shows the rainfall for each day of that week.

Amount of Rainfall at the Park



[16] Betty Longbottom / Flooded Playground! - Cliffe Avenue / CC BY-SA 2.0

Part A

Based on the data, on which day was the park **most likely** flooded?

A. Sunday
B. Monday
C. Tuesday
D. Wednesday

Key: D

Part B

Based on your understanding of weather, which piece of evidence **best** supports your answer in Part A?

A. This day was rainy.
B. The rain started falling on this day.
C. This day had more rainfall than any other day in the week did.
D. The amount of rainfall on this day was lower than on any other day.

Key: C

Exhibit A.3. Locusts Stimulus

Grade and discipline: Grade 8, Life Science and Earth and Space Sciences
Possible NAEP DCs, SEPs, and CCCs that could be included with the stimulus:

NAEP DCs	NAEP SEPs	NAEP CCCs
L8.9: Ecosystems are dynamic in nature; their characteristics can vary over time. Disruptions to any physical or biological component of an ecosystem can lead to shifts in all its populations, therefore helping or hurting the health of the ecosystem, including its biodiversity. E8.12: Human activities have significantly altered the biosphere, atmosphere, and geosphere, sometimes damaging or destroying ecosystems and causing the extinction of organisms. Human choices can minimize harm to other organisms and risks to the health of the regional environment. E8.13: Human activities that release greenhouse gasses, such as production and combustion of fossil fuels, are major factors in the current rise in Earth's temperature. Monitoring the production and reducing the use of fossil fuels can slow the increase in global temperatures as well as the effects of climate change.	S8.9: Analyze data to provide evidence to support or reject a model or explanation or to use to improve a design solution. S8.15: Develop, use, and/or revise a model to describe, explain, and/or predict phenomena by identifying relationships among parts and/or quantities in a system, including both visible and invisible quantities. S8.18: Construct or revise an explanation that uses a chain of cause and effect or evidence-based associations between factors to account for the qualitative or quantitative relationships between variables in a phenomenon. S8.22: Identify evidence that could be used to refute a claim about a phenomenon. S8.24: Compare and critique two arguments on the same question to analyze their fit with the evidence and/or whether they emphasize similar or different evidence and/or interpretations.	C8.3: Cause-and-effect relationships may be used to predict phenomena in natural or designed systems. C8.4: Phenomena may have more than one cause, and some cause-and-effect relationships in systems can only be described using probability. C8.5: The observed function of natural and designed systems may change with scale. Phenomena that can be observed at one scale may not be observable at another scale.

A desert locust is an insect that undergoes changes to its body in certain environmental conditions. Figure 1 describes some differences between two modes of a desert locust.

Figure 1. Desert Locust[17]

Mode 1: Grasshopper (Dry, warm or cool weather)	Mode 2: Locust (Wet/rainy, warm or hot weather)
• Behave independently • Stay away from other desert locusts • Mostly walk slowly and jump • Limited diet • Small, scattered populations that stay in one place • Very stable population; females lay eggs but most don't hatch until the environment is wet and hot.	• Behave as a united group (swarms) • Gather together with other desert locusts • Walk quickly and fly long distances • Broad diet, including crops • Tens of billions of locusts in a swarm that can travel up to 100 miles per day • Population can increase 400x in six months.

When these insects are in Mode 2, they are able to swarm. A single swarm of locusts can cover an area of up to 100 square miles, with 40 to 80 million locusts in each square mile. Swarms can travel up to 100 miles a day. Figure 2 shows the effect of three months of locust swarms on available vegetation in an area of Africa.

[17] Bernard DuPont Bird Locust Attribution-ShareAlike (CC BY-SA 2.0); Magnus Ullman, CC BY-SA 3.0, via Wikimedia Commons

Figure 2. Available Vegetation Before and After Locust Swarms

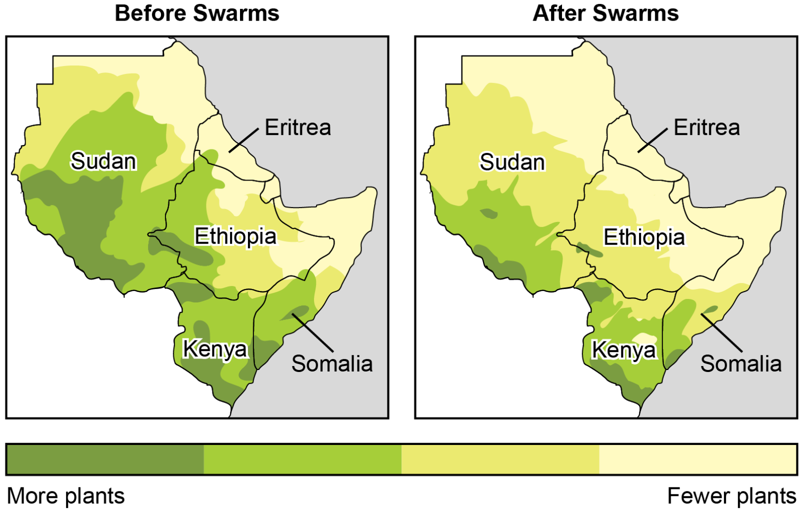


Exhibit A.4. Plant Growth, Version 1

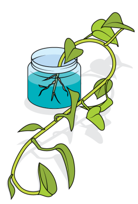
Item ID: Plant Growth (adapted from the Next Generation Science Assessment Projects)
Grade and discipline: Grade 4, Life Science
Item type: Multipart item, single-select multiple choice and short constructed response
Alignment: This item is a **two-dimensional item**, measuring parts of the following:

- DC: L4.3: All animals need food, water, and air in order to live and grow. They obtain their food from their surroundings—from plants or from other animals. Plants need air, water, minerals (in the soil), and light to live and grow.
- SEP: S4.7: Make predictions about what would happen if a variable changes.

Phenomenon: Soil and water conditions affect plant growth.

The plants shown were placed in a classroom on the same day. They are all the same kind of plant. The plants were placed on the same side of the room near a window so they receive the same amount of light each day. Students in the class want to find out what the plants need the most in order to grow. They grow the plants using the conditions shown in the table.

Conditions for Growing Plants

Plant	Planted in Soil	Water
 Plant A	No	Water added regularly for one month
 Plant B	Yes	Water added regularly for one month
 Plant C	Yes	No water added

Part A

Which plant will likely grow the **least** over the next month?

A. Plant A
B. Plant B
C. Plant C

Key: C

Part B

Provide one reason the plant you chose in Part A will grow the least over the next month.

Exhibit A.5. Plant Growth Part B Constructed Response Scoring Notes

- Reasons students provide should leverage understanding of what plants need to grow (water, air, minerals from soil).
- Note that while a complete answer might include comparisons among plants (e.g., Plant A and B have X, but Plant C does not), this is not a requirement.
- Possible reasons include:
    - Plant C does not get water.
    - Plant C does not get minerals.
    - Plant C does not get water or minerals.
- Students should receive credit as long as their reason supports their choice, with an accurate understanding of plant needs for growth.

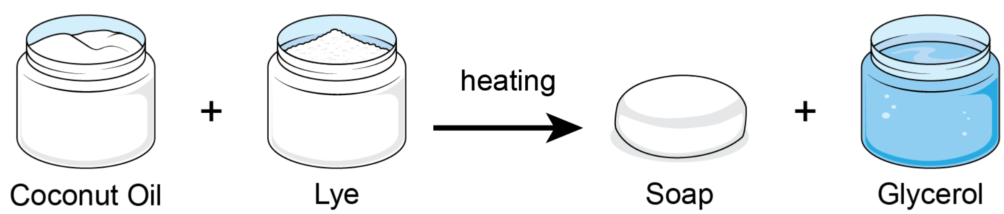
Exhibit A.6. Making Soap

Item ID: Making Soap (adapted from the Next Generation Science Assessment Projects)
Grade and discipline: Grade 8, Physical Science
Item type: Single-select multiple choice
Alignment: This item is a **three-dimensional item**, measuring parts of the following:

- DC: P8.4: In a chemical reaction, the atoms of the reacting substances are regrouped in characteristic ways into new substances with different properties. Atoms only rearrange. As such, the amount of matter does not change.
- SEP: S8.9: Analyze data to provide evidence to support or reject a model or explanation or to use to improve a design solution.
- CCC: C8.1: Patterns in data can be identified and represented using graphs, charts, and tables. Analyzing patterns can help identify cause-and-effect relationships and estimate probabilities of events.

Phenomenon: Heating a combination of coconut oil and lye produces soap.

One way to make soap is to heat a combination of coconut oil and lye. The diagram shows a simplified model of the soapmaking process.



The data table shows properties of each substance in the model of the soapmaking process.

Properties of Soapmaking Substances

Substance	Mass (g)	Odor	Density (g/cm^3)	Melting point (℃)
Coconut oil	100	Coconut	0.93	27
Lye	20	Odorless	2.13	318
Soap	115	Coconut	0.95	48
Glycerol	5	Odorless	1.26	17.8

Which data provide evidence that making soap involves a chemical reaction?

A. Coconut oil and soap both smell like coconut.
B. The density of soap is different from the density of glycerol.
C. The total mass of soap and glycerol is the same as the total mass of coconut oil and lye.
D. The melting points of soap and glycerol are different from the melting points of coconut oil and lye.

Key: D

Exhibit A.7. Human Migration to Appalachia

Item ID: Human Migration to Appalachia
Grade and discipline: Grade 12, Earth and Space Sciences and Life Science
Item set with the following item types: Single-select multiple choice, constructed response
Alignment: This item set is a **three-dimensional item set at medium complexity**, measuring parts of the following:

	NAEP DC	NAEP SEP	NAEP CCC
Item 1	E12.9: Resource availability guides the development of human societies. All forms of energy production and resource extraction have associated economic, social, and environmental cost-benefit factors. L12.13: Humans depend on the living world for the resources and other benefits provided by biodiversity. Changes in biodiversity can influence resources and ecosystem services that humans rely on. high complexity	S12.15: Apply mathematical expressions, computer programs, algorithms, or simulations of a process or system to evaluate the model by comparing the outcomes with what is known about the phenomena or design problem. low complexity	C12.3: Cause-and-effect relationships can explain and predict complex natural and human-designed systems. Such explanations may require examining and modeling small scale mechanisms within the system. [Note: in a scenario-based task (SBT), this CCC could be further explored by examining the mechanisms within any of the subsystems included as part of the inputs of this model.] low complexity
Item 2	L12.12: Changes induced by human activity in the environment—such as habitat destruction, pollution, introduction of invasive species, overexploitation, and climate change—can disrupt an ecosystem, reduce biodiversity, and threaten the survival of some species. low complexity	S12.14: Apply or revise algorithms when analyzing data or designing, programming, testing, and revising scientific models, explanations, and design solutions. medium complexity	C12.3: Cause-and-effect relationships can explain and predict complex natural and human-designed systems. Such explanations may require examining and modeling small-scale mechanisms within the system. medium complexity

Phenomenon/Problem: Human migration to Appalachia has been greater than predicted by computer models. The model used to make the prediction needs to be revised to better reflect the factors that influence migration into different regions of Appalachia.

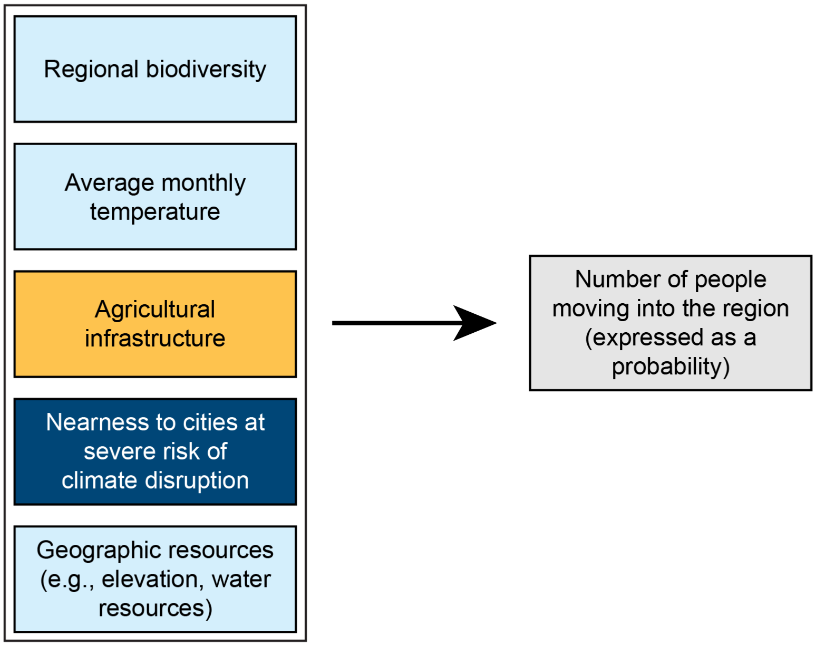
Appalachia is considered “climate resilient.” This means that the area can successfully handle the impacts of changes to climate and can prevent those impacts from growing worse. The green areas in Figure 1 show where Appalachia is located in the United States.

Figure 1. Map of Appalachia

Computational models predict that many people will move into the Appalachian region over the next 20 years as they seek to find places to live that are safer and more stable.

Figure 2 shows one model local leaders are using to predict how many people will move into Appalachia. Blue indicates factors that are expected to increase migration, and orange indicates factors expected to decrease migration. Darker colors indicate more weight on that factor in the model. Agricultural infrastructure includes farms, markets and businesses that support farms, and the transportation and communication systems in the area.

Figure 2. Computational Model for Predicting Human Migration



When this model was tested against recent population growth due to migration into two locations in Appalachia, leaders noticed some differences between what the model predicted and what the data showed. The table shows these differences for the two locations, along with information about how high or low each location is rated on several factors.

Predicted and Actual Population Growth in Pittsburgh and the Shenandoah Valley

Location	Pittsburgh	Shenandoah Valley
Predicted population growth	high	low
Actual population growth	low	high
Relative biodiversity	low	high
Average monthly temperature range	29–73° F	32–74° F
Relative agricultural infrastructure	low	medium
Nearness to climate-impacted urban centers	high	high
Access to usable water	medium	high

Item 1

Based on your understanding of what human societies need to be successful, which idea **best** explains why the model prediction was different from the observed data?

A. Pittsburgh has highly limited access to water. The model did not account for how this shortage would limit the growth of local businesses.
B. Human societies require reasonable annual temperature ranges. The model incorrectly assumed that average monthly temperature was a less important factor.
C. Human societies require access to food and water. The model did not account for how important access to food and water would be for human migration to Appalachia.
D. The Shenandoah Valley has very little access to fresh fruits and vegetables. The model incorrectly assumed that agricultural infrastructure would not promote migration.

Key: C

Item 2

Human migration to areas such as Appalachia can result in rural gentrification. People who are currently living in low-cost, natural resource-rich areas are forced to leave, resulting in the loss of access to the resources and communities they have actively contributed to developing and maintaining. As a result, they can no longer enjoy the benefits of these regions.

Describe **one** way you could revise or build on the computational model to better understand how migration into Appalachia could impact current residents' access to natural resources. Be sure to explain how the change you describe will provide a better understanding of how migration into Appalachia could impact current residents' access to natural resources.

Exhibit A.8. Human Migration to Appalachia Item 2 Constructed Response Scoring Notes

Emphasis here is on an understanding of how to revise the model parameters to better understand more nuanced population-environment dynamics. Student ideas can be wide-reaching but should be justifiable as at least one of the following: (a) updating the model to better understand who moves into and out of the region, and/or (b) determining the feedback impacts of population growth on biodiversity, and natural resources such as water. This item specifically focuses on the computational reasoning aspect of this SEP and can include either quantitative or qualitative reasoning from students. Appropriate lines of reasoning here can include, but are not limited to, the following:

- updating the model to account for demographic subgroups
- considering housing costs / other metrics for socioeconomic status as part of the inputs and/or outputs, as a mechanism to better understand the relative wealth/characteristics of who lives in the area
- calculations of migration out of Appalachian regions
- feedback mechanisms that influence biodiversity and natural resource availability (Note: This could be specific to particular resources, or general at the level of the categories included in the model.)
- relative factors for scaling variables (quantitative or qualitative)

Exhibit A.9. Permafrost Melting, Version 1

Item ID: Permafrost Melting (adapted from OpenSciEd)
Grade and discipline: Grade 12, Physical Science
Item type: Selected response, matching
Alignment: This item is a **three-dimensional item**, measuring parts of the following:

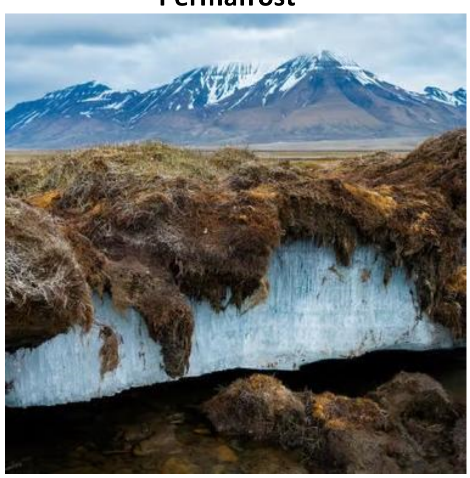
- DC: P12.14: When sunlight is absorbed at Earth's surface, it is eventually reradiated as infrared radiation that transfers heat into the atmosphere. The average temperature of the atmosphere is determined by how long the energy stays in the system until it is reradiated into space from the top of the atmosphere.
- SEP: S12.16: Develop, use, and/or revise a model that includes mathematical relationships (including both visible and invisible quantities) to describe, explain, and/or predict phenomena or to test a proposed design solution.
- CCC: C12.13: Feedback mechanisms within a system are important elements for explaining or designing for either the stability or instability of the system.

Phenomenon: Permafrost is melting, and that melting seems to be associated with increased global temperatures.

Permafrost is a layer of soil and ice that is just below the surface in the Arctic. An example of permafrost is shown in the picture.

Historically, permafrost stayed frozen for many years at a time. However, in some areas permafrost is now melting, which can cause many changes to Earth's surface and living things. One major concern is that permafrost contains carbon dioxide.

Permafrost[18]



[18] Permafrost from OpenSciEd, CC BY-SA 4.0

Use the statements to develop a model that shows the relationships between thawing permafrost and rising global temperatures. Drag the statements into the correct boxes to complete the model. Each statement will be used once.

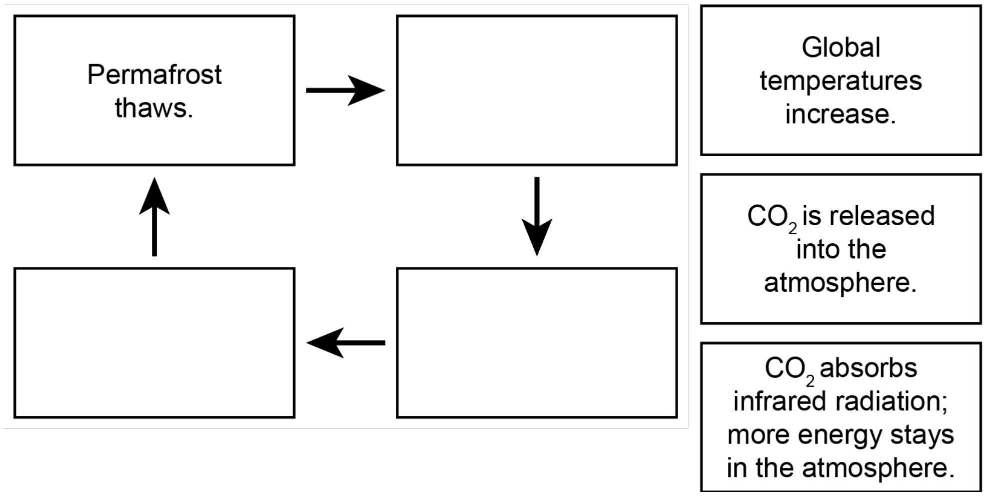


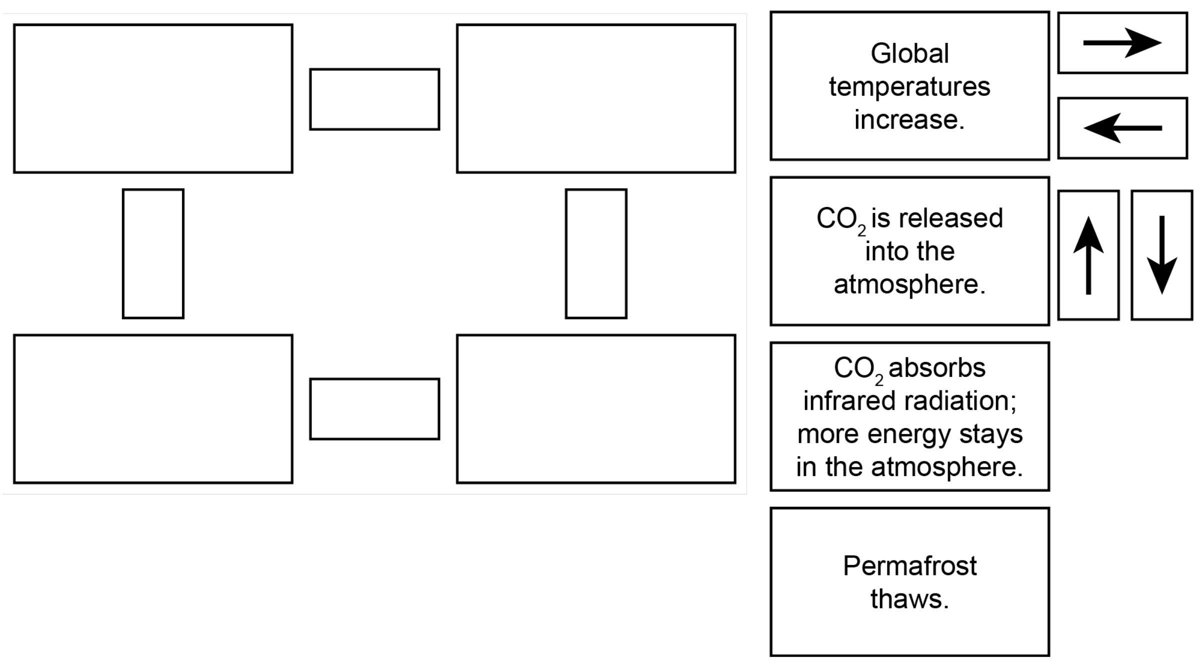
Exhibit A.10. Permafrost Melting, Version 2

Permafrost is a layer of soil and ice that is just below the surface in the Arctic. An example of permafrost is shown in the picture.

Historically, permafrost stayed frozen for many years at a time. However, in some areas, permafrost is now melting, which can cause many changes to Earth's surface and living things. One major concern is that permafrost contains carbon dioxide.

Permafrost[19]

Use the statements and arrows to develop a model that shows the relationships between thawing permafrost and rising global temperatures. Drag the statements and arrows into the correct boxes to develop the model. Each statement and each arrow will be used once.



[19] Permafrost from OpenSciEd,

Exhibit A.11. Permafrost Melting, Version 3

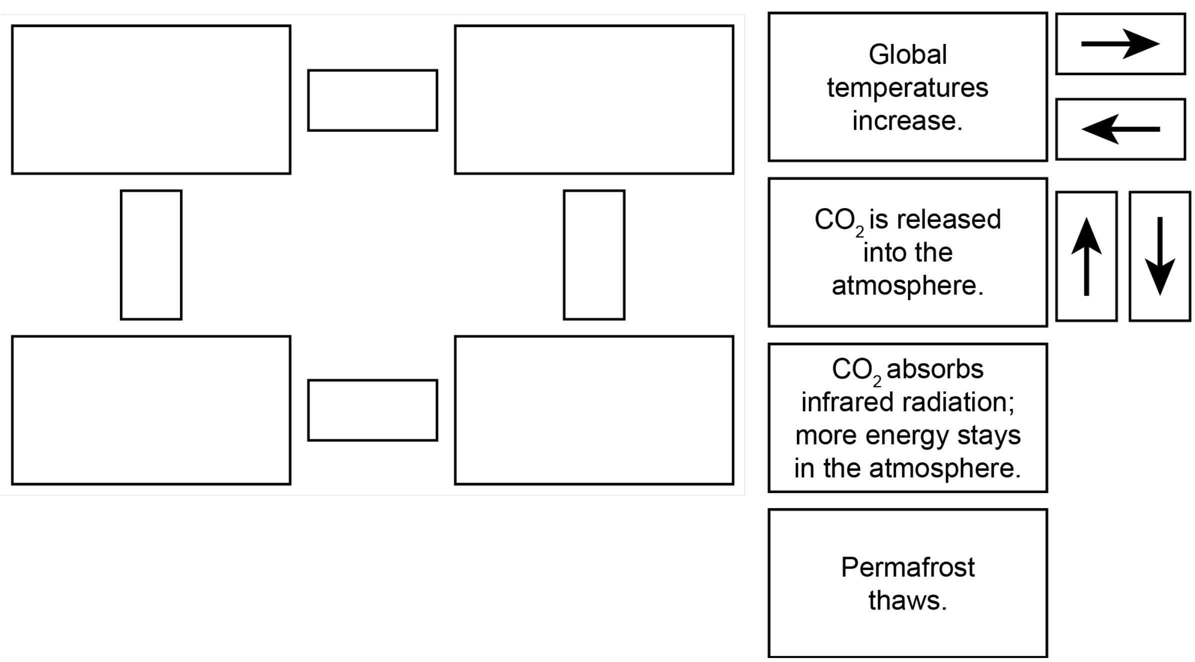
Permafrost is a layer of soil and ice that is just below the surface in the Arctic. An example of permafrost is shown in the picture.

Historically, permafrost stayed frozen for many years at a time. However, in some areas, permafrost is now melting, which can cause many changes to Earth's surface and living things. One major concern is that permafrost contains carbon dioxide.

Permafrost[20]

Part A

Use the statements to develop a model that shows the relationships between thawing permafrost and rising global temperatures. Drag the statements and arrows into the correct boxes to complete the model. Each statement and arrow will be used once.



[20] Permafrost from OpenSciEd, CC BY-SA 4.0

Part B

Complete the sentence by choosing the correct answer from the drop-down menu.

The model predicts that the rate of melting of the permafrost will likely (increase / stay the same / decrease) over the next 50 years.

Key: increase

Part C

Describe a limitation of the model you developed in Part A.

Exhibit A.12. Permafrost Melting Part C Constructed Response Scoring Notes

- Students provide one limitation of the model, for example:
  - The model is missing factors such as interactions with other systems and feedback loops.
  - Thawing may occur at different rates in different locations around the world.
  - There may be a time lag between temperature increases and thawing.

Exhibit A.13. Limu Kohu

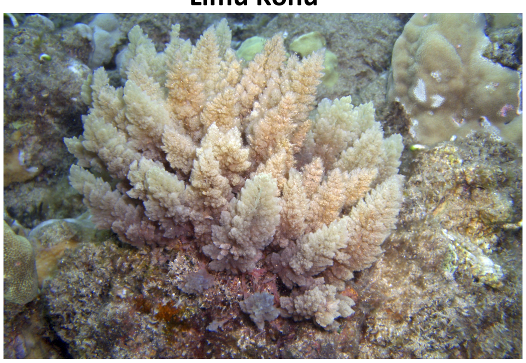
Item ID: Limu Kohu (adapted from the State Performance Assessment Learning Community)
Grade and discipline: Grade 8, Earth and Space Sciences
Item type: Multipart, single-select multiple choice, short response constructed response
Could part A be used as a standalone item? Yes
Alignment: This item is a **three-dimensional item**, measuring parts of the following:

- DC: E8.12: Human activities have significantly altered the biosphere, atmosphere, and geosphere, sometimes damaging or destroying ecosystems and causing the extinction of organisms. Human choices can minimize harm to other organisms and risks to the health of the regional environment.
- SEP: S8.22: Identify evidence that could be used to refute a claim about a phenomenon.
- C8.3: Cause-and-effect relationships may be used to predict phenomena in natural or designed systems.

Phenomenon and engineering design problem: Limu kohu populations have been declining. Students are asked to make sense of the role of human activities in causing this problem.

Limu kohu is a type of seaweed that is native to the waters around Honolulu, Hawaiʻi. It is an important part of food systems as well as cultural and religious practices. Although limu kohu was easy to find for hundreds of years, limu kohu populations around Honolulu have been rapidly declining over the past 60 years. An example of limu kohu seaweed is shown in the picture. The table describes observations of limu kohu.

Limu Kohu[21]



Observations from Generations of Hawaiian Elders about Limu Kohu Growth and Harvesting
• Limu kohu needs warm water and high salinity to grow. • Limu kohu grows and reproduces well on the edges of coral reefs. • When limu kohu is trimmed, it regrows. • When the base of the limu kohu is harvested, it cannot regrow.

[21] MDC Seamarc Maldives, CC BY-SA 4.0, via Wikimedia Commons

Part A

Which human activity is **least** likely to cause harm to limu kohu populations?

A. Companies using industrial methods of harvesting limu kohu remove the whole limu kohu plant.
B. Restaurants using traditional methods of harvesting limu kohu remove the top of the limu kohu.
C. Industrial runoff changes the temperature and salinity of the water in coastal regions where limu kohu live.
D. Ships visiting Hawaiʻi introduce invasive seaweed species that use the same resources as limu kohu into coastal regions where limu kohu live.

Key: B

Part B

Use the information provided and your understanding of the impacts of human activities on the environment to support your answer to Part A.

Exhibit A.14. Limu Kohu Part B Constructed Response Scoring Notes

Students provide one statement that is based on the information provided, shows understanding of the impacts of human activities on the environment, and supports the answer to Part A. For example:

- Traditional harvesting practices that focus on harvesting only the top portions of limu kohu are less destructive compared to other methods that involve uprooting the entire plant. When only the top is harvested, the base of the plant remains intact, allowing the limu kohu to potentially regrow.
- The information provided indicates that limu kohu has the ability to regrow when it is trimmed. This suggests that harvesting the upper parts of the seaweed allows it to regenerate, ensuring the sustainability of the population over time.

APPENDIX B: Achievement Level Descriptions

The NAEP Achievement Level Descriptions (ALDs) in this appendix provide examples of what students performing at the *NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced* achievement levels should know and be able to do in terms of the science disciplinary content, science and engineering practices, and crosscutting concepts identified in the framework.

The ALDs in the 2028 NAEP Science Framework have changed, relative to ALDs presented in previous frameworks. The differences reflect not only changes to the science knowledge, skills, and abilities assessed (science disciplinary content, science and engineering practices, and crosscutting concepts) but also an effort to develop ALDs that provide explicit guidance for item developers. Specifically, across grade levels, the 2028 NAEP Science Framework ALDs have changed in the following ways:

- Updates to the grade-level objectives in Chapter 2 of the framework are reflected in the content foci described in each grade-level ALD.
- The science practices from previous science frameworks have been expanded to Science and Engineering Practices (SEPs) for the 2028 NAEP Science Framework, and Crosscutting Concepts (CCCs) have been added. A new paragraph was developed to show the progression of application of the SEPs and CCCs to make sense of science phenomena.
- To provide specific and unambiguous guidance to item developers, these ALDs provide more explicit elaborations of the knowledge and skills students should demonstrate and the actions they should perform at each grade level and within each achievement level. In addition to the overall section and the section on SEPs and CCCs, the ALDs continue to be broken out by science domain: Life Science, Physical Science, and Earth and Space Sciences. The ALDs provide samples of how the SEPs and CCCs can be applied to specific concepts within each domain.

Within each grade level, the shifts from one achievement level to the next have commonalities, and the content of each achievement level can be described generally. Descriptions at each achievement level, for all grade levels, are as follows:

- Descriptions at the *NAEP Basic* achievement level focus on partial understanding of grade-appropriate concepts and simple applications of SEPs and CCCs to that content to make sense of real-world situations and common phenomena.
- Descriptions at the *NAEP Proficient* achievement level focus on solid understanding of grade-appropriate concepts and skillful application of SEPs and CCCs to that content to reason with and interpret real-world situations and phenomena.
- Descriptions at the *NAEP Advanced* achievement level focus on superior understanding of grade-appropriate concepts and expert engagement with SEPs and CCCs to that content to interpret, explain, and predict real-world situations and phenomena.

Text that elaborates on these statements is included within the ALD tables. The ALDs are organized into three sections. The first demonstrates how the DCs progress in depth of understanding from *NAEP Basic* to *NAEP Proficient* to *NAEP Advanced*. The second section illustrates the same progression for the SEPs and CCCs. However, the intent is that none of these dimensions is assessed in isolation. Therefore, the third section provides examples of the intersection of the DCs with the SEPs and CCCs. The third section is subdivided by discipline, with four examples per discipline. Although it wasn't possible to write ALDs for every possible crossing of DCs with SEPs and CCCs, the four selected represent big ideas that can also be shown to progress from grade 4 to grade 8 to grade 12.

To add clarity and specificity, the NAEP Science Assessment and Item Specifications include example performance expectations targeting each achievement level within each grade level. In Appendix A of the Assessment and Item Specifications, three items (one each for grades 4, 8, and 12), along with annotations that describe items across the achievement levels, illustrate the knowledge and skills required at different NAEP achievement levels.

Exhibit B.1. NAEP Grade 4 Science Achievement Level Descriptions

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
Students at this level should be able to demonstrate partial mastery and competency in making sense of common phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. different types of matter (materials) have different properties, B. a force acting on an object at rest can move the object, C. water and light are needed for a plant's survival, D. the location of rocks and fossils can be used to establish Earth's history, E. Earth surface features can be changed by natural processes, such as wind or water or living organisms, F. humans can impact the land, water, and air where they live.	Students at this level should be able to demonstrate solid academic performance and competency in making sense of phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. matter (materials) can be classified based on its properties, B. a change in motion requires unequal forces acting on an object, C. varying amounts of water and light may affect a plant's growth, D. fossils can provide evidence for the nature of an environment where organisms lived long ago, E. some changes to Earth's surface features by wind, water, and living organisms can be observable, F. humans can impact the land, water, and air around the world.	Students at this level should be able to demonstrate superior performance and competency in making sense of complex phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. different types of matter (materials) have multiple different properties; it is therefore necessary to consider pros and cons when selecting a material for a specific purpose, B. two objects can each exert a force on the other; the change in motion of either object depends on all the forces that act on it, C. some animals obtain the matter they need for growth and survival from plants or from other animals, D. the location of fossils within rock layers can be used to show the changes that occurred to Earth and life on Earth over time, E. small changes to Earth's surface features by water, wind, or organisms can result in large changes over time, F. humans can impact the land, water, and air around the world in positive and negative ways.

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
Working with the disciplinary concepts, students require a well-defined set of actions to be able to apply science and engineering practices and crosscutting concepts such as: A. asking a simple question about an observation, B. using models to describe a phenomenon or describe a design proposal, C. using a diagram to identify one way that changes might affect a phenomenon, D. identifying an evidence-based argument, E. describing quantitative evidence related to a design problem, F. using evidence to support the solution to a design problem while considering the criteria that the solution should meet, G. developing an evidence-based description of a phenomenon.	Working with the disciplinary concepts, students require some cueing to be able to apply science and engineering practices and crosscutting concepts such as: A. asking questions about observed data to aid in interpretation, B. describing how the parts and relationships in a model represent a phenomenon or proposed design solution, C. describing observations or measurements that can be used as evidence to explain a phenomenon, D. evaluating the merits of an evidence-based argument, E. organizing data to reveal patterns that can be used to solve a design problem, F. making a claim about the solution to a design problem using evidence while considering criteria and constraints, G. using patterns in information to develop an evidence-based explanation of a phenomenon.	Working with the disciplinary concepts, students require limited cueing to be able to apply science and engineering practices and crosscutting concepts such as: A. using questions as a tool to clarify an argument or investigate a problem, B. identifying the limitations of a model used to represent a phenomenon or proposed design solution, C. predicting the outcome of an experiment designed to explore changes to a phenomenon, D. comparing evidence-based arguments about the changes of a system based on the evidence or the reasoning the arguments include, E. estimating or predicting data points using patterns in recorded data to solve a design problem, F. proposing a solution to a design problem using evidence to help ensure it will meet criteria and constraints, G. developing an evidence-based explanation of a phenomenon supported by reasoning about cause-and-effect relationships.

<table>
<tr><th>NAEP Basic Level</th><th>NAEP Proficient Level</th><th>NAEP Advanced Level</th></tr>
<tr><td>In Physical Science, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the simple sensemaking of common phenomena when provided a well-defined set of actions to perform tasks such as:
A. using evidence to describe that temperature affects the physical state of a material (solid vs. liquid),
B. proposing appropriate variables and tests when planning an investigation to determine whether objects exert a force on each other when they collide,
C. organizing data to identify a pattern between how fast an object moves and its energy,
D. asking questions to clarify the relationship between the parts of a simple model used to represent that an object can be seen when light produced by the object or reflected from its surface enters the eyes.</td><td>In Physical Science, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the moderate degree of sensemaking of phenomena when provided some cues to perform tasks such as:
A. making a claim using data about the relationship between temperature and the physical state of a material (solid vs. liquid),
B. planning an investigation considering the variables to control and/or the number of trials to conduct to produce data about whether objects exert forces on each other when they are touching or colliding,
C. describing patterns in data to support the claim that the motion of an object is related to its energy,
D. asking questions to evaluate a model that represents that objects can be seen only when light produced by the object or reflected from its surface enters the eyes.</td><td>In Physical Science, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the high degree of sensemaking of complex phenomena when provided limited cues to perform tasks such as:
A. evaluating a claim based on the evidence or reasoning it includes about how a material might function differently at a different temperature due to the physical state of the material (solid vs. liquid) at that temperature,
B. evaluating an investigation plan designed to explore a scientific question about the way that objects exert forces on each other when they touch or collide,
C. analyzing patterns in data gathered from the motion of two different objects to support claims about the relationship between the motion and speed of objects and their energy,
D. asking questions to identify the limitations of a model that represents the conditions under which various objects and materials can be seen.</td></tr>
</table>

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *simple sensemaking* of common phenomena when provided *a well-defined set of actions* to perform tasks such as: A. evaluating textual information to determine whether it is related to given data about the similarities and differences between the life cycles of different animals, B. using data to describe that some of the matter that an organism needs to survive comes from other organisms, C. supporting a claim that inherited characteristics can affect what an organism looks like, D. identifying evidence for an argument about characteristics of an individual providing advantages in surviving.	In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *moderate degree of sensemaking* of phenomena when provided *some cues* to perform tasks such as: A. evaluating whether textual information accurately summarizes data about the diverse life cycles of plants or animals, B. analyzing data that can be used to support a claim that much of the matter organisms need to grow and survive comes from other organisms, C. making an evidence-based claim about the relationship between organisms' characteristics and their look and function, D. supporting an argument with evidence about different characteristics of individuals providing advantages in surviving and finding mates.	In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *high degree of sensemaking* of complex phenomena when provided *limited cues* to perform tasks such as: A. evaluating whether textual information summarizing a table of data comparing the life cycles of different plants or animals accurately reflects a claim about the essential nature of reproduction for all organisms, B. analyzing two different sets of data to determine which one can be used to support a claim that much of the matter organisms need to grow and survive comes from other organisms and that same matter is used again later by other organisms, C. evaluating the evidence to support or reject various claims about whether the characteristics of an organism are inherited, result from interactions with the environment, or both, D. evaluating evidence and reasoning of arguments about whether different characteristics of individuals can provide advantages in surviving and finding mates when a habitat changes.

<table>
<tr><th>NAEP Basic Level</th><th>NAEP Proficient Level</th><th>NAEP Advanced Level</th></tr>
<tr>
<td>In Earth and Space Sciences, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the simple sensemaking of common phenomena when provided a well-defined set of actions to perform tasks such as:
A. using a model to describe that objects in the sky are not always visible from Earth,
B. identifying evidence to support arguments that the Earth's surface was different in the past,
C. analyzing patterns in data to describe a weather event that occurred in a region,
D. identify evidence that can be used to support a claim about how natural processes can cause hazards for humans in some areas.</td>
<td>In Earth and Space Sciences, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the moderate degree of sensemaking of phenomena when provided some cues to perform tasks such as:
A. developing a model to describe that objects in the sky are not always visible due to Earth's rotation,
B. making arguments based on evidence about differences of the surface of the Earth between the present and in the past,
C. analyzing patterns in data to describe the kind of weather expected in a region,
D. supporting an argument with evidence for how natural processes can cause hazards for humans in some areas but not others.</td>
<td>In Earth and Space Sciences, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the high degree of sensemaking of complex phenomena when provided limited cues to perform tasks such as:
A. developing and using a model to predict which objects in the sky may not be visible at certain points of Earth's rotation,
B. evaluating an argument based on the evidence or reasoning it includes about changes to the surface of the Earth and life on Earth over time,
C. analyzing patterns in data to determine whether the data support or refute predictions about the kind of weather expected in a region,
D. constructing an argument with evidence about a cause-and-effect relationship between natural and human-designed processes and hazards that occur for humans in some areas.</td>
</tr>
</table>

Exhibit B.2. NAEP Grade 8 Science Achievement Level Descriptions

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
Students performing at this level should be able to demonstrate partial mastery and competency in making sense of common phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. in solids atoms are close together, in liquids atoms are close together but are moving relative to one another, and in gasses the atoms are relatively far apart , B. an object at rest will remain at rest if the net force on it is zero, C. photosynthetic organisms use energy from light to make food, D. the fossil record documents the existence and extinction of many life-forms throughout Earth's history of life, E. the movement of water within the water cycle is a function of phase changes and downhill movement, F. human activities have caused changes in the areas where they live, bringing about major changes in the land, water, and air.	Students performing at this level should be able to demonstrate solid academic performance and competency in making sense of phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. in solids atoms are close together, in liquids atoms are close together but are moving relative to one another and, in gasses the atoms are relatively far apart and move freely and movement of atoms are influenced by temperature and move freely in gasses, B. the change in motion of an object is determined by the sum of the forces acting on it, C. photosynthetic organisms use energy from light, carbon dioxide and water to make sugars and release oxygen, D. the fossil record documents the existence, diversity, extinction, and change of many life-forms throughout the history of the Earth, supporting inferences of lines of evolutionary descent, E. the movement of water within the water cycle is a function of phase changes and is driven by gravity,	Students performing at this level should be able to demonstrate superior performance and competency in making sense of complex phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. in solids atoms are close together vibrating in place, in liquids atoms are close together but are moving relative to one another, and in gasses, the atoms are relatively far apart and move rapidly and freely and the higher the temperature, the more rapid the movement, B. the greater the mass of the object, the greater the force needed to achieve the same change in motion, C. through a series of chemical reactions that release energy photosynthetic organisms use energy from light, carbon dioxide, and water to form new molecules, sugars to support growth and oxygen which is released to the atmosphere, D. the changes in life forms documented through the fossil record can be used to reconstruct an evolutionary history for organisms. E. the movement of water within the water cycle is driven by gravity and

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
	F. human activities have significantly altered the biosphere, atmosphere, and geosphere, sometimes causing the extinction of many organisms.	energy from the sun and continually cycles, changing the land on and below the Earth's surface, F. human activity can significantly alter the biosphere, atmosphere, and geosphere, and human choices can increase or decrease harm to organisms and the environment.
Working with the disciplinary concepts, students require a well-defined set of actions to be able to apply science and engineering practices and crosscutting concepts such as: A. asking questions that arise from observations of phenomena, B. using a model to describe a phenomenon, C. evaluating whether a simple experimental design would meet the goals of an investigation, D. identifying evidence to support an argument for a proposed model or explanation of a phenomenon, E. applying simple mathematical concepts (such as basic operations and simple computations) to scientific questions or designed problems, F. using graphical displays of data to identify relationships between variables,	Working with the disciplinary concepts, students require some cueing to be able to apply science and engineering practices and crosscutting concepts such as: A. asking questions to clarify or refine the explanation for a phenomenon, B. developing and using a model to explain a phenomenon by identifying relationships among parts and or quantities in a system, C. planning an experimental design to produce data that can be used as evidence that meets the goals of an investigation, D. using evidence to support an argument for a proposed model or explanation of a phenomenon, E. applying simple mathematical concepts (such as ratios or proportional thinking) to scientific questions or designed problems, F. constructing graphical displays to identify relationships (linear vs. nonlinear) between variables,	Working with the disciplinary concepts, students require limited cueing to be able to apply science and engineering practices and crosscutting concepts such as: A. asking questions that can be answered with empirical evidence to refine an explanation of cause-and-effect relationships in phenomena, B. revising a model to explain a phenomenon by identifying relationships among parts and or quantities in a system, C. revising an experimental design to help ensure it produces data that can be used as evidence that meets the goals of an investigation, D. revising an argument with evidence for a proposed model or explanation of a phenomenon, E. applying mathematical concepts (such as ratios, rates, or percent) to scientific questions or designed problems, F. evaluating the limitation of data

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
G. describing a solution to a design problem while considering prioritized criteria, H. identifying flaws in a science-related argument in text (poor assumptions).	G. evaluating a solution to a design problem while considering criteria and constraints, H. identifying and critiquing flaws in science-related arguments in text (faulty explanations or overgeneralizations from limited data).	presented in graphical displays of data to identify causal versus correlational relationships between variables, G. evaluating the merits of a solution to a design problem using evidence while considering criteria and constraints, H. identifying and critiquing flaws in science-related arguments in text (cause vs. correlation).
In **Physical Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in *simple sensemaking* of common phenomena when provided *a well-defined set of actions* to perform tasks such as: A. using a simple model to describe patterns associated with the position and movement of atoms relative to one another in solids, liquids, or gasses. B. evaluating whether a simple experimental design provides evidence that the net force on an object is zero when an object is at rest, C. using graphical displays of data on a pair of interacting objects to identify the relationship between the force exerted by the first object on the	In **Physical Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *moderate degree of sensemaking* of phenomena when provided *some cues* to perform tasks such as: A. developing a model to describe the patterns associated with the position and motion of atoms relative to one another in solids, liquids, or gasses and how temperature influences the position and movement of atoms, B. planning an experimental design to produce data that can be used as evidence that the change in motion of an object is determined by the sum of the forces on it, C. constructing graphical displays of data to showcase a relationship between the forces between any pair of interacting objects,	In **Physical Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *high degree of sensemaking* of complex phenomena when provided *limited cues* to perform tasks such as: A. revising a model-based explanation to describe the patterns associated with the position and motion of atoms relative to one another in solids, liquids, or gasses, including the cause-and-effect relationship between temperature and states of matter, B. revising an experimental design to produce data that can serve as evidence that the greater the mass of the object, the greater the force needed to achieve the same change in motion,

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
second object and the force exerted by the second object on the first object, D. asking questions based on observations for how the material an object is made of influences the reflection or transmission of light shining on the object.	D. asking questions to refine an explanation for how the material an object is made of influences whether light shining on the object is reflected or transmitted.	C. constructing graphical displays of data to showcase the causal or correlational relationship between changes in motion of any pair of interacting objects due to the fact the force exerted by the first object on the second and that exerted by the second on the first are second object are equal in strength but in the opposite direction, D. asking questions that can elicit empirical evidence to refine an explanation for how the material an object is made of, or the frequency (color) of the light, influences the reflection, absorption, or transmission of the light shining on the object.

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in *simple sensemaking* of common phenomena when provided *a well-defined set of actions* to perform tasks such as: A. identifying weak assumptions in an argument about the species involved in predatory interactions in an ecosystem, B. using a simple model to explain that matter is transferred between producers, consumers, and decomposers, C. identifying evidence to support an argument about how variations in inherited traits between parent and offspring arise from the subset of genes inherited, D. using evidence to support an explanation about the changes a population undergoes over time in response to a change in physical components of an ecosystem.	In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *moderate degree of sensemaking* of phenomena when provided *some cues* to perform tasks such as: A. identifying and critiquing faulty explanations or overgeneralizations from limited data in an argument about the impact of predatory actions on a population of organisms, B. developing and using a model to explain that matter and energy are transferred between producers, consumers, and decomposers, C. revising an argument using new evidence about how variations in inherited traits between parent and offspring arise from the subset of genes inherited, D. constructing an explanation that uses a chain of cause-and-effect associations between the changes a population undergoes over time and changes in physical or biological components of that population's ecosystem.	In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *high degree of sensemaking* of complex phenomena when provided *limited cues* to perform tasks such as: A. identifying and critiquing flaws in scientific arguments about patterns of predatory interactions across various ecosystems, B. revising a proposed model to explain that the atoms that make up the organisms in an ecosystem are cycled repeatedly between the living and nonliving parts of the ecosystem, C. revising an argument using new evidence to support or reject an explanation that genetic mutations may result in changes in the structure and function of the proteins encoded by genes, D. revising an explanation that uses a chain of cause-and-effect associations between the changes a population undergoes over time in response to changes in physical or biological components of an ecosystem, therefore helping or hurting the health of the ecosystem, including biodiversity.

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
In **Earth and Space Sciences**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in *simple sensemaking* of common phenomena when provided *a well-defined set of actions* to perform tasks such as: A. using a simple model to describe an observable pattern in the motion of an object in the sky relative to Earth, B. making a claim about the relative timing of major events in Earth's history based on the sequence of rock strata, C. interpreting graphical displays of data to identify a relationship between the sunlight, the ocean, and the weather patterns in a given location, D. identifying evidence that predicts future natural hazards through the patterns that precede those hazardous events.	In **Earth and Space Sciences**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *moderate degree of sensemaking* of phenomena when provided *some cues* to perform tasks such as: A. developing and using a model to explain the observable patterns in the motion of objects in the sky relative to Earth, B. identifying evidence that can be used to refute a claim about the relative timing of major events in Earth's history based on the sequence of rock strata and fossil records, C. constructing graphical displays of data to describe relationships between sunlight, the ocean, the atmosphere, ice, or landforms and the weather patterns in a given location, D. constructing an argument using evidence for how observable phenomena that precede the occurrence of some natural hazards can help forecast future events.	In **Earth and Space Sciences**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *high degree of sensemaking* of complex phenomena when provided *limited cues* to perform tasks such as: A. revising a model based on observable patterns in the motion of objects in the sky relative to Earth to make predictions about the future motion or positions of objects in the sky, B. evaluating evidence used to refute a claim about the relative timing of major events in Earth's history based on fossil records and rock types, C. constructing graphical displays of data to explain relationships between sunlight, the ocean, the atmosphere, ice, and landforms and the weather patterns in a given location, D. evaluating arguments about minimizing the risk from hazards using evidence from observable phenomena that precede the occurrence of some natural hazards.

Exhibit B.3. NAEP Grade 12 Science Achievement Level Descriptions

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
Students performing at this level should be able to demonstrate partial mastery and competency in making sense of common phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. all matter is made of atoms, B. the mass and speed of a moving object determine its momentum, C. photosynthesis converts light energy to stored chemical energy, D. DNA sequences vary among species, but there are many sequence similarities, E. the decay of radioactive isotopes in rocks provides a way to date rock formations, F. water's unique properties include expanding upon freezing, G. humans can mitigate the negative impacts on the environment resulting from the use of Earth's resources by applying engineering solutions.	Students performing at this level should be able to demonstrate solid academic performance and competency in making sense of phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. all matter is made of atoms that contain protons, neutrons, and electrons, B. momentum is always conserved, C. photosynthesis converts light energy to stored chemical energy by converting carbon dioxide plus water into sugars plus released oxygen, D. DNA sequences vary among species, but there are many overlaps, providing evidence of evolution, E. the decay of radioactive isotopes in rocks from Earth, moon rocks, and meteorites provides a way to date rock formations that can be used as evidence for Earth's early history, F. water's unique properties include expanding upon freezing, dissolving and transporting solid material, and separating different chemical elements, G. humans can mitigate the negative impacts on the environment resulting from the use of Earth's resources and	Students performing at this level should be able to demonstrate superior performance and competency in making sense of complex phenomena or designing solutions using science and engineering practices and/or crosscutting concepts together with disciplinary concepts such as: A. the electrostatic forces between subatomic particles explain both the structure of isolated atoms and why atoms combine to form molecules, compounds, and extended materials, B. momentum is always conserved because the forces between any two interacting objects are equal and opposite and thus result in equal and opposite changes in momentum, C. photosynthesis converts light energy to stored chemical energy by converting carbon dioxide plus water into sugars, which have more chemical bonds than does carbon dioxide, plus released oxygen, D. evidence of evolution includes overlaps of DNA sequences among species, similarities/differences in amino acid sequences, and anatomical and embryological evidence, E. the decay of radioactive isotopes in rocks from Earth, moon rocks, and

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
	waste disposal by applying engineering and design solutions.	meteorites provides a way to date rock formations that can be used as evidence for Earth's formation and early history, F. interactions between the atmosphere, hydrosphere, and geosphere are influenced by water's unique properties, including expanding upon freezing, dissolving and transporting solid materials, and separating different chemical elements, G. humans can mitigate negative impacts on the environment from resource extraction and waste disposal by applying engineering and design solutions, but when the sources of such problems are not well understood, some actions could magnify the problems.
Working with the disciplinary concepts, students require a well-defined set of actions to be able to apply science and engineering practices and crosscutting concepts such as: A. asking a question that arises from examining a model to clarify the model, B. using a simple model of a system that includes a mathematical relationship to describe phenomena, C. planning an investigation that will produce data that can support the	Working with the disciplinary concepts, students require some cueing to be able to apply science and engineering practices and crosscutting concepts such as: A. asking a question that arises from examining a model to identify additional needed information, B. developing a simple model of a system that includes scale, proportion, and other mathematical relationships to describe phenomena,	Working with the disciplinary concepts, students require limited cueing to be able to apply science and engineering practices and crosscutting concepts such as: A. asking multiple questions that arise from examining a model to identify all needed additional information, B. revising and using a model of a system that includes scale, proportion, and other mathematical relationships to explain phenomena,

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
scientific explanation of a phenomenon, D. using evidence to support an argument about a proposed explanation for structure-function relationships in a system, E. applying simple statistical reasoning to represent design problems, F. analyzing data to identify evidence that could support a model, G. identifying an explanation of a phenomenon that uses a chain of evidence-based associations between variables, H. identifying a flaw in a science-related argument.	C. planning an investigation that considers the appropriate variables to control to produce data that can be used as evidence for cause-and-effect relationships in a phenomenon, D. constructing an argument with evidence and scientific reasoning to support a proposed explanation for structure-function relationships in a system, E. applying statistical reasoning to solve design problems, F. analyzing patterns in data to provide evidence to support or reject a model, G. constructing an explanation of a phenomenon that uses a chain of evidence-based associations between variables, H. identifying and critiquing a flaw related to overgeneralization from limited data in a science-related argument.	C. evaluating the design of an investigation intended to produce data that can be used as evidence for cause-and-effect relationships in a phenomenon, considering possible confounding variables, D. revising an argument to support or reject a proposed explanation for structure-function relationships in a system, addressing new evidence, E. applying statistical reasoning in the context of complicated measurement problems to represent and solve design problems, F. analyzing patterns in data to provide evidence of cause-and-effect relationships that could support or reject a model, G. revising an explanation of a phenomenon that uses a chain of cause-and-effect associations between factors to account for relationships between variables, H. critiquing science-related arguments by identifying multiple flaws, including flaws related to overgeneralizations from limited data.

<table>
<tr><th>NAEP Basic Level</th><th>NAEP Proficient Level</th><th>NAEP Advanced Level</th></tr>
<tr>
<td>In Physical Science, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in simple sensemaking of common phenomena when provided a well-defined set of actions to perform tasks such as:
A. evaluating whether data could support a relationship between temperature and pressure in a gas, at fixed volume,
B. identifying relevant and irrelevant variables in the design of an investigation about the relationship between the forces acting on an object and the change in motion of the object,
C. identifying scientific questions that arise from examining an explanation of the relationship between the energy within a system and the motion and interactions of matter and radiation within that system,
D. using models to compare visible light with X-rays.</td>
<td>In Physical Science, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the moderate degree of sensemaking of phenomena when provided some cues to perform tasks such as:
A. evaluating whether data are sufficient in quantity to support a qualitative statement of the relationship between temperature and pressure of a gas at fixed volume,
B. planning an investigation that will produce data to serve as the basis for evidence about the relationship between the sum of two aligned forces acting on an object and the change in motion of the object,
C. asking questions that arise from examining an explanation related to the way the energy available within a system depends on the motion and interactions of matter and radiation within that system,
D. developing models of the mathematical relationships between visible light and X-rays.</td>
<td>In Physical Science, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the high degree of sensemaking of complex phenomena when provided limited cues to perform tasks such as:
A. evaluating whether data are sufficient in quantity, accuracy, and reliability to support or reject a claim that there is a relationship between temperature and pressure or between temperature and the rate of chemical reactions in gasses, at fixed volume,
B. planning an investigation to produce data that can serve as evidence for an explanation of the relationship between the relative magnitudes of two aligned forces acting on an object and the change in motion of the object, considering possible confounding variables,
C. evaluating questions that arise from examining a model that illustrates that the quantity of energy available for processes within a system depends on the motion and interactions of matter and radiation within that system,
D. revising models of the mathematical relationships between visible light, X-rays, and radio waves.</td>
</tr>
</table>

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in *simple sensemaking* of common phenomena when provided *a well-defined set of actions* to perform tasks such as: A. describing the merits of a simulation of an ecosystem that illustrates that not all of the matter consumed at one level of a food web is transferred to other levels in a food web, B. revising an argument that genes code for the formation of proteins that determine traits, C. evaluating whether the quantity of the data are sufficient to support an explanation about whether natural selection can result from competition for resources, D. identifying the merits and limitations of a model of cell division.	In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *moderate degree of sensemaking* of phenomena when provided *some cues* to perform tasks such as: A. evaluating a simulation of an ecosystem that illustrates that only a small fraction of the matter consumed at lower levels of a food web is transferred to upper levels in a food web, B. revising an argument that genes code for the formation of proteins, which determine traits, addressing new evidence about DNA sequences that do not code for a protein, C. evaluating whether the quantity and accuracy of the data used as evidence are sufficient to support an explanation about the four factors that can cause natural selection, D. evaluating the merits and limitations of two different models of the reproduction of genetic information in mitosis to select a model that best fits the evidence.	In **Life Science**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *high degree of sensemaking* of complex phenomena when provided *limited cues* to perform tasks such as: A. evaluating simulations of an ecosystem by comparing the outcomes of the simulation with what is known about the problems in the ecosystem that result from the fact that only a small fraction of the matter consumed at lower trophic levels is transferred to upper trophic levels in a food web to produce growth and release energy at the higher level, B. revising an argument to support or reject an explanation about the various functions of genes and whether all genes code for the formation of proteins that determine traits, C. evaluating whether the quantity, accuracy, and reliability of the data used as evidence are sufficient to support an explanation about the interaction of the four factors that can cause natural selection, D. evaluating the merits and limitations of two different models of the role of mitosis in the growth of an organism

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
		to revise the model that best fits the evidence.
In **Earth and Space Science**s, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in *simple sensemaking* of common phenomena when provided *a well-defined set of actions* to perform tasks such as: A. analyzing data to identify evidence that changes in the orientation of Earth's axis of rotation affects the amount of sunlight falling on the planet, B. identifying evidence that the decay of radioactive isotopes in rocks provides a way to date rock formations, C. using models to describe ways that ocean and atmospheric circulations influence climate, D. identifying evidence that the size of a human population has been affected by a natural hazard.	In **Earth and Space Sciences**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *moderate degree of sensemaking* of phenomena when provided *some cues* to perform tasks such as: A. analyzing data to support an explanation for changes in the shape of Earth's orbit and the orientation of its axis of rotation driving changes in the amount of sunlight falling on the planet, B. evaluating evidence that uses measurements of the decay of radioactive elements in minerals and rocks to provide evidence of Earth's early history, C. using system models that include mathematical relationships to describe how ocean and atmospheric circulations influence temporal and spatial patterns on Earth's climate, D. developing an argument that the size and location of a human population has been affected by natural hazards.	In **Earth and Space Sciences**, students should be able to integrate disciplinary concepts, science and engineering practices, and crosscutting concepts to engage in the *high degree of sensemaking* of complex phenomena when provided *limited cues* to perform tasks such as: A. analyzing global temperature data to reveal patterns that can support an explanation for changes in the shape of Earth's orbit and the orientation of its axis of rotation altering the intensity and distribution of sunlight falling on the planet, B. evaluating evidence from multiple sources using measurements of the decay of radioactive elements in minerals and rocks to provide evidence for Earth's formation and early history, C. developing system models that employ mathematical relationships to describe ways that the absorption, reflection, storage, and redistribution of energy from the sun lead to temporal and spatial patterns in Earth's climate system, D. evaluating an argument that feedback mechanisms could magnify or mitigate the intensity of the effect

NAEP Basic Level	NAEP Proficient Level	NAEP Advanced Level
		on a human population by a natural hazard.

APPENDIX C: Glossary

The following terms are used in the NAEP Science Assessment Framework and the NAEP Science Assessment and Item Specifications. Additional terms may be found in the *NAEP Glossary of Terms*.

achievement level descriptions (ALDs): Descriptions of students' performance at achievement levels (*NAEP Basic*, *NAEP Proficient*, and *NAEP Advanced*), detailing what students should know and be able to do in terms of the science disciplinary concepts, science and engineering practices, and crosscutting concepts on the NAEP assessment.

alignment: The coordination of goals, instruction, and assessment in a mutually reinforcing educational system.

argument: A process of evaluating competing claims, models, or explanations of a phenomenon based on available evidence (in science) or to a process of evaluating prospective designs based on specifications and constraints (in engineering).

construct: An image, idea, or theory, especially a complex one formed from a number of simpler elements and often embedded in a web of related ideas.

constructed response (CR): An open-ended, text-based response. Every constructed response item has a scoring guide that defines the criteria used to evaluate students' responses.

Constructed response item types that may be used on the 2028 NAEP Science Assessment are listed below.

- **Short constructed response** – Students respond by giving a short response, from a single word or number to a few sentences.
- **Extended constructed response** – Students respond by giving a description or explanation that requires more than a few words.

context: All the information presented to a student in framing a task and the prompt that elicits a student response. The same phenomenon or problem can be addressed through many different contexts and thus can frame many tasks. All stimulus information provided to students (e.g., written descriptions, images, videos, simulations, long-form texts, infographics, data tables, graphs, etc.) used to present the phenomenon is considered context, offering background information necessary for students' sensemaking.

contextual variables: Students, teacher, administrator, and school factors that shape students' opportunities to learn, including instructional time, content, strategies, and resources.

crosscutting concepts: Ideas that transcend disciplinary boundaries and prove fruitful in explanation, in theory, in observation, and in design. These ideas are conceptual tools that guide effective and reflective practice in all fields of science and engineering.

digital tools: Any technology that stores and transmits data electronically. In a digitally based environment, students require tools to make and enter calculations, build models, run simulations, and to create and modify graphical representations of data.

dimensions: Three broad sets of expectations with respect to a student's knowledge and skills: Science and Engineering Practices (SEPs), Disciplinary Concepts (DCs), and Crosscutting Concepts (CCCs).

disciplinary concepts (DCs): Well-tested theories and explanations developed by scientists organized into three major disciplinary groupings: physical science, life science, and Earth and space sciences.

discrete item (DI): A single, standalone item.

English Learners: Active learners of the English language who may benefit from various types of language support programs; students from a diverse set of backgrounds who often come from non-English-speaking homes and backgrounds and who typically require specialized or modified instruction in both the English language and in their academic courses.

engineering: A discipline involved in the definition and solution of problems. Engineering often requires development of a design to solve the problem that meets the criteria for a successful solution within constraints such as time and budget. The term *engineering* includes many areas of application (e.g., medicine, agriculture, infrastructure, environmental management).

evidence: A body of facts or observations that can provide information about whether a belief or proposition is true or valid.

exhibit: Tables and figures in this framework. Exhibits are numbered consecutively within each chapter. For example, the first three exhibits in Chapter 3 are labeled Exhibit 3.1, Exhibit 3.2, and Exhibit 3.3.

item: The questions students answer, or the tasks they must complete, as part of an educational assessment.

item part: The smallest element requiring a response within an item. For example, a two-part item might consist of a selected response item part followed by a constructed response item part that asks the student to explain the answer chosen in the selected response item part.

item set: A group of independent items that uses common stimulus material. Item sets make it possible to take advantage of efficiency by presenting rich and engaging stimulus material, then asking several questions to collect evidence. Since the items do not depend on each other, questions in an item set each receive a separate score. Item sets should include at least one item that is three-dimensional.

item subtype: A specific format available within an item type (e.g., multiple choice and multiple select are subtypes of the selected response item type).

item type: A description of the format of an assessment item. Item types may be categorized by their overall structure and complexity, such as discrete, multipart, item set, and scenario-based task. Items may also be categorized by the kind of response required, such as selected response, constructed response, and technology enhanced.

multidimensional: Items that integrate two or all three dimensions.

multipart item (MPI): An item that includes multiple parts that are dependent on each other. For example, a multipart item might ask students to make a choice or decision and follow up with another question to explain their reasoning.

opportunity to learn: Inputs and processes that enable student achievement of intended outcomes.

phenomena: Observable events that occur in either natural or human-designed systems.

problem: A challenge that arises from a human need or want. In the 2028 NAEP Science Framework, the term *problem* is used to describe a real-world issue that requires a designed solution; as such, it is an engineering problem.

process data: Information collected as students navigate the digital assessment, including the time taken to engage in item stem or stimulus and respond to questions, how often they return to the stem or stimulus to answer questions, and their use of digital tools. Process data reflect interactions in which the student engages and may provide relevant evidence about whether the student possesses a skill that is an assessment target. Thus, process data can be captured, measured, and interpreted to generate a score.

scenario-based task (SBT): A sequence of items presented through an unfolding context, often with rich and engaging stimulus material such as images and video. SBTs are often interactive, asking students to respond to several short tasks and questions. However, the task does not have to be interactive to be a scenario-based task. SBTs typically present meaningful and compelling phenomena and problems, including those that require a large amount of background information. Scenario-based tasks should include at least one item that is three-dimensional.

science achievement expectation: An assessable statement of what students should know and be able to do. Formulating a multidimensional science achievement expectation—expressed as a performance—is often the starting point in developing an assessment item.

science and engineering practices: Ways of working to develop scientific explanations of phenomena or design engineering solutions to problems.

selected response (SR): Assessment responses that involve a student selecting one or more response options from a given, limited set of choices.

Different types of selected response items that may be used on the 2028 NAEP Science Assessment include the following.

- **Grid** – Students evaluate statements, such as claims or explanations, or classify components of a system based on their properties or interactions. The answer is entered by selecting cells in a table.
- **Inline choice** – Students respond by selecting one option from one or more drop-down menus that may appear in various sections of an item.
- **Matching** – Students respond by inserting (i.e., dragging and dropping) one or more source elements (e.g., an image) into target fields (e.g., a table).
- **Multiple-selection multiple choice** – Students respond by selecting two or more choices that meet the condition stated in the stem of the item.
- **Single-selection multiple choice** – Students respond by selecting a single choice from a set of given choices.
- **Zones** – Students respond by selecting one or more regions on a graphic stimulus.

sensemaking: Actively applying disciplinary concepts, science and engineering practices, and crosscutting concepts to figure out a phenomenon or address a real-world problem. The degree and nature of sensemaking required by students is determined by the complexity of each item, and the degree to which each dimension contributes to that complexity in each item.

sophisticated/sophistication: For the purpose of this framework, increasing sophistication refers to a student expression of understanding that is more thorough, more precise, more accurate, and more coherent throughout.

stem: The item question or prompt to which the student responds.

stimulus: Written descriptions, images, videos, simulations, long-form texts, infographics, data tables, graphs, and all other information provided to students in a NAEP test question. Item sets and scenario-based tasks include stimuli that are shared among multiple items.

target: Assessable knowledge and skills. For an item or item part in an item set, the target consists of the evidence statements and associated parts of the dimensions included in the evidence statement.

APPENDIX D: Full Description of Exhibit 3.4

Summary: Exhibit 3.4 is an infographic titled 'Visualizing the Sensemaking Process'

Description: Five graphics depict different stages in student sensemaking, accompanied by descriptions of each step. Each image includes the same four symbols, which are shown being more closely connected together with each progressive stage.

- **Stage 1**: Four symbols shown inside of a circle. Related text: 'Each dot represents ideas, abilities, and experiences that students have developed. These can include disciplinary concepts, crosscutting concepts, and practices as well as explanations for particular kinds of phenomena or solutions to problems. While sensemaking may have been involved in developing these understandings, students no longer need to figure them out to respond to relevant questions or situations. Importantly, students have not connected these ideas.'
- **Stage 2**: The same four symbols inside of a circle, with a question mark centered between them. Related text: 'An assessment item presents a phenomenon that the student cannot immediately explain, or a problem that the student cannot solve right away. Uncertainty activates prior ideas, abilities, and experiences. This may happen through cuing, scaffolding, or because of connections students make themselves.'
- **Stage 3**: The same four symbols inside of a circle with the question mark centered between them, but with arrows from each symbol point to the question mark, and dotted lines connect the symbols to each other. Related text: 'Sensemaking occurs when students connect their previously developed ideas, abilities, and experiences to address the unexplained phenomenon or unsolved problem. This sensemaking can be heavily or lightly scaffolded, or collaboratively pursued, as long as students are responsible for putting the pieces together.'
- **Stage 4**: The same four symbols inside of a circle, now connected by solid lines to an unlocked padlock in the center. A solid line connects the symbols to each other. Related text: 'If students are asked to address the same (or similar) phenomena and problems repeatedly, over time the connections between ideas require less figuring out and become increasingly easier for students to make.'
- **Stage 5**: The same four symbols inside of a circle, now connected by solid lines to a star in the center. A solid line connects the symbols to each other. The space between these connecting lines is now shaded in. Related text: 'Eventually, the whole experience becomes one that students can leverage as a schema or an understanding that they have figured out and can connect with others to explain new phenomena or solve new problems.'

REFERENCES

Achieve (2019a). *Task Annotation Project in Science: Sense-making*. https://issuu.com/achieveinc/docs/sense-making_02142019__7_

Achieve (2019b). *A framework to evaluate cognitive complexity in science assessments*. https://www.nextgenscience.org/sites/default/files/Science%20Cognitive%20Complexity%20Framework_Final_093019.pdf

American Association for the Advancement of Science (1990). *Science for all Americans: A Project 2061 report on literacy goals in science, mathematics, and technology*.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Joint Committee on Standards for Educational and Psychological Testing.

Anderson, C. W., de los Santos, E. X., Bodbyl, S., Covitt, B., Edwards, K. D., & Hancock, J. (2018). Designing educational systems to support enactment of the Next Generation Science Standards. *Journal of Research in Science Teaching, 55*(7), 1026–1052.

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research.

Brown, B. A. (2019). *Science in the city: Culturally relevant STEM education.* Harvard Education Press.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, *64*(8), 1–9.

Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, *18*(1), 26–31.

Davis, E. A., & Stephens, A. (2022). *Science and engineering in preschool through elementary grades: The brilliance of children and the strengths of educators—a consensus study report*. National Academies Press.

Ferguson, H. B., Bovaird, S., & Mueller, M. P. (2007). The impact of poverty on educational outcomes for children. *Paediatrics & Child Health*, *12*(8), 701–706.

Joint Task Force on Assessment of the International Reading Association and the National Council of Teachers of English. (2010). *Standards for the assessment of reading and writing* (Rev. ed.). NCTE; IRA.

Kohlhaas, K., Lin, H. H., & Chu, K. L. (2010). Disaggregated outcomes of gender, ethnicity, and poverty on fifth grade science performance. *RMLE Online*, *33*(7), 1–12.

National Academies of Sciences, Engineering, and Medicine. (2017). *Seeing students learn science: Integrating assessment and instruction in the classroom.* National Academies Press. https://doi.org/10.17226/23548

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

National Academies of Sciences, Engineering, and Medicine. (2019). *Science and engineering for grades 6–12: Investigation and design at the center*. National Academies Press.

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

National Assessment Governing Board. (2018). Technology and engineering literacy framework for the 2018 National Assessment of Educational Progress. https://files.eric.ed.gov/fulltext/ED594359.pdf

National Assessment Governing Board. (2019). Science framework for the 2019 National Assessment of Educational Progress. https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/science/2019-science-framework.pdf

National Assessment Governing Board. (2022). NAEP assessment tutorials. Retrieved from https://nces.ed.gov/nationsreportcard/experience

National Research Council. (1996). *National Science Education Standards.* National Academies Press.

National Research Council. (2005). *How students learn: Science in the classroom.* National Academies Press. https://doi.org/10.17226/11102

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K–8*. National Academies Press. https://doi.org/10.17226/11625

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts.* National Academies Press.

National Research Council. (2015). *Guide to implementing the next generation science standards*. Washington, DC: National Academies Press.

Penuel, W. R., Harris, C. J., & DeBarger, A. H. (2015). Implementing the next generation science standards. *Phi Delta Kappan, 96*(6), 45–49.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). Academic Press.

Stembridge, A. (2019). *Culturally responsive education in the classroom: An equity framework for pedagogy*. Routledge.

Tekkumru-Kisa, M., Stein, M. K., & Schunn, C. (2015). A framework for analyzing cognitive demand and content-practices integration: Task analysis guide in science. *Journal of Research in Science Teaching*, *52*(5), 659–685.

WestEd, Center on Standards and Assessment Implementation, & Delaware Department of Education. (2019). *Cognitive loading in three-dimensional NGSS assessment: Knowledge, skills, and know-how*. https://csaa.wested.org/resource/cognitive-loading-in-three-dimensional-ngss-assessment-knowledge-skills-and-know-how/