Science Assessment and Item Specifications for the 2009 National Assessment of Educational Progress

SCIENCE

NAEP

2009

Developed by WestEd
and the Council of Chief State School Officers
under contract to the National Assessment Governing Board
Contract # ED04CO0148

Spring 2007

Table of Contents

List of Figures and Tables

Figures

Tables

PREFACE

The *Science Assessment and Item Specifications for the 2009 NAEP* (the *Specifications*) translates the *Science Framework for the 2009 NAEP* (the *Framework*) into guidelines for developing items and for developing the assessment as a whole. The primary purpose of the *Specifications* is to provide the National Center for Education Statistics (NCES) and the NAEP assessment development contractor with information to ensure that the NAEP Science Assessment reflects the intent of the National Assessment Governing Board's (NAGB) science framework development process.

The *Specifications* are structured so that the assessment developer and item writers have a single document to which to refer when they are working on the assessment. Therefore, portions of the *Framework* that specify characteristics of the assessment or items are repeated in the *Specifications*.

Portions of the *Specifications* are drawn from and informed by the 1996-2005 *Science Assessment and Exercise Specifications for the National Assessment of Educational Progress* and the *2005 NAEP Mathematics Assessment and Item Specifications.*

This preface provides a brief overview of the subsequent chapters.

Chapter One, Overview, provides background information on the new *Framework* and *Specifications,* including need and context. There is an introduction to the *Specifications* and a description of the document's development. The 1996-2005 and 2009 NAEP Science Specifications are compared. The chapter ends with guidelines for aligning the assessment with the *Framework* and *Specifications*.

Chapter Two, **Science Content,** presents the content for the 2009 NAEP Science Assessment. First is a discussion of how the content statements were developed, followed by how they are organized. Key components of this chapter include overviews of each science content area—Physical Science, Life Science, and Earth and Space Science—and content statements. These content statements define key science principles along with the facts, concepts, laws, and theories to be assessed. This chapter also provides content boundaries, elaborations on the content statements that delineate what is and is not appropriate to assess for a given subtopic at grades 4, 8, and 12. The chapter concludes with a discussion of content that cuts across the science content areas and the suggested distribution of items by content area for the NAEP assessment.

Chapter Three, Science Practices, defines what students should be able to do with the science content presented in the previous chapter. Key science practices are Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design. This chapter also presents four cognitive demands with which the science practices can be associated. Discussed in detail are "knowing that," "knowing how," "knowing why," and "knowing when and where to apply knowledge." The chapter ends with a suggested distribution of items by science practice.

Chapter Four, Generating and Interpreting Items, provides examples of performance expectations for grades 4, 8, and 12. Assessment items stem from performance expectations, the intersection of content statements and science practices. A complete illustration includes the content statement itself, commentary on the content, performance expectations, and sample items. The chapter concludes with a discussion of learning progressions, sequences of successively more complex ways of reasoning about a set of ideas.

Chapter Five, Types of Items, focuses on the nature of assessment items to be found on the NAEP Science Assessment. Topics include principles of good item writing such as clear intent of an item and the use of clear and concise language. The types of items to be used on the assessment include selected-response (multiple-choice) items and constructed-response items (which include short and extended constructed-response items as well as concept-mapping tasks). Some "combination items" may require more than one response. Such combination items include item clusters, Predict-Observe-Explain (POE) item sets, hands-on performance tasks, and interactive computer tasks. The responses requested may be all selected-response, all constructed-response, or a mixture. These different types of items are discussed in detail, including scoring; and numerous illustrations are provided. The chapter ends with a suggested distribution of items by item type and a discussion of how item development and review should be carried out.

Chapter Six, Administration of the Assessment, presents guidelines for selecting students for participation, assessment administration, reporting requirements, and achievement levels. Special emphasis is on making NAEP assessments as accessible as possible through either construction of the assessment itself or accommodations for students with special needs. Special needs students are English language learners and students with disabilities. This chapter also discusses the uses of NAEP data and the challenges of developing a NAEP assessment. It concludes with summaries of the special studies recommended by the project's Steering Committee and Planning Committee.

The extensive **Appendices** include lists of members of the Steering Committee and Planning Committee, the Steering Committee's guidelines for development of the *Framework* and *Specifications*, NAEP Science preliminary achievement level descriptions, sample items and scoring guides, additional special studies, examples of generating and interpreting items, detail on learning progressions, and the NAEP Item Development and Review Policy Statement.

CHAPTER ONE: OVERVIEW

Introduction

Science seeks to increase our understanding of the natural world through empirical evidence. Such evidence gathered through observation and measurement allows explanation and prediction of natural phenomena. Hence, a scientifically literate person is familiar with the natural world and understands key facts, concepts, principles, laws, and theories of science, such as the motion of objects, the function of cells in living organisms, and the properties of Earth materials. Further, a scientifically literate person can connect ideas across disciplines, for example, the conservation of energy in physical, life, Earth, and space systems. Scientific literacy also encompasses understanding the use of scientific principles and ways of thinking to advance our knowledge of the natural world, as well as the use of science to solve problems in real-world contexts, which this document refers to as "Using Technological Design."

The National Assessment of Educational Progress (NAEP) measures student science achievement nationally, state-by-state, and most recently across selected urban school districts. Periodically, the framework and specifications underlying the science assessment is revised or updated. This document, the *Science Assessment and Item Specifications for the 2009 NAEP* (herein called the *Specifications*), contains a new set of recommendations for the NAEP Science Assessment to be administered in 2009 and beyond. The *Specifications* provides guidance on the science content to be assessed, the types of assessment questions, and the administration of the assessment.

For more than 35 years, NAEP has gathered information on student achievement in selected academic subjects. Originally, assessments were age-based samples of students 9-, 13-, and 17-years old. Beginning in 1983, the assessment also has included grade-based samples of students in grades 4, 8, and 12. Currently, long-term trend NAEP continues to assess 9-, 13-, and 17-year-olds in mathematics and reading, while main NAEP assesses students in grades 4, 8, and 12. For more information about differences between long-term trend and main NAEP, see the following website: http://nces.ed.gov/nationsreportcard/about/ltt_main_diff.asp (National Center for Education Statistics [NCES], 2005b).

NAEP has become an important source of information on what U.S. students know and are able to do in reading, mathematics, science, U.S. history, writing, and other subjects. In addition, NAEP provides information on how student performance has changed over time. Since the 1990s, in addition to the national-level assessments, NAEP has conducted and reported state-level assessments at grades 4 and 8 in reading, mathematics, writing, and science. State-level, as well as national, science assessments were conducted in 1996, 2000, and 2005. The resulting data on student knowledge and performance have been accompanied by background information that allows analyses of a number of student demographic and instructional factors related to achievement. The assessments have been designed to allow comparisons of student performance over time and among subgroups of students according to region, parental education, gender, and race/ethnicity. In 2002, NAEP began a Trial Urban District Assessment (TUDA) in districts that volunteered to participate. The TUDA has continued through 2005 when ten districts took part in NAEP assessments that produced district-level results.

Need for New Framework and Specifications

The framework and specifications documents that guided the last three NAEP Science Assessments (administered in 1996, 2000, and 2005) were developed some 15 years ago. Since then, the following developments have taken place, making it necessary to create a new set of specifications for assessing science in 2009 and beyond:

- Publication, for the first time, of national standards for science literacy in *National Science Education Standards* (National Research Council [NRC], 1996) (herein called *National Standards*) and *Benchmarks for Scientific Literacy* (American Association for the Advancement of Science [AAAS], 1993) (herein called *Benchmarks*). Since their publication, these two national documents have informed state science standards.
- Advances in science research (e.g., on the relationship between human activity and the natural world) that have increased knowledge and as a consequence influenced the school curriculum in the fields of physical, life, and Earth and space sciences.
- Advances in cognitive research (e.g., on how students learn increasingly complex material over time) that have yielded new insights into how and what students learn about science (NRC, 1999c, 2001, 2005b). For example, there is new information about what is appropriate for students to learn at various grades (Catley, Lehrer, & Reiser, 2005; Metz, 1995; Smith, Wiser, Anderson, Krajcik, & Coppola, 2004).
- Growth in the prevalence of science assessments nationally and internationally, including the requirements in the current federal education legislation, *No Child Left Behind*, for science assessment starting in 2007; the ongoing international assessment, *Trends in International Mathematics and Science Study* (TIMSS) (Mullis et al., 2001; see http://timss.bc.edu); and *Programme for International Student Assessment* (PISA) (Organisation for Economic Cooperation and Development, 2005; see www.pisa.oecd.org). These assessments, including frameworks, curricular analyses, and reports on results, are increasing the visibility of and interest in comparing and contrasting student science performances.[1] For example, a recent TIMSS report (Gonzales et al., 2004) lists ten NCES publications related to TIMSS (see http://nces.ed.gov/timss) and 16 additional publications available through Boston College, the TIMSS International Study Center (see http://timss.bc.edu).

[1] An influential component of TIMSS has been a comprehensive and detailed analysis of science curricula in more than 40 countries, including the U.S. (Schmidt, Raizen, Britton, Bianchi, & Wolfe, 1997). A recent comparison of the science content in the NAEP 2000 and TIMSS 2003 assessments found considerable overlap but also key differences between these national and international assessments (Neidorf, Binkley, & Stephens, 2006).

- Growth in innovative assessment approaches that probe students' understanding of science in greater depth than before (e.g., clusters of items tapping students' conceptions of the natural world), sometimes with the use of computer technology (NRC, 2001).
- Increased inclusion of formerly excluded groups in science assessments (e.g., students with disabilities and English language learners), requiring a new assessment to be as accessible as possible and also to incorporate accommodations so that the widest possible range of students can be fairly assessed (e.g., Lee, 1999). Accommodations should not alter the science constructs being measured.

Context for Planning the Framework and Specifications

Any NAEP framework and specifications must be guided by NAEP purposes as well as the policies and procedures of the National Assessment Governing Board (NAGB), which oversees NAEP. For the NAEP Science Assessment, the main purpose of the *Framework* and *Specifications* is to establish what students should know and be able to do in science for the 2009 and future assessments. Meeting this purpose requires recommendations built around what communities involved in science and science education consider as a rigorous body of science knowledge and skills that are most important for NAEP to assess.

In prioritizing the content, the *Framework* and *Specifications* developers used the guidance from the NAEP Science Assessment Steering Committee, which recommended the two national documents, *National Standards* and *Benchmarks*, as representative of the leading science communities and their expectations for what students should know and be able to do in science. As curriculum frameworks, however, these documents cover a very wide range of science content and performance. The inclusive nature of both these documents demonstrates the difficulty of identifying a key body of knowledge for students to learn in science and, therefore, what should be assessed. Neither document limits or prioritizes content as is necessary for developing an assessment, posing a considerable challenge to the *Framework* and *Specifications* developers. The development of the *Framework* and *Specifications* also was informed by research in science and science education, best practices, international assessment frameworks, and state standards.

The Specifications Development Process

In September 2004, NAGB awarded a contract to WestEd and the Council of Chief State School Officers (CCSSO) to develop a recommended *Framework* and *Specifications*. WestEd and CCSSO, in collaboration with the American Association for the Advancement of Science (AAAS), the Council of State Science Supervisors (CSSS), and the National Science Teachers Association (NSTA), used a process designed to accomplish the purposes of this project with special attention given to the assessment issues that are specific to K-12 science achievement. The process for developing the

Framework, Specifications, and related products was inclusive and deliberate, designed to achieve as much broad-based input as possible.

A two-tiered committee structure, consisting of a Steering Committee and a Planning Committee, provided the expertise to develop the *Framework* and *Specifications* as specified by NAGB. (See Appendix A for lists of committee members.) The two committees were composed of members who were diverse in terms of role, gender, race/ethnicity, region of the country, perspective, and expertise regarding the content of the assessment to be developed.

Comprised of leaders in science, science education, general education, assessment, and various public constituencies, the Science Assessment Steering Committee set the course for the project. Functioning as a policy and oversight body, this group developed a charge that outlined what the Planning Committee should attend to in the development of the *Framework* and *Specifications.* (See Appendix B for Steering Committee Guidelines.) The committee also reviewed and provided feedback on drafts of these documents and related materials.

The Science Assessment Planning Committee, supported by the project staff, was the development and production group responsible for drafting the *Framework*, the *Specifications*, recommendations for background variables, designs for one or more small-scale studies, and preliminary science achievement level descriptions (see Appendix C). This committee was made up of science teachers, district and state science personnel, science educators in higher education, scientists, and assessment experts. The Planning Committee's work was guided by policies, goals, and principles identified by the Steering Committee. In addition, the Planning Committee used a number of resources to facilitate their work. These included an *Issues and Recommendations* paper (Champagne, Bergin, Bybee, Duschl, & Gallagher, 2004) developed specifically for this NAEP project; the frameworks and specifications for the 2005 NAEP Mathematics Assessment and 1996-2005 NAEP Science Assessments; other NAEP reports and documents produced by NAGB and NCES; international assessment frameworks; syntheses of state and national curriculum standards; and research papers and resources provided by Steering and Planning Committee members and project staff.

The structure for conducting the work consisted of a series of meetings. From December 2004 through September 2005, the Steering Committee met three times and the Planning Committee met six times; two of the Steering Committee meetings overlapped with Planning Committee meetings. In January 2006, a subgroup of the Planning Committee met to further refine the *Specifications*. NAGB staff supported and participated in the work of the committees during the meetings. Additionally, between formal work sessions, NAGB members and staff provided ongoing feedback and guidance on project documents and processes.

During spring 2005, CCSSO led a series of outreach efforts to solicit feedback on draft versions of the *Framework*. Formal activities included the following:

- A series of 13 regional meetings held across the country and hosted by CCSSO and members of the Council of State Science Supervisors (CSSS)
- A national meeting of CSSS representatives
- A web-based survey of science teachers distributed through the National Science Teachers Association (NSTA)
- An invitational science and industry feedback forum held in Atlanta in conjunction with a NAGB meeting

These activities are discussed in *A Summary of National Feedback Provided on Preliminary Drafts Gathered from Surveys and Regional and National Feedback Meetings* (CCSSO, 2005). Feedback from these sessions has been incorporated into the *Framework* and *Specifications*. Examples include reduction of the number of statements of science content to be assessed; a comparison of the old and new Science Frameworks; and ensuring a high level of consistency in scope, specificity, language, and format across the science content areas.

Other related outreach activities included but were not limited to presentations and sessions held with the American Association for the Advancement of Science (AAAS); the CSSS annual conference; the NSTA national and regional conventions; meetings of the National Research Council (NRC)'s Board on Science Education and Committee on Science Learning K-8; and CCSSO's Mega-SCASS (State Collaborative on Assessment and Student Standards) conference, Large-Scale Assessment conference, and Education Information Management Advisory Consortium (EIMAC). NAGB engaged an external review panel to evaluate the draft *Framework* and convened a public hearing to gather additional input during the development process. Additionally, members of CSSS conducted a review of the draft *Specifications* in January 2006. (See Appendix A for more complete lists of individuals and organizations that contributed to the development of the *Framework* and *Specifications*.) The Planning Committee reviewed feedback from these groups as well as that from the Steering Committee and made changes as it deemed appropriate. After final approval from the Steering Committee, the *Framework,* the *Specifications*, and related products were submitted to NAGB for action. The Governing Board unanimously approved the *Framework* on November 18, 2005 and the *Specifications* on March 3, 2006.

Introduction to the Specifications

Science comprises both content and practices. The NAEP Science Assessment provides a snapshot view of what the nation's 4^{th}, 8^{th}, and 12^{th} graders know and can do in science. One expects students, as a result of their education and life experiences, to have learned about the principles (along with the facts, concepts, laws, and theories) that have been verified by the community of scientists, as well as how scientists discover regularities in the natural world.[2] NAEP will assess students' abilities to identify and use science principles, as well as use scientific inquiry and technological design. (See Chapter Three.) While the *Specifications* distinguishes content from practice, the two are closely linked in assessment as in science itself.

The *Framework* and *Specifications* address scientific knowledge and processes. Science is a way of knowing about the natural world based on tested explanations supported by accumulated empirical evidence. Explanations of natural phenomena that rely on non-scientific views are not reflected in the *Framework* and *Specifications*. The committees responsible for the development of the *Framework* and *Specifications* relied on *National Standards*, *Benchmarks*, international frameworks, and state standards for content about the nature and practice of science. As stated in the *National Standards*,

> Scientific explanations must meet certain criteria. First and foremost, they must be consistent with experimental and observational evidence about nature, and must make accurate predictions, when appropriate, about systems being studied. They should also be logical, respect the rules of evidence, be open to criticism, report methods and procedures, and make knowledge public (p. 201).

[2] "Science principles" is used throughout the *Framework* and *Specifications* to denote not only the principles but also the facts, concepts, laws, and theories of science.

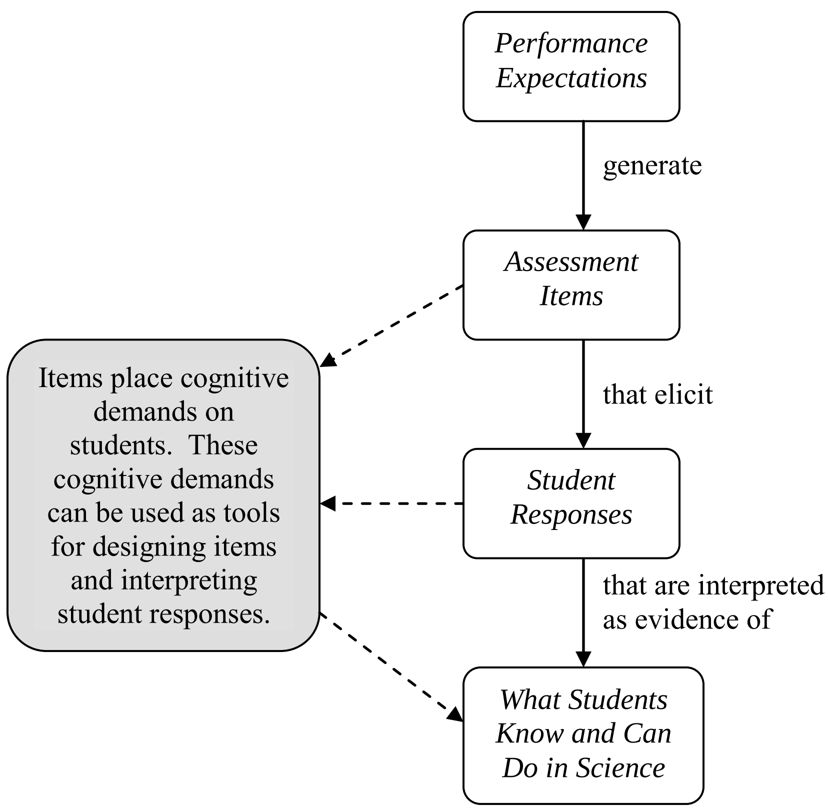
The design of the NAEP Science Assessment is guided by the *Specifications*' descriptions of the science content and practices to be assessed. Figure 1 illustrates how content and practices are combined ("crossed") to generate performance expectations. The columns contain the science content (defined by content statements—propositions that express science principles—in three broad areas), and the rows contain the four science practices. A double dashed line distinguishes Identifying Science Principles and Using Science Principles from Using Scientific Inquiry and Using Technological Design. The former two practices can be generally considered as "knowing science," and the latter two practices can be considered as the application of that knowledge to "doing science" and "using science to solve real-world problems." The cells at the intersection of content (columns) and practices (rows) contain student performance expectations. Since content and practice categories are not entirely distinct, some overlap in the resultant performance expectations is to be expected, as denoted by dashed lines.

Figure 1. Crossing Content and Practices to Generate Performance Expectations

		Science Content		
		Physical Science content statements	**Life Science** content statements	**Earth and Space Science** content statements
Science Practices	**Identifying Science Principles**	*Performance Expectations*	*Performance Expectations*	*Performance Expectations*
	Using Science Principles	*Performance Expectations*	*Performance Expectations*	*Performance Expectations*
	Using Scientific Inquiry	*Performance Expectations*	*Performance Expectations*	*Performance Expectations*
	Using Technological Design	*Performance Expectations*	*Performance Expectations*	*Performance Expectations*

Figure 2 illustrates the fuller process of generating assessment items and interpreting student responses. An item is an individual question or exercise on the NAEP Science Assessment and is used to gather information about students' knowledge and abilities. Figure 2 begins with student performance expectations, which describe in observable terms what students are expected to know and do on the assessment. These performance expectations guide the development of assessment items. Sample assessment items and scoring rubrics are provided throughout this document and in Appendix D. The cognitive demands (see Chapter Three) of the items can then be used to interpret students' responses as evidence of what students know and can do in science (see Ruiz-Primo, Shavelson, Li, and Schultz [2001] for more on cognitive interpretations of performances on assessment tasks). Figure 2 suggests a linear process, but the development of an assessment is iterative (e.g., assessment items are modified based on student responses provided on trials of pilot versions).

Figure 2. Generating Items and Interpreting Responses



The Science Assessment at a Glance

The NAEP Science Assessment will include items sampled from the domain of science achievement identified by the intersection of the content areas and science practices (i.e., performance expectations) at grades 4, 8, and 12. The types of items to be used on the assessment include selected-response (multiple-choice) items and constructed-response items (which include short and extended constructed-response items as well as concept-mapping tasks). Some “combination items” may require more than one response. Such combination items include item clusters, Predict-Observe-Explain (POE) item sets, hands-on performance tasks, and interactive computer tasks. The responses requested may be all selected-response, all constructed-response, or a mixture. At each of grades 4, 8, and 12, student assessment time will be divided evenly (50%-50%) between selected-response items and constructed-response items. Extra assessment time will be provided for a portion of the student sample so that hands-on performance tasks and interactive computer tasks can be administered.

At grade 4, the items will be distributed approximately evenly among Physical Science, Life Science, and Earth and Space Science. At grade 8, the balance shifts toward a somewhat greater emphasis on Earth and Space Science, whereas at grade 12, the balance shifts toward the Physical and Life Sciences with a lesser emphasis on Earth and Space Science.

Finally, the distribution of items across the science practices will be approximately 60% combining Identifying Science Principles and Using Science Principles, 30% Using Scientific Inquiry, and 10% Using Technological Design. Moving from grades 4 to 8 to 12, the emphasis on Using Science Principles increases, while the emphasis on Identifying Science Principles decreases. The expectation is that, as students move up through the grades, their critical response skills and methodological and analytical capabilities will increase.

Comparing the NAEP Science Specifications: 1996-2005 → 2009

The chapter overviews provided in the Preface reflect the differences between the NAEP 1996-2005 science specifications and the new *Specifications* to be used for 2009 and future NAEP science assessments. The assessment resulting from the 2009 *Specifications* will start a new NAEP "science trend" (i.e., measure of student progress in science). One difference between the two specifications documents is that the 2009 Steering and Planning Committees drew upon a variety of new standards and assessments. They also were able to extract findings from research in science, science education, and cognition, as well as consider the use of technology to increase the options for assessment administration. Table 1 lists the major differences between the 1996-2005 and 2009 NAEP science specifications documents.

Table 1. NAEP Science Specifications: 1996-2005 → 2009

	1996-2005 Specifications	2009 Specifications
Science Content	Few science standards available on which to base the content to be assessed	Content drawn from existing standards and assessment frameworks: *National Standards*, *Benchmarks,* TIMSS, PISA, and state standards
	Content areas organized into Physical Science, Life Science, and Earth Science	Content areas organized into Physical Science, Life Science, and Earth and Space Science
	Recommendations on distribution of questions by fields of science and grade: approximately equal distribution in grades 4 and 12; a somewhat heavier emphasis on Life Science in grade 8	Recommendations on distribution of questions by science content area and grade: equal weight for all three sciences in grade 4; emphasis on Earth and Space Science in grade 8; emphasis on Physical Science and Life Science at grade 12
	Content presented as bullets and short phrases	Content presented as statements in tables organized by science content subtopics (e.g., "Forces Affecting Motion" from Physical Science) and by grade level
		Boundary tables in *Specifications* clarify the intent of the content statements, i.e., what is included or excluded
	Framework and *Specifications* employed three abstract themes: Systems, Models, and Patterns of Change	
		Framework and *Specifications* employ crosscutting content among Life, Physical, and Earth and Space Sciences
	Assessment asked questions about the nature of science	

	1996-2005 Specifications	2009 Specifications
Science Practices	Knowing and Doing dimension organized into Conceptual Understanding, Scientific Investigation, and Practical Reasoning	Science Practices dimension organized into Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design
		Nature of science treated within science practices, particularly Using Science Principles and Using Scientific Inquiry
	Science practices assessed were largely experience-based	Science practices assessed take into account extant research and the cognitive complexity of the items
	Assessment included items on practical reasoning (i.e., applying science to suggest effective solutions to everyday problems)	
		Assessment includes questions on technological design (i.e., the systematic process of applying science knowledge and skills to solve problems in a real-world context)
	Forty-five percent of the assessment focused on conceptual understanding	Sixty percent of the assessment focuses on conceptual understanding (Identifying and Using Science Principles)
		Learning progressions (i.e., connected sequences of science performances across grade spans) are included
Assessment Items		Assessment uses the history of science and the relationship between science and technology as contexts for questions
	Assessment included both paper-and-pencil and hands-on performance tasks	Assessment includes paper-and-pencil questions, hands-on performance tasks, and interactive computer tasks
	No illustrative items to convey science knowledge or practices in the *Framework*; only a few suggested ideas for items provided in the *Specifications*	Illustrative items that convey science knowledge and practices are included in both the *Framework* and the *Specifications*
		Framework and *Specifications* include guidelines for assessing students with disabilities and English language learners
		Framework and *Specifications* include examples showing how questions are generated and interpreted
		Students' naïve conceptions about science principles are explicitly assessed

Aligning the Assessment with the Framework and Specifications

The assessment should be developed so that it is aligned with the performance expectations defined by the intersection of content statements and science practices, as set forth in the *Framework* and *Specifications*. More specifically:

1. The content of the assessment should be matched with the content of the *Framework* and *Specifications*. The assessment as a whole should reflect the breadth of knowledge covered by the topics, subtopics, and content statements in the *Framework* and *Specifications.* The content of the assessment should not go beyond the content boundaries, as defined in this document. The assessment should represent the balance of science content at each grade as described in the *Framework* and *Specifications.*
2. The science practices and cognitive demands on the assessment should be matched to those in the *Framework* and *Specifications*. The assessment should represent the balance of science practices at each grade as described in the *Framework* and *Specifications*.
3. The emphases of the assessment should match the emphases of topics, subtopics, content statements, science practices, and cognitive demands. The emphases of the assessment should also take into account assessment item contexts such as the history and nature of science and the relationship between science and technology (see Chapter Five).
4. While it is not possible to cover every possible content statement and performance expectation in the *Framework* and *Specifications* on one assessment, appropriate alignment between the assessment and the *Framework* and *Specifications* at each grade should be maintained in the item pools. The assessment should be built so that the constructs represented by the performance expectations in each content area are adequately represented. The breadth and relative emphasis of science knowledge (as well as practices and cognitive demands) covered in each content area, as presented in the *Framework* and *Specifications*, should be represented on the assessment as a whole. The developer should avoid under- or over-emphasizing particular content statements or performance expectations, the goal being to ensure broad coverage in any given year's item pool and coverage of all content statements over time.
5. The assessment should represent the balance of types of items specified in the *Framework* and *Specifications* and give appropriate emphasis to the conditions in which students are expected to demonstrate their science achievement, reflecting the use of laboratory equipment and materials, simulations, and real-world settings.
6. The assessment should report and interpret scores based on the *Framework, Specifications,* and NAEP achievement level descriptions. That is, the assessment should be developed so that scores will reflect both the performance expectations in the *Framework* and *Specifications* and the range of performances illustrated in the preliminary NAEP science achievement level descriptions.
7. The assessment design should match the characteristics of the targeted assessment population. That is, the assessment should give all students tested a reasonable

opportunity to demonstrate their knowledge and skills in the content areas, science practices, and cognitive demands covered by the *Framework* and *Specifications*.

CHAPTER TWO: SCIENCE CONTENT

Introduction

This chapter presents a series of statements that describe the science content of the 2009 NAEP Science Assessment. The content statements contain key science principles for NAEP assessment. Note that, in this document, the phrase "science principles" is broadly conceived and encompasses not only the key principles but also the facts, concepts, laws, and theories of science. In order to specify the science that should be assessed at each grade level, the *Framework* and *Specifications* organize the science content into the three broad content areas that generally make up the K-12 school curriculum to which students are exposed:

- Physical Science
- Life Science
- Earth and Space Science

Classifying statements into one primary content area is not always clear-cut and is artificial to some extent. For example, Ernest Rutherford's discovery of the nucleus earned this physicist the Nobel Prize in Chemistry, and Rosalyn Yalow's work on radioimmunoassay earned this physicist the Nobel Prize in Medicine. However, using three broad content areas as an organizer helps ensure that key science content is assessed in a balanced way.

In the interest of clarity, tables are used to depict the content statements at each grade level. The content statements are based on the assumption that a person literate in science is one who understands key science ideas, is aware that science and technology are interdependent human enterprises with strengths and limitations, is familiar with the natural world and recognizes both its diversity and unity, and uses scientific knowledge and ways of thinking for individual and social purposes (see AAAS, 1994, p. xvii).

Two types of textboxes are used throughout this chapter. *Clarification* textboxes provide details on potentially confusing content, such as the distinction between mass and weight. *Item Suggestion* textboxes present item ideas to illustrate points made in the text; these suggestions require further development in that they are neither published nor field-tested items. Answers to selected-response item suggestions are indicated within each textbox. Although the items in these textboxes may assess more than one content statement or practice, only the primary content and practice designations are provided. This follows NAEP practice, which uses only primary designations for items in the analysis and reporting of student responses.

Development of the Content Statements

The selection and generation of specific content statements at each grade level followed a similar approach across the three broad content areas:

- The *National Standards* and *Benchmarks* were used as key documents for identifying the science content to be assessed, pursuant to the charge from the Steering Committee. Various tools, primarily crosswalks between *National Standards* and *Benchmarks* (AAAS, 1997; Kendall & Marzano, 2004), were used to crosscheck the documents' content standards and benchmark statements, and those that were common to both documents were generally given priority. On a case-by-case basis, content not represented in both documents was discussed and decisions made about inclusion or exclusion. Additions were made where warranted by scientific advances in the decade or more since the development of the *National Standards* and *Benchmarks*, or as a consequence of international assessment results from TIMSS and PISA. Some of the NAEP Science content statements are verbatim reproductions of statements from *National Standards* and *Benchmarks*.[3]
- The focus in the selection process was on the central principles of each discipline. The content statements in the *Specifications* represent foundational and pervasive knowledge, key points of scientific theories, and underpinnings upon which complex understandings are built; and/or they demonstrate connectivity to other central content.
- A primary consideration was the grade-level appropriateness and accuracy within grade level of content statements.
- Once key content was identified within subtopics, the progression of ideas and performances, informed by available research, was tracked through grades 4, 8, and 12.
- A deliberate attempt was made to limit the breadth of science content to be assessed so that some important topics could be measured in-depth. Once core content was identified in each science area, additional content statements could be added only if others previously included were eliminated.

The selection and generation of content statements for inclusion in the *Framework* and *Specifications* was not a linear process. While the Planning Committee attempted to use clear and concise language, the complexities associated with the task of defining what students should know and be able to do in science at particular points in their development necessitated an iterative approach that included many perspectives. In addition to internal review by the project Committees and staff, outreach activities gathered external feedback on the content statements from a variety of stakeholders (e.g., teachers, school and district administrators, state science education personnel,

[3] Statements are reprinted from *National Science Education Standards* (NRC, 1996) with permission from the National Academy of Sciences, courtesy of the National Academies Press, Washington, DC. Statements are reprinted from *Benchmarks for Science Literacy* (AAAS, 1993) with permission from Project 2061, on behalf of the American Association for the Advancement of Science, Washington, DC.

policymakers, scientists, and members of business, industry, and post-secondary communities). The *Framework* and *Specifications* should be revisited in the future as new research becomes available and as the influence of new developments in science takes shape in the K-12 curriculum. See Chapter Six and Appendix E for suggested areas of new research in the form of NAEP special studies.

Organization of Science Content

As described above, the *Framework* and *Specifications* organize science content into three broad *content areas* (Physical Science, Life Science, and Earth and Space Science). The content is further organized into *topics* (such as Motion), *subtopics* (such as Forces Affecting Motion), and, finally, grade-specific *content statements*. The description of each broad content area follows this structure of increasing specificity and is presented in two ways: narrative introductions and content statements presented in tables (see Tables 6, 9, and 12 in the sections for Physical Science, Life Science, and Earth and Space Science, respectively). The following table summarizes the 2009 NAEP Science content topics and subtopics.

Table 2. 2009 NAEP Science Content Topics and Subtopics

Physical Science	Life Science	Earth and Space Science
Matter • Properties of Matter • Changes in Matter **Energy** • Forms of Energy • Energy Transfer and Conservation **Motion** • Motion at the Macroscopic Level • Forces Affecting Motion	**Structures and Functions of Living Systems** • Organization and Development • Matter and Energy Transformations • Interdependence **Changes in Living Systems** • Heredity and Reproduction • Evolution and Diversity	**Earth in Space and Time** • Objects in the Universe • History of Earth **Earth Structures** • Properties of Earth Materials • Tectonics **Earth Systems** • Energy in Earth Systems • Climate and Weather • Biogeochemical Cycles

As an organizational tool in Tables 6, 9, and 12, each content statement is preceded by a specific code in bold (e.g., "**L12.10:** Sorting and recombination of genes in sexual reproduction results in a great variety of possible gene combinations from the offspring of any two parents."). Within a code, the letter denotes broad content area ("P" for Physical Science, "L" for Life Science, and "E" for Earth and Space Science); the number before the period denotes grade level (grade 4, 8, or 12); the number following the period denotes the content statement's order of appearance within a given content area and grade. Thus, **L12.10** denotes that this is the tenth content statement to appear in the grade 12 section of the Life Science content statements table. Since the numbering within each content area and grade is strictly sequential, code numbers do *not* necessarily indicate any relationships *across* grades (see, for example, P4.13, P8.13, and P12.13).

Interpretation of the Content Statements

In the *Specifications*, the content statements generally follow a form that is consistent with the *National Standards*, *Benchmarks,* and the practice of the scientific community. The content statements are phrased as propositions that express science principles. Based on evidence, these principles have been verified by the scientific community and are under constant review. An example of how one grade 8 physical science principle is represented in the *Specifications*, *National Standards,* and *Benchmarks* appears below in Table 3.

Table 3. One Grade 8 Physical Science Principle Represented in the *Specifications, National Standards,* and *Benchmarks*

Specifications	***National Standards***	***Benchmarks***
P8.16: Forces have magnitude and direction. Forces can be added. The net force on an object is the sum of all the forces acting on the object. A non-zero net force on an object changes the object's motion; that is, the object's speed and/or direction of motion changes. A net force of zero on an object does not change the object's motion; that is, the object remains at rest or continues to move at a constant speed in a straight line (p. 36)[4]	If more than one force acts on an object along a straight line, then the forces will reinforce or cancel one another, depending on their directions and magnitude. Unbalanced forces will cause changes in the speed or direction of an object's motion (p. 154).	An unbalanced force acting on an object changes its speed or direction of motion, or both. If the force acts toward a single center, the object's path may curve into an orbit around the center (p. 90).

The content statements form the basis for explaining or predicting naturally occurring phenomena. For example, the above content statement about objects in motion can be used to explain and predict the motions of many different specific objects: an ice skater, an automobile, an electron, or a planet.

The content statements do not include observations of phenomena. As the content statements are written, the empirical foundations of the science principles they represent are not detailed. Instead, knowledge is presented in general terms, such as patterns in

[4] This content statement is longer in the *Specifications* than in *Benchmarks* or *National Standards,* not because it introduces additional science principles, but because it has adopted more detailed language.

observations or theoretical models that account for these patterns. Because the NAEP assessment will require students to apply content statements to specific observations of phenomena, the range of specific phenomena needs to be clarified. Examples of appropriate phenomena are provided in the content boundaries for Physical Science, Life Science, and Earth and Space Science (see Tables 7, 10, and 13).

In order to fully understand the content statements and their intent, readers of the *Framework* and *Specifications* should be cognizant of the following:

- While all content statements have been assigned a primary classification, some are likely to fall into more than one content area.
- Some assessment items may draw on more than one content statement at a time.
- Empty cells in the content statement tables denote that a particular subtopic is not recommended for assessment at that grade level.
- Retention of foundational knowledge from one grade to the next is assumed; however, if the relevant content statement does not appear in a succeeding grade level, it should not be assessed.
- The content statements listed in the *Framework* and *Specifications* describe the whole of what is to be assessed on the 2009 NAEP. The content statements should not be interpreted as a complete description of the school science curriculum that should be taught.

Setting Boundaries

Boundaries elaborate the content statements (with linkages to the science practices) contained in the *Framework* and *Specifications*. They serve the function of informing judgments about the appropriateness of an item for a given grade level and subtopic content area (e.g., “Properties of Matter,” which falls under the topic “Matter” in the Physical Science content area). Presented in Tables 7, 10, and 13, the boundaries follow the content statement tables for Physical Science, Life Science, and Earth and Space Science, respectively. General guidelines for the use of boundaries in assessment development are given below; specific guidelines applicable to each category of boundaries follow.

- The boundaries are intended as “notes to the item writer,” not as comprehensive descriptions of the full range of science content to be included on the 2009 NAEP Science Assessment. In the boundary statements, the terms “such as,” “including,” “e.g.,” and “etc.” are used to denote suggestions. The examples to be used on the assessment are not necessarily limited to these suggestions.
- The boundaries do not stand alone and should be considered in conjunction with the relevant content statements and narrative introductions for each of the content areas—Physical Science, Life Science, and Earth and Space Science. Some content statements are very detailed and require less specification of boundaries (e.g., see the detailed content statements for the subtopic “Forces Affecting Motion” in grade 12 on p. 36 and the accompanying brief “examples” category on p. 48).

- Although the boundaries relevant to a given subtopic may focus more heavily on some content statements than others, this is not intended to denote a sense of content priority.
- Throughout the boundaries, the term “Exclusion(s)” is used to denote content that is not appropriate for inclusion on the 2009 NAEP Science Assessment.

The boundaries are formulated using four categories of parameters: (1) phenomena, examples, and observations; (2) instruments, measurement, and representations; (3) technical vocabulary; and (4) clarification/other. Although boundaries for all subtopics are presented using these categories, individual boundaries may be written in slightly different styles, reflecting variation across subtopics in the nature of the science content.

Examples, Observations, and Phenomena

<u>The Boundary Category</u>

This parameter refers to the breadth and depth of topic coverage in the assessment for grades 4, 8, and 12. There may be limits to the phenomena, examples, and observations that should be used in assessment items. Using Physical Science as an example, if the science principle to be assessed is the particulate nature of matter, phenomena that constitute empirical evidence for this model of matter and the canonical version of the model are appropriate for items, including but not limited to physical properties of solids, liquids, and gases; compressibility; changes of state; thermal expansion and thermal conductivity; electrical conductivity; Brownian motion; solution of solid (sugar) in a liquid (water); and diffusion. Taking grade 4 as a further example, it is fair to expect that all students at this level will have experienced most or all of these phenomena, either in their daily lives or as a part of the science curriculum, and they will be able to make simple predictions on the basis of these experiences. However, students will not be expected to relate these observations to the idea that matter is composed of tiny particles in motion. By contrast, at grades 8 and 12, students can be expected to explain that the prediction and other observations they have made are consistent with the idea that the model of matter is composed of particles in motion.

Because it is impractical to list every phenomenon, example, and observation that is appropriate for inclusion on the 2009 NAEP Science Assessment, the boundaries in Tables 7, 10, and 13 often refer to those that are “common” or “familiar” to students (e.g., common plants and animals). “Common” or “familiar” examples are those that appear frequently in curriculum materials (e.g., textbooks, science kits), as well as those that appear frequently in most students’ experiences outside of school (e.g., museums, zoos, children’s periodicals, television). Items will be reviewed to ensure that they are free of racial, cultural, gender, and regional bias (see section on “Item Development and Review” in Chapter Five).

The Boundary Guidelines

In general, the following criteria should be considered when developing and reviewing items:

- Students' reasoning and experience at each grade level (4, 8, or 12)
- The probability that phenomena related to the topic are familiar; that is, the probability that phenomena are experienced in daily life or in the science curriculum
- Available research on children's learning

It is important to note that some boundary statements pertain, in the first instance, to a particular subtopic but are applicable across several subtopics or even across all of the content areas. One such example is "states of matter." Defining the boundary for the pertinent content statements occurs under the Physical Science subtopic "Properties of Matter": only solid, liquid, and gas should be dealt with in the assessment (not states having unusual properties such as plasma, colloids, gels, or superfluid matter). However, this boundary applies across all three of the science content areas, not just in Physical Science.

Instruments, Measurement, and Representations

The Boundary Category

In some cases, boundaries describe the instruments students are expected to be able to use and the level of precision expected of students in measuring or classifying phenomena, or in interpreting measurements. Moreover, scientific information can be represented in different forms. Data are contained in tables and represented graphically and pictorially. Science principles can be represented verbally, through diagrams and photographs, and in mathematical equations. Concept maps represent principles as nodes and arrows illustrating relationships between and among concepts. Diagrams and pictures of objects, including scientific instruments (rulers, graduated cylinders, thermometers); of particles (atoms, molecules, electrons); and of forces between and among particles also are appropriate for assessment items. Further, boundaries attend to the kinds of mathematical operations or equations and chemical formulas that should appear on assessment items.

The Boundary Guidelines

- Unless otherwise stated, inclusion of a type of instrument means that students are expected to both use the instrument and interpret data that are collected by that instrument.
- All measurements will use the metric system.
- Properties to be measured are specified at the subtopic level, such as for "Properties of Matter": mass (weight in grade 4), volume, density.

- Unless otherwise stated, inclusion of a type of representation means that students are expected to both generate and interpret such a representation, as appropriate to their grade level.
- Only representations that students have had the opportunity to use (e.g., in common curricula) should be a part of assessment items. These may include representations not explicitly mentioned in a content statement but subsumed by it, such as images taken by space-based telescopes for the Earth and Space Science subtopic “Objects in the Universe.”
- Mathematics required for understanding and responding to assessment items should not be so complex that an item is assessing mathematical understanding and facility rather than science understanding. Students should have been exposed to the mathematics required for an item one or two years prior to its use on a science assessment item. Computations should be simple so that computation time required is small.

It is impractical to list details about every type of instrument, measurement, or representation that is appropriate for inclusion on the 2009 NAEP Science Assessment. Thus, these are often described in general terms, and item writers have choice in how to translate boundaries into assessment items. For example, “images of fossils” may be included as a type of representation, and it is expected that items will use fossil images that are grade-appropriate and commonly found in the curriculum. Most decisions (e.g., about specific types of fossils and expected student performances related to interpreting these images) are left to the item writers.

Technical Vocabulary

<u>The Boundary Category</u>

In general, the NAEP Science Assessment will concern itself with science principles and their use and application. Yet, vocabulary is important to science communication. Consequently, knowing definitions of science terms has a place in the science assessment. Particularly at grades 8 and 12, students should know the meanings and be able to use science vocabulary associated with the topics to which they have been exposed.

<u>The Boundary Guidelines</u>

- Unless otherwise stated, it should be assumed that the science terms included in assessment items are generally limited to those that appear in the content statements.
- Boundaries specifically note terms beyond those that appear in the content statements that could be included in assessment items; conversely, some terms may be considered off-limits as noted in the “Exclusions.”

Assessing the ability to define or use a science term is distinct from assessing knowledge of a principle with which the term is associated. Some items are designed to assess knowledge of the meaning of science vocabulary, others to assess understanding of science principles. If the intent of the item is to assess the understanding of a principle, then science vocabulary should not impede a student's ability to respond to the item. The items below illustrate this guideline. The following item is designed to measure a student's knowledge of the definition of a science term.

Item Suggestion

Sublimation is the change of

A. a liquid to a solid.
B. a gas to a liquid.
C. a solid to a gas.
D. a liquid to a gas.

Key: C

P8.6, Identifying Science Principles

The next item is designed to assess understanding of the phenomenon of the change of solids into the vapor state. Form A of the stem assesses knowledge of the definition of the science term used for this phenomenon and the understanding of the phenomenon. Form B of the stem does not require knowledge of the term, "sublimation," to demonstrate an understanding of the science principle.

Item Suggestion

[Form A]

Which of the statements below explains sublimation?

[Form B]

Which of the statements below best explains why ice cubes left in the freezer for long periods of time get smaller and smaller?

A. Ice is composed of tiny particles that slowly melt then turn into a vapor.
B. Air in the freezer is warmer than the ice so the ice melts.
C. The motion of the particles of ice gets slower after the ice is in the cold freezer for a long time.
D. Ice is composed of tiny particles in motion and some of the particles move from the surface of the ice cube into the air around it in the freezer.

Key: D

P8.6, Using Science Principles

Clarification

In addition to the three boundary parameters described above, Tables 7, 10, and 13 contain a fourth row allowing for clarifying remarks. These clarifying remarks may include restatement of a “key idea” encompassed by the content statement(s); specification of the intent of a boundary statement; elaboration on suitable types of representations; suggestion of what might be the basis for appropriate assessment tasks; or reference to related content statements, subtopics, or other specific sections of this document. Attention paid to certain science content in this category (e.g., the label “key idea”) is not intended to denote a sense of content priority.

Commentary on the Content Statements

Beyond consideration of the boundary statements developed for each content subtopic (see Tables 7, 10, and 13), the process of item development and making inferences about students’ knowledge and abilities (see Figures 1 and 2, Chapter One) may necessitate further clarification of the content statements themselves. For example, this may involve “detailing” the meanings of individual content statements (e.g., “boiling point” assumes standard atmospheric pressure) or *further* defining the boundaries of the content to be assessed (e.g., grade 12 students are expected to know that DNA provides instructions for assembling proteins but not the details of DNA transcription and translation). See Table 4 for an example of commentary on a grade 8 Physical Science content statement. Although the commentary describes key ideas about waves relevant through grade 12, note that only three of these key ideas are recommended for assessment at grade 8. Whenever possible, commentary on content statements was informed by available research on learning and cognition. Commentary is intended as helpful notes to item writers and is provided for a number of content statements in the *Specifications* (see Table 4 below, Chapter Four, and Appendices F, G, and H). Commentary is not always of the same kind, and the examples provided do not exhaustively cover the content statements. Assessment developers should continue the process of “detailing” the grade-appropriate principles to be assessed for all content statements sampled for a particular NAEP Science Assessment.

Table 4. Commentary on a Physical Science Content Statement

Grade 8: Energy—Forms of Energy
Content Statement **P8.10:** Energy is transferred from place to place. Light energy from the sun travels through space to Earth (radiation). Thermal energy travels from a flame through the metal of a cooking pan to the water in the pan (conduction). Air warmed by a fireplace moves around a room (convection). Waves—including sound and seismic waves, waves on water, and light waves—have energy and transfer energy when they interact with matter.
Commentary Wave principles recommended for assessment at grade 8: • Waves involve transfer of energy without a transfer of matter. • Waves are caused by disturbances and are also themselves disturbances. Some of the energy of these disturbances is transmitted by the wave. • Water, sound, and seismic waves transfer energy through a material. Wave principles that are related but not recommended for assessment at grade 8: • Some waves are transverse (water, seismic), and other waves are longitudinal (sound, seismic). • In transverse waves, the direction of the motion is perpendicular to the disturbance. • In longitudinal waves, the direction of motion is parallel to the disturbance. • Waves (e.g., light waves) traveling from one material to another undergo transmission, reflection, and/or changes in speed. • Waves can be described by their wavelength, amplitude, frequency, and speed (speed is frequency multiplied by wavelength; energy is a function of the amplitude for non-electromagnetic waves). • Light has dual wave-particle properties. Energy and refraction calculations are also not recommended for assessment at grade 8. Note that a quantitative understanding of electromagnetic waves is expected at grade 12. See P12.10.

Physical Science

Physical science principles, including fundamental ideas about matter, energy, and motion, are powerful conceptual tools for making sense of phenomena in physical, living, Earth, and space systems. Familiar changes—an ice cube melting, a baseball changing direction after being struck by a bat, the appearance of a bolt of lightning, the formation and erosion of mountains, and the growth of a plant—can be explained using these fundamental ideas.[5]

Energy is the constant in an ever-changing world. Energy from the sun fuels electrical storms, hurricanes, tornados, and photosynthesis. In turn, the products of photosynthesis (carbohydrates and oxygen) react during respiration to fuel life processes, such as growth and reproduction of plants and animals. Consequently, it is important for students to develop an understanding of physical science principles early and to appreciate their usefulness across Physical Science, Life Science, and Earth and Space Science.

The physical science principles to be assessed are sorted into three topics—Matter, Energy, and Motion. Matter is the "stuff" of the natural world. Energy is involved in all changes in matter. Motion of the heavenly bodies, of objects found in daily experiences (e.g., balls, birds, cars), and of the tiny particles (atoms, molecules, and their component parts) composing all objects and substances is the result of interactions of matter and energy. The content statements have been divided into topics and subtopics as summarized in Table 5.

Table 5. Physical Science Content Topics and Subtopics

Matter
Properties of Matter Changes in Matter
Energy
Forms of Energy Energy Transfer and Conservation
Motion
Motion at the Macroscopic Level Forces Affecting Motion

[5] The importance of developing understanding early and making connections among the physical, life, Earth, and space sciences has motivated increased attention on physical science in elementary school and prompted consideration of rearranging the usual Earth–life–chemistry–physics curriculum sequence.

Matter

The topic, Matter, is divided into two subtopics: Properties of Matter and Changes in Matter. Conservation of mass, the particulate model of matter, and the Periodic Table of the Elements are the conceptual glue tying together these two subtopics and their related principles.

Properties of Matter

Matter has physical and chemical properties. Physical properties common to all matter as well as those physical properties unique to solids, liquids, and gases are included in the *Framework* and *Specifications,* as are chemical properties. All objects and substances in the natural world are composed of matter. Matter has two fundamental properties: matter takes up space, and matter has inertia—it changes motion only when under the influence of a non-zero net force (grade 4). See the following textbox, "A Matter of Mass," for more on mass versus weight and the treatment of this distinction in the *Framework* and *Specifications*.

Clarification: A Matter of Mass

Mass is a property common to all objects. It is the amount of matter (or "stuff") in an object. The more mass an object has, the more inertia (or "sluggishness") it displays when attempts are made to change its speed or direction. Mass is measured in grams (g) or kilograms (kg) (1 kg=1000 g) using a beam or electronic balance.[6]

Weight, on the other hand, is a measure of the force of attraction (gravitational force) between an object and Earth. Every object exerts gravitational force on every other object. The force depends on how much mass the objects have and on how far apart they are. Force and weight are measured in newtons (N) using a spring scale.

Changing an object's position (say from Earth to the moon) will change its weight, but not its mass. For example, on the surface of Earth, a cannon ball has a mass of 10 kg and a weight of 98 N. On the surface of the moon, that same cannon ball still has a mass of 10 kg, but its weight is only 16 N. So, the cannon ball weighs less on the moon than on Earth, even though nothing has been taken away. Why? Because of the moon's lesser mass and smaller radius, the force of attraction between the moon and the cannon ball is less than the force of attraction between Earth and the cannon ball. Hence, it is said that an object on the moon weighs less than the same object weighs on Earth.

These concepts of mass and weight are complicated and potentially confusing to 4th grade students. Hence, the *Framework* and *Specifications* use the more familiar term "weight" in grade 4 to stand for both weight and mass, and this usage is denoted as "weight (mass)." By grades 8 and 12, students are expected to understand the distinction between mass and weight, and thus, both terms will be used as appropriate.

[6] As found in current NAEP practice, metric units of measure are used for grades 4, 8, and 12.

Matter exists in several different physical states, each of which has unique properties (grade 4). Three of the most commonly encountered are solids, liquids, and gases. Shape and compressibility are examples of properties that distinguish solids, liquids, and gases (grade 4).

The particulate model of matter can be used to explain and predict the properties of states of matter, such as why ice is harder than liquid water and why ice (once formed) has a shape independent of its container while liquid water takes the shape of whatever container it is in (grade 4). In the particulate model of matter, the molecules or atoms of which matter is composed are assumed to be tiny particles in motion (grade 8). The motion is translational, rotational, and vibrational (grade 12). This model can be used to explain the properties of solids, liquids, and gases, as well as changes of state. The particulate model can be used to explain the unique properties of water, as described in the following textbox.

Clarification: Unique Properties of Water

Grade 12: Matter—Changes in Matter
P12.5: Changes of state require a transfer of energy. Water has a very high specific heat, meaning it can absorb a large amount of energy while producing only small changes in temperature.

The unique properties of water have important consequences for Earth Systems and Life Science, including the origin and existence of life on Earth. Understanding the substance of water requires knowledge across the Physical Science categories of Matter, Energy, and Motion.

As with all kinds of matter, water's unique properties can be explained by the shape of its molecule, the forces between its molecules in solid (ice) and liquid states, and the resulting arrangement of molecules in solid and liquid states. In particular, in the solid state, the molecular bonds are such that they separate the molecules more than in the liquid state, resulting in ice being less dense than liquid water. The strong intermolecular forces of liquid water account for its high specific heat. (The specific heat of a substance is the amount of energy required to change 1g of the substance by 1^0C.)

The detailed structures of molecules and atoms that compose them serve as models that explain the forces of attraction between molecules. The structure of atoms, especially the outermost electrons, explains the chemical properties of the elements and the formation of the chemical bonds made and broken during chemical reactions (grade 12). The Periodic Table of the Elements (introduced at grade 8) is another way in which order can be made out of the complexity of the variety of types of matter. (In grade 8, the emphasis is on observed periodicity of properties.) The Periodic Table demonstrates the relationship between the atomic number of the elements and their chemical and physical properties and provides a structure for inquiry into the characteristics of the chemical elements (grade 12).

Two classes of chemical substances serve as exemplars of chemical properties. One class is metals (elements), and the other class is acids (compounds). A chemical property of metals is to react with non-metals to form salts. Included among the properties of acids is the formation of characteristic colors when interacting with acid/base indicators and the interaction with bases to produce salts and water (i.e., neutralization) (grade 8).

Changes in Matter

Matter can undergo a variety of changes. Changes are physical if the relations between the molecules of the material are changed such as changing from a solid to a liquid or from a liquid to a gas (grade 4). When matter undergoes physical change, generally no changes occur in the structure of the molecules or atoms composing the matter (grade 8), though there are exceptions (e.g., sulfur). Changes are chemical if they involve the rearrangement of how atoms are bound to one another, thereby changing the molecules of the material. These are changes in the configuration of the outermost electron shell surrounding the nuclei of the interacting atoms (grade 12). Changes are nuclear if the particles are emitted from or absorbed into the nucleus of the atom, changing the atoms themselves into isotopes or different elements (grade 12).

That mass is conserved when matter undergoes physical and chemical changes is a powerful principle for understanding the natural world and was influential in the development of chemical theory. Adherence to the principle discourages the conclusion that something “disappears” (as water seems to disappear from a puddle) and encourages the search for the “missing” matter.

Most nuclear reactions, involving changes in the nuclei of atoms, are very high-energy and result in the formation of elements or nuclei different from those that began the process. In nuclear reactions, a measurable amount of mass is converted into energy (grade 12).

Energy[7]

The topic, energy, is divided into two subtopics, one addressing the forms of energy and the other energy transfer and conservation.

Forms of Energy

Knowing the characteristics of familiar forms of energy (grade 4) and the scientific categories of potential and kinetic energy (grade 8) are useful in coming to the understanding that, for the most part, the natural world can be explained and is predictable. The most basic characteristics of thermal, light, sound, electrical, and

[7] There are different approaches to helping students understand the concept of energy and related observable phenomena, such as light, heat, and sound. These differences are reflected in the *Framework*'s and *Specifications*’ reference documents and have therefore influenced the framing of the Physical Science, Life Science, and Earth and Space Science content statements.

mechanical energy and the relationship between changes in the natural world and energy are included in the *Framework* and *Specifications*.[8] For example, that two objects, one at a higher temperature than the other, come to the same temperature when placed in contact with each other is a familiar experience. Heat as a concept can be used to explain this experience (grade 8).

Energy Transfer and Conservation

That energy is conserved can be demonstrated by keeping track of the familiar forms of energy as they are transferred from one object to another. The chemical potential energy in a battery is transferred by electric current to a light bulb, which in turn transfers the energy in the form of heat (thermal energy) and light to its surroundings (grade 4). The energy stored in the battery decreases as its surroundings are heated. The loss in chemical potential energy equals the light and heat (thermal energy) transferred by the bulb and the wires to their surroundings. Quantitative accounting is complex; however, on a qualitative basis, both the ability to trace energy transfer and the understanding that energy is conserved (grade 8) are of great explanatory and predictive value. Chemical reactions either release energy to the surroundings or cause energy to flow from the surroundings into the system (grade 12). The sun as the main energy source for the Earth provides opportunity at all grade levels to make important connections between the science disciplines (see discussion of crosscutting content on p. 105).

Motion

The topic, motion, is divided into two subtopics. The first addresses motion at the macroscopic level, and the second addresses the forces that affect motion.

Motion at the Macroscopic Level

Objects observed in daily life undergo different kinds of motion (grade 4). The *Framework* and *Specifications* distinguish three kinds of motion (translational, rotational, vibrational) with emphasis on the translational motion of objects in the natural environment (grade 12). Translational motion is more difficult to describe than it appears because descriptions depend on the position of the observer and the frame of reference used. Speed (grades 4 and 8), velocity (grade 12), and acceleration (grade 12) of objects in translational motion are described in terms of change in direction and position in a time interval.

Forces Affecting Motion

It takes energy to change the motion of objects. The energy change is understood in terms of forces. For example, it takes energy for a baseball pitcher to set the ball in motion toward the batter. Also, pushes and pulls applied to objects often result in changes in motion (grade 4). Principles germane to the relationship of forces and motion serve to

---

[8] The term "heat" in Physical Science is used in grade 4 to stand for thermal energy; this usage is denoted as follows: "heat (thermal energy)." "Thermal energy" is used in grades 8 and 12.

motivate the search for forces when objects change their motion or when an object remains at rest even though it seems that the forces acting on it should result in setting it in motion (grade 8).

Some forces act through physical contact of objects while others act at a distance. The force of a bat on a ball and the downward push of a lead block resting on a tabletop are contact forces. Gravitational and magnetic forces act at a distance (grade 8). Magnets do not need to be in contact to attract or repel each other. The Earth and an airplane do not need to be in contact for a force of attraction to exist between them. Qualitative relationships (grade 8) and quantitative relationships (grade 12) between the mass of an object, the magnitude and direction of the net force on the object, and its acceleration are powerful ideas to explain and predict changes in the natural world.

Table 6. Physical Science Content Statements for Grades 4, 8, and 12

Grade 4	Grade 8	Grade 12
Matter		
Properties of Matter: From physical properties common to all objects and substances and physical properties common to solids, liquids and gases (4) to chemical properties, particulate nature of matter, and the Periodic Table of Elements (8) to characteristics of sub-atomic particles and atomic structure (12).		
P4.1: Objects and substances have properties. Weight (mass) and volume are properties that can be measured using appropriate tools.[9] **P4.2:** Objects vary in the extent to which they absorb and reflect light and conduct heat (thermal energy) and electricity. **P4.3:** Matter exists in several different states; the most commonly encountered are solid, liquid, and gas. Each state of matter has unique properties. For instance, gases are easily compressed while solids and liquids are not. The shape of a solid is independent of its container; liquids and gases take the shape of their containers. **P4.4:** Some objects are composed of a single substance; others are composed of more than one substance. **P4.5:** Magnets can repel or attract other magnets. They can also attract certain non-magnetic objects at a distance.	**P8.1:** Properties of solids, liquids, and gases are explained by a model of matter that is composed of tiny particles in motion. **P8.2:** Chemical properties of substances are explained by the arrangement of atoms and molecules. **P8.3**: All substances are composed of one or more of approximately one hundred elements. The Periodic Table organizes the elements into families of elements with similar properties. **P8.4:** Elements are a class of substances composed of a single kind of atom. Compounds are composed of two or more different elements. Each element and compound has physical and chemical properties, such as boiling point, density, color, and conductivity, which are independent of the amount of the sample.[10] **P8.5:** Substances are classified according to their physical and chemical properties. Metals and acids are examples of such classes. Metals are a class of elements that exhibit common physical properties such as conductivity and common chemical properties such as reacting with non-metals to produce salts. Acids are a class of compounds that exhibit common chemical properties including a sour taste, characteristic color changes with litmus and other acid/base indicators, and the tendency to react with bases to produce a salt and water.	**P12.1:** Differences in the physical properties of solids, liquids, and gases are explained by the ways in which the atoms, ions, or molecules of the substances are arranged and the strength of the forces of attraction between the atoms, ions, or molecules. **P12.2:** Electrons, protons, and neutrons are parts of the atom and have measurable properties including mass and, in the case of protons and electrons, charge. The nuclei of atoms are composed of protons and neutrons. A kind of force that is only evident at nuclear distances holds the particles of the nucleus together against the electrical repulsion between the protons. **P12.3:** In the Periodic Table, elements are arranged according to the number of protons (called the atomic number). This organization illustrates commonality and patterns of physical and chemical properties among the elements. **P12.4:** In a neutral atom, the positively charged nucleus is surrounded by the same number of negatively charged electrons. Atoms of an element whose nuclei have different numbers of neutrons are called isotopes.

[9] See textbox on p. 26 for more detail on the distinction between weight and mass.
[10] While this content statement generally holds, there are some compounds that decompose before boiling.

Table 6. Physical Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Matter		
Changes in Matter: From changes of state (4) to physical and chemical changes and conservation of mass (8) to particulate nature of matter, unique physical characteristics of water, and changes at the atomic and molecular level during chemical changes (12).		
P4.6: One way to change matter from one state to another and back again is by heating and cooling.	**P8.6:** Changes of state are explained by a model of matter composed of tiny particles that are in motion. When substances undergo changes of state, neither atoms nor molecules themselves are changed in structure. Mass is conserved when substances undergo changes of state. **P8.7:** Chemical changes can occur when two substances, elements, or compounds react and produce one or more different substances, whose physical and chemical properties are different from the reacting substances. When substances undergo chemical change, the number and kinds of atoms in the reactants are the same as the number and kinds of atoms in the products. Mass is conserved when substances undergo chemical change. The mass of the reactants is the same as the mass of the products.	**P12.5:** Changes of state require a transfer of energy. Water has a very high specific heat, meaning it can absorb a large amount of energy while producing only small changes in temperature.[11] **P12.6:** An atom's electron configuration, particularly of the outermost electrons, determines how the atom can interact with other atoms. The interactions between atoms that hold them together in molecules or between oppositely charged ions are called chemical bonds. **P12.7:** A large number of important reactions involve the transfer of either electrons (oxidation/reduction reactions) or hydrogen ions (acid/base reactions) between reacting ions, molecules, or atoms. In other chemical reactions, atoms interact with one another by sharing electrons to create a bond. An important example is carbon atoms, which can bond to one another in chains, rings, and branching networks to form, along with other kinds of atoms—hydrogen, oxygen, nitrogen, and sulfur—a variety of structures, including synthetic polymers, oils, and the large molecules essential to life.

[11] See textbox on p. 27 for more detail on the unique properties of water.

Table 6. Physical Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Energy		
Forms of Energy: From examples of forms of energy (4) to kinetic energy, potential energy, and light energy from the sun (8) to nuclear energy and waves (12).		
P4.7: Heat (thermal energy), electricity, light, and sound are forms of energy.[12] **P4.8:** Heat (thermal energy) results when substances burn, when certain kinds of materials rub against each other, and when electricity flows though wires. Metals are good conductors of heat (thermal energy) and electricity. Increasing the temperature of any substance requires the addition of energy. **P4.9:** Light travels in straight lines. When light strikes substances and objects through which it cannot pass, shadows result. When light travels obliquely from one substance to another (air and water), it changes direction. **P4.10:** Vibrating objects produce sound. The pitch of sound can be varied by changing the rate of vibration.	**P8.8:** Objects and substances in motion have kinetic energy. For example, a moving baseball can break a window; water flowing down a stream moves pebbles and floating objects along with it. **P8.9:** Three forms of potential energy are gravitational, elastic, and chemical. Gravitational potential energy changes in a system as the relative positions of objects are changed. Objects can have elastic potential energy due to their compression, or chemical potential energy due to the nature and arrangement of the atoms. **P8.10:** Energy is transferred from place to place. Light energy from the sun travels through space to Earth (radiation). Thermal energy travels from a flame through the metal of a cooking pan to the water in the pan (conduction). Air warmed by a fireplace moves around a room (convection). Waves—including sound and seismic waves, waves on water, and light waves—have energy and transfer energy when they interact with matter. **P8.11:** A tiny fraction of the light energy from the sun reaches Earth. Light energy from the sun is Earth's primary source of energy, heating Earth surfaces and providing the energy that results in wind, ocean currents, and storms.	**P12.8:** Atoms and molecules that compose matter are in constant motion (translational, rotational, or vibrational). **P12.9:** Energy may be transferred from one object to another during collisions. **P12.10:** Electromagnetic waves are produced by changing the motion of charges or by changing magnetic fields. The energy of electromagnetic waves is transferred to matter in packets. The energy content of the packets is directly proportional to the frequency of the electromagnetic waves. **P12.11:** Fission and fusion are reactions involving changes in the nuclei of atoms. Fission is the splitting of a large nucleus into smaller nuclei and particles. Fusion involves joining of two relatively light nuclei at extremely high temperature and pressure. Fusion is the process responsible for the energy of the sun and other stars.

[12] See footnote on p. 29 for more detail on the use of the terms "heat" and "thermal energy" in the *Framework* and *Specifications*.

Table 6. Physical Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Energy		
Energy Transfer and Conservation: From electrical circuits (4) to energy transfer and conservation of energy (8) to translational, rotational, and vibrational energy of atoms and molecules, and chemical and nuclear reactions (12).		
P4.11: Electricity flowing through an electrical circuit produces magnetic effects in the wires. In an electrical circuit containing a battery, a bulb, and a bell, energy from the battery is transferred to the bulb and the bell, which in turn transfer the energy to their surroundings as light, sound, and heat (thermal energy).	**P8.12:** When energy is transferred from one system to another, the quantity of energy before transfer equals the quantity of energy after transfer. For example, as an object falls, its potential energy decreases as its speed, and consequently, its kinetic energy increases. While an object is falling, some of the object's kinetic energy is transferred to the medium through which it falls, setting the medium into motion and heating it. **P8.13:** Nuclear reactions take place in the sun. In plants, light from the sun is transferred to oxygen and carbon compounds, which, in combination, have chemical potential energy (photosynthesis).	**P12.12:** Heating increases the translational, rotational, and vibrational energy of the atoms composing elements and the molecules or ions composing compounds. As the translational energy of the atoms, molecules, or ions increases, the temperature of the matter increases. Heating a sample of a crystalline solid increases the vibrational energy of the atoms, molecules, or ions. When the vibrational energy becomes great enough, the crystalline structure breaks down and the solid melts. **P12.13:** The potential energy of an object on Earth's surface is increased when the object's position is changed from one closer to Earth's surface to one farther from Earth's surface. **P12.14:** Chemical reactions either release energy to the environment (exothermic) or absorb energy from the environment (endothermic). **P12.15:** Nuclear reactions—fission and fusion—convert very small amounts of matter into appreciable amounts of energy. **P12.16:** Total energy is conserved in a closed system.

Table 6. Physical Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Motion		
Motion at the Macroscopic Level: From descriptions of position and motion (4) to speed as a quantitative description of motion and graphical representations of speed (8) to velocity and acceleration as quantitative descriptions of motion and the representation of linear velocity and acceleration in tables and graphs (12).		
P4.12: An object's position can be described by locating the object relative to other objects or a background. The description of an object's motion from one observer's view may be different from that reported from a different observer's view. **P4.13:** An object is in motion when its position is changing. The speed of an object is defined by how far it travels divided by the amount of time it took to travel that far.	**P8.14:** An object's motion can be described by its speed and the direction in which it is moving. An object's position can be measured and graphed as a function of time. An object's speed can be measured and graphed as a function of time.	**P12.17:** The motion of an object can be described by its position and velocity as functions of time and by its average speed and average acceleration during intervals of time. **P12.18:** Objects undergo different kinds of motion—translational, rotational, and vibrational.

Table 6. Physical Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Motion		
Forces Affecting Motion: From the association of changes in motion with forces and the association of objects falling toward Earth with gravitational force (4) to qualitative descriptions of magnitude and direction as characteristics of forces, addition of forces, contact forces, forces that act at a distance, and net force on an object and its relationship to the object's motion (8) to quantitative descriptions of universal gravitational and electric forces, and relationships among force, mass, and acceleration (12).		
P4.14: The motion of objects can be changed by pushing or pulling. The size of the change is related to the size of the force (push or pull) and the weight (mass) of the object on which the force is exerted. When an object does not move in response to a push or a pull, it is because another push or pull (friction) is being applied by the environment. **P4.15:** Earth pulls down on all objects with a force called gravity. With a few exceptions (helium filled balloons), objects fall to the ground no matter where the object is on Earth.	**P8.15:** Some forces between objects act when the objects are in direct contact or when they are not touching. Magnetic, electrical, and gravitational forces can act at a distance. **P8.16:** Forces have magnitude and direction. Forces can be added. The net force on an object is the sum of all the forces acting on the object. A non-zero net force on an object changes the object's motion; that is, the object's speed and/or direction of motion changes. A net force of zero on an object does not change the object's motion; that is, the object remains at rest or continues to move at a constant speed in a straight line.	**P12.19:** The motion of an object changes only when a net force is applied. **P12.20:** The magnitude of acceleration of an object depends directly on the strength of the net force and inversely on the mass of the object. This relationship ($a=F_{net}/m$) is independent of the nature of the force. **P12.21:** Whenever one object exerts force on another, a force equal in magnitude and opposite in direction is exerted by the second object back on the first object. In closed systems, momentum is the quantity of motion that is conserved. Conservation of momentum can be used to help validate the relationship $a=F_{net}/m$. **P12.22:** Gravitation is a universal attractive force that each mass exerts on any other mass. The strength of the gravitational force between two masses is proportional to the masses and inversely proportional to the square of the distance between them. **P12.23:** Electric force is a universal force that exists between any two charged objects. Opposite charges attract while like charges repel. The strength of the electric force is proportional to the magnitudes of the charges and inversely proportional to the square of the distance between them. Between any two charged particles, the electric force is vastly greater than the gravitational force.

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12

<table>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td colspan="3">Properties of Matter: Examples, Observations, and Phenomena</td></tr>
<tr>
<td>Common objects, solid and liquid substances found in students' daily lives

<u>Exclusions:</u>
• Unusual substances and mixtures such as mud that cannot be clearly classified as solid or liquid
• Classification of viscous fluids such as glass and obsidian (However, viscous fluids can sometimes be appropriate as examples in assessment tasks. For example, given a description of the physical properties of glass, students could be asked which of the properties are properties of fluids and which are properties of solids.)
• Plasma—a fourth state of matter that has unusual physical properties and not often found in students' experience</td>
<td>Common objects, solid and liquid substances found in students' daily lives

Some common gases: oxygen, carbon dioxide, air, water vapor, common odors, and volatile liquids

<u>Exclusions:</u>
• Colloids
• Viscous fluids (see grade 4)
• Plasma (see grade 4)
• Isotopes</td>
<td>Elements in the first two rows of the Periodic Table (Nos. 1-18)

Other common elements: iron, gold, silver, mercury, iodine, potassium, titanium, chromium, copper

Common compounds of elements cited above

Mixtures of elements cited above and compounds, including solutions

Isotopes (e.g., of iodine, cobalt, uranium)

<u>Exclusions:</u>
• Colloids
• Viscous fluids (see grade 4)
• Plasma (see grade 4)</td>
</tr>
</table>

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

<table>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td colspan="3">Properties of Matter: Instruments, Measurement, and Representations</td></tr>
<tr>
<td>Qualitative descriptions using common properties (e.g., solid and liquid, color names, stretchy, springy)

Linear measures, using rulers or meter sticks, in cm and m

Volume measures using repeated cubes, measuring cups, or graduated cylinders, in mL or L

Weight (mass) measures or qualitative comparisons using spring scales or double pan balances in g or kg

Simple bar graphs and tables, pictures, or diagrams

<u>Exclusions:</u>
• Volume measures using cm^3
• Newtons (especially since weight/mass distinction is inappropriate at this level)</td>
<td>Qualitative descriptions of properties of solids, liquids, and gases

Measures of mass, volume, and linear measures, accurate to tenths of g, mL, cm^3, cm, using standard measuring instruments and techniques, including triple beam balances and water displacement

Calculated measures of density in g/mL or g/cm^3, including relative densities (sinking and floating)

Classification of elements as metals or non-metals

Drawings of common atoms and molecules (such as H_2O and CO_2) showing how atoms are linked to make molecules (not internal structure of atoms)

Graphs and tables

<u>Exclusions:</u>
• Chemical formulas (except for common molecules such as H_2O and CO_2)</td>
<td>Measures of linear dimensions, mass, volume, density, melting point, and boiling point using standard laboratory instruments

Qualitative descriptions of other chemical and physical properties (limited to conductivity of heat and electricity, solubility)

Chemical formulas and equations involving elements and compounds listed in the above “Examples, Observations, and Phenomena” category

Drawings of atoms and molecules (using Lewis structures or equivalent—note exclusion below of “Lewis structures” as a vocabulary term)

Graphs and tables of physical or chemical properties

Use of arrows to represent forces of attraction or repulsion among molecules</td>
</tr>
<tr><td colspan="3">Properties of Matter: Technical Vocabulary</td></tr>
<tr>
<td><u>Exclusions:</u>
• Compressed
• Independent
• Mass</td>
<td>[No special vocabulary]</td>
<td>Elements and compounds listed in the above “Examples, Observations, and Phenomena” category

<u>Exclusions:</u>
• The term, “Lewis structures”</td>
</tr>
</table>

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Properties of Matter: Clarification		
Concrete observable properties of matter and liquid, solid, and gas states See textbox on p. 26 for more detail on the distinction between weight and mass. See Appendix F for elaboration of P4.1 and P4.2.	Measurable chemical and physical properties, particulate nature of matter, and the Periodic Table as a logical organization of the elements according to observed periodicity of properties Note connection between this subtopic and content statements P8.6 and P8.7. See textbox on p. 26 for more detail on the distinction between weight and mass.	Key idea: Power of the Periodic Table to enable predictions (e.g., properties, reactivity) See clarification statements under content boundaries for "Changes in Matter." Note connections between this subtopic and content statements P12.6 and P12.23. See textbox on p. 26 for more detail on the distinction between weight and mass. See Appendix F for elaboration of P12.1.

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Changes in Matter: Examples, Observations, and Phenomena		
Changes of state at the macroscopic level: melting, freezing, evaporation, and condensation Temperature changes associated with physical changes Familiar relationships between changes in matter and heating and cooling	Use of the particulate nature of matter to explain physical changes (changes of state including sublimation, diffusion, thermal expansion, solution) at the molecular/atomic level Chemical changes that illustrate evidence that a change is chemical rather than physical Differences between physical and chemical changes (e.g., physical changes are reversible; chemical changes involve production of a new substance) Qualitative measures of temperature change and thermal energy associated with chemical and physical changes Influences (qualitative only) on the temperature of freezing and boiling, such as pressure or contamination (e.g., by salt) See content statement P8.5 for examples of chemical changes (metals reacting with non-metals and acids reacting with bases).	Classification of chemical changes as oxidation-reduction, acid-base (neutralization), synthesis, or decomposition Photosynthesis included as a synthesis reaction Respiration included as an oxidation-reduction reaction Use of particulate nature of matter to explain chemical changes and to explain the difference between physical and chemical changes Measures of temperature and thermal energy changes associated with physical changes Distinction between physical changes that result in taking energy from the surroundings and those that add energy to the surroundings Phenomena that are used to support a theory of matter composed of small particles in motion including those from "Properties of Matter" and qualitative relationships among temperature, pressure, and volume of gases, Brownian motion, diffusions of a soluble substance in a liquid, and sublimation

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Changes in Matter: Instruments, Measurement, and Representations		
Beakers, graduated cylinders, balances, and thermometers Metric units of weight (mass), volume, and temperature	Chemical formulas for familiar chemical reactions (such as $2H_2 + O_2 \rightarrow 2H_2O$; $C + O_2 \rightarrow CO_2$) Diagrams representing particles and forces among them in solids, liquids and gases Drawings of common atoms and molecules (such as H_2O and CO_2) showing how atoms are linked to make molecules (not internal structure of atoms) Interpretation of chemical symbols and formulas	Symbols and structural diagrams (Lewis structures) for elements and compounds of importance in Life and Earth and Space Sciences such as sugars and Earth materials (minerals) Chemical equations for oxidation-reduction, neutralization, synthesis, and decomposition reactions of importance in Life and Earth and Space Sciences Simple calculations relating temperature change and thermal energy measured in calories
Changes in Matter: Technical Vocabulary		
Exclusions: • Macroscopic	Condensation, evaporation, freezing, melting	Exclusions: • The term, “Lewis structures”

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Changes in Matter: Clarification		
See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of changes in matter. See Appendix F for elaboration of P4.6.	Key idea: changes of state and conservation of mass are explained by a particulate model of matter See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of changes in matter. See Appendix F for elaboration of P8.6.	Energy transfers and relationships among interacting particles as a means of understanding chemical reactions On the atomic scale, electrical forces being responsible for holding an individual molecule together as well as for the attraction between neighboring molecules Note connections between this subtopic and content statements P12.2, P12.4, and P12.23. See textbox on p. 27 for more on the unique properties of water. See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of changes in matter. See Appendix F for elaboration of P12.5. Exclusions: • Details of van der Waals dispersion forces (London forces) and dipole-dipole attractions

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Forms of Energy: Examples, Observations, and Phenomena		
Observations of heat (thermal), electrical, light, and sound energy through observations of changes in familiar systems, such as: • Addition of heat (thermal) energy: changes of state • Addition of heat (thermal) energy: burning organic substances (wood, wax, sugar, nuts), qualitative relationships between amount of fuel and heat/light produced • Addition of light energy: warming of objects that absorb light • Addition of sound energy: breaking or vibration of a glass • Electricity: heating of wires, lighting of bulb, ringing of bell, powering of small toys and electronic games	Observations of kinetic energy through observations of changes in familiar systems, such as: • Addition of kinetic energy: wind turning a windmill Observations of potential energy through appearance of related form of energy, such as: • Potential energy stored in food: appearance of thermal energy and motion of muscles • Potential energy stored in wax: appearance of heat and light Observations of thermal energy through conduction, convection, radiation, simple examples of absorption	Observations of nuclear energy through observations of changes in systems containing radioactive substances, such as: • Water used to cool down nuclear reactions in nuclear power plants: observable temperature increase in the water • Radioactive isotopes of elements: emission of particles • Thermonuclear reactions: light emission
Forms of Energy: Instruments, Measurement, and Representations		
Primarily verbal descriptions of evidence of energy in familiar systems, that is, if a change is observed, a form of energy is identified as a probable cause of the change Qualitative descriptions of energy	Primarily verbal descriptions of evidence of energy in familiar systems, that is, if a change is observed, a form of energy is identified as a probable cause of the change Qualitative relationships between the mass of combustible materials and the quantity of thermal energy produced. (For example, when candles are used to heat water, the wax "disappears," and the temperature of the water increases. The mass of wax that burns is directly related to the increase in the temperature of the water.)	Primarily verbal descriptions of evidence of energy in familiar systems, that is, if a change is observed, a form of energy is identified as a probable cause of the change Quantitative measurements of wave properties: frequency, amplitude, wavelength, and velocity
Forms of Energy: Technical Vocabulary		
<u>Exclusions:</u> • Thermal energy	[No special vocabulary]	Alpha, beta, and gamma particles

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Forms of Energy: Clarification		
See footnote on p. 29 for use of the terms “heat” and “thermal energy.” Refer to boundaries for “Energy Transfer and Conservation” for appropriate related item content. See pp. 105-106, “Energy Sources and Transfer” and “Uses, Transformations, and Conservation of Energy,” for more on the crosscutting nature of energy. See Appendix F for elaboration of P4.8.	See footnote on p. 29 for use of the terms “heat” and “thermal energy.” Refer to boundaries for “Energy Transfer and Conservation” for appropriate related item content. See pp. 105-106, “Energy Sources and Transfer” and “Uses, Transformations, and Conservation of Energy,” for more on the crosscutting nature of energy. See Appendix F for elaboration of P8.8, P8.9, and P8.10.	See footnote on p. 29 for use of the terms “heat” and “thermal energy.” Refer to boundaries for “Energy Transfer and Conservation” for appropriate related item content. See pp. 105-106, “Energy Sources and Transfer” and “Uses, Transformations, and Conservation of Energy,” for more on the crosscutting nature of energy. See Appendix F for elaboration of P12.8.

Grade 4	Grade 8	Grade 12
Energy Transfer and Conservation: Examples, Observations, and Phenomena		
Electrical circuits: transfer of energy from a battery to light, sound, motion (in wires, light bulbs, buzzers, bells, fans); qualitative relationships between batteries and light bulbs	Falling objects: transfer of gravitational potential energy to kinetic energy and thermal energy produced by friction Transfer of energy from a person lifting an object to gravitational potential energy Power plants: transfer of gravitational potential energy (water power) to mechanical energy to electrical energy; transfer of chemical energy from burning coal or natural gas to mechanical and to electrical energy Plants: transfer of light energy to chemical energy	Calculations of the temperature of a mixture of two liquids of different temperatures and volumes Distinction between chemical changes that either release energy to the surroundings or cause energy to flow from the surroundings into the system

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Energy Transfer and Conservation: Instruments, Measurement, and Representations		
Drawing of a simple electrical circuit containing a battery, bulb and switch (with electrical symbols for wires, bulb, and battery)	Representations of systems involving energy transfer and conservation (e.g., basketball thrown up in the air and falling down to the ground, power plant burning coal for electrical energy)	Calorie as a measure of thermal energy and stored chemical energy Joule as a unit of energy Graphical representation of changes of state as water is heated at a constant rate Use of energy transfer and conservation to explain graphical representations of phase change as thermal energy is added to a system at a constant rate Mathematical reasoning and representations: • Calorimetry: Calculation of changes in temperatures of objects in closed systems • Qualitative comparisons of changes in potential energy with corresponding changes in kinetic energy • Calculations of gravitational potential energy (GPE) of an object very close to Earth's surface and the change in GPE when the distance of the object from Earth's surface is increased (GPE=mgh) • Calculations of kinetic energy and speed of a falling object very close to Earth's surface as the object's GPE decreases ($mg\Delta h+\Delta \frac{1}{2} mv^2=0$)
Energy Transfer and Conservation: Technical Vocabulary		
[No special vocabulary]	[No special vocabulary]	Calorie, joule, exothermic, endothermic

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Energy Transfer and Conservation: Clarification		
Key idea: As one form of energy “disappears” from a system, other forms of energy “appear.” Although not explicitly stated, knowing that a closed electrical circuit is necessary for the transfer of energy is implied by content statement P4.11. Refer to boundaries for “Forms of Energy” for appropriate related item content. See pp. 105-106, “Energy Sources and Transfer” and “Uses, Transformations, and Conservation of Energy,” for more on the crosscutting nature of energy.	Key idea: As the quantity of one form of energy in a system decreases, the quantity of other forms of energy increases. Note connections between this subtopic and content statements L8.4, E8.11. Refer to boundaries for “Forms of Energy” for appropriate related item content. See pp. 105-106, “Energy Sources and Transfer” and “Uses, Transformations, and Conservation of Energy,” for more on the crosscutting nature of energy.	Key idea: Kinetic molecular theory of matter is based on observations, phenomena, and measurements of the type listed above. Kinetic molecular theory of matter accounts for conservation of energy. Note connections between this subtopic and content statements L12.4, L12.5, L12.6, E12.9, E12.12. Refer to boundaries for “Forms of Energy” for related item content. See pp. 105-106, “Energy Sources and Transfer” and “Uses, Transformations, and Conservation of Energy,” for more on the crosscutting nature of energy. See Appendix F for elaboration of P12.12 and P12.16.

Grade 4	Grade 8	Grade 12
Motion at the Macroscopic Level: Examples, Observations, and Phenomena		
Comparisons of speed by comparing distances traveled in equal times or times to travel equal distances Position relative to reference object (e.g., positions of balls on horizontal surfaces)	Quantitative comparisons of speeds in different parts of a journey (e.g., to/from inclined surfaces, along horizontal surfaces) Descriptions of the motion of one object relative to another where both are moving along a straight line Change in speed or direction for particular intervals of time	Quantitative comparisons of average speeds and average accelerations in different parts of a journey and/or for different objects Curvilinear paths: circular, parabolic, elliptical, or simply “around a corner” Translational motion of objects or objects moving from place to place (e.g., falling stone); rotational motion of objects or objects turning about an axis (e.g., carousel); vibrational motion of objects or objects moving rapidly back and forth in place (e.g., vibrating guitar string) <u>Exclusions:</u> • Independence of horizontal and vertical motion for projectiles

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Motion at the Macroscopic Level: Instruments, Measurement, and Representations		
Time (at an instant) as read on a clock Time (interval) found by subtracting the first clock-reading from the final clock-reading Measures of time: hours, minutes, seconds Position and distance scales Measuring distance with a ruler, meter stick Measures of distance: cm, m, km Pictures, drawings, or diagrams of positions of objects at different times relative to reference objects or distance scales Pictures, diagrams, or drawings of objects on horizontal surfaces Cardinal points (north, south, east, west) on a compass as reference points for motion in a single plane	Reading speedometers “Freeze-frame” pictures with distance and time scales Tables and graphs of position versus time (clock-readings) and of speed versus time (clock-readings) Pictures, diagrams, or drawings of objects on horizontal surfaces, inclined surfaces Measures of time: hours, minutes, seconds Measures of distance: cm, m, km	“Freeze-frame” pictures with distance and time scales Tables and graphs of position versus time and of speed versus time (same key ideas as at grade 8 but with more complex motions such as speeds that change moment by moment or several different intervals of different constant speeds) Use of vectors to describe and interpret the motion of an object moving along a curvilinear path Pictures of projectiles Calculations of average speeds and average accelerations for different parts of a journey Quantitative comparisons of rates of translation, rotation, and vibration Measures of time: hours, minutes, seconds Measures of distance: cm, m, km
Motion at the Macroscopic Level: Technical Vocabulary		
Distance (from one position on a reference scale to another) Exclusions: • Velocity • Reference object	Exclusions: • Velocity	Average velocity as “a vector including magnitude and direction”

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Motion at the Macroscopic Level: Clarification		
Description of position and motion of objects Relationships between objects and changes in point of view See Appendix F for elaboration of P4.12 and P4.13.	See Appendix F for elaboration of P8.14.	Velocity and acceleration as means of describing the motion of objects Translational motion of objects (difficult to describe because descriptions depend on observer's position and reference frame used) See Appendix F for elaboration of P12.18.

Grade 4	Grade 8	Grade 12
Forces Affecting Motion: Examples, Observations, and Phenomena		
Objects (e.g., toys) moving across horizontal surfaces If no force keeps them moving, objects slowing down at rates depending on the kinds of surface (friction) Dropped objects falling to the ground <u>Exclusions:</u> • Passive forces (e.g., support force by a table holding up an object)	Low friction surfaces such as air hockey pucks on air, block of ice on smooth table surface, low friction wheeled cart Surfaces showing the effects of friction on objects moving across them (e.g., toy car on carpet)	Two different masses (different densities but equal volumes) falling from equal heights in approximately equal times (ignoring air resistance) Some objects falling more slowly than others when they have substantial air resistance (e.g., parachute) The ability of a charged rod to lift up a tiny piece of paper, demonstrating that between two charged particles, the electric force is larger than Earth's gravitational force <u>Exclusions:</u> • Quantitative comparisons of electric and gravitational forces

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Forces Affecting Motion: Instruments, Measurement, and Representations		
Measuring push or pull by the relative stretch or compression of an object (e.g., a spring or rubber ring) Equal arm balances used to compare masses	Spring scales to pull objects across different surfaces Force diagrams with arrows to represent the relative magnitude and direction of forces acting on an object of interest (limited to forces acting in a straight line)	Force diagrams on each of two interacting objects; force diagrams with relative magnitudes can be used to compare the forces acting on each object of the pair Use of force diagrams and equations to show qualitatively how two objects of different densities can fall in approximately equal times, if air resistance is not a major factor, and explanations of how major air resistance on an object affects the force diagram and resulting motion of the object $a=F_{net}/m$ to predict or compare accelerations or masses of objects, or the net force acting on objects Conservation of momentum (where momentum=mv) used to predict relative motions or relative masses of two interacting objects along a straight line Relationship of distance to gravitational force: doubling (or tripling) the distance between two masses reduces the magnitude of the gravitational force to one quarter (or one ninth). Relationship of distance to electrical force: doubling (or tripling) the distance between two charges reduces the magnitude of the electrical force to one quarter (or one ninth). Exclusions: • Momentum in two dimensions
Forces Affecting Motion: Technical Vocabulary		
Weight (mass) as measure of heaviness Exclusions: • Friction	Attract, repel (as related to forces)	[No special vocabulary]

Table 7. Physical Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Forces Affecting Motion: Clarification		
More force applied to one movable object will make it move (speed up) faster. With the same force applied to two objects, the heavier one will move (speed up) more slowly. To get a heavier object and a lighter object to move at the same speed, the heavier object requires a bigger push/pull.	Resolution of forces on an object should be confined to motion in a straight line. This subtopic includes a qualitative understanding of Newton's first two laws of motion applied to an object.	Resolution of forces should be confined to horizontal, vertical, or inclines of 30 or 45 degrees. Resolution of orthogonal forces should result in a vector at roughly 30, 60 or 45 degrees relative to one of the "first" vectors. This subtopic includes all three of Newton's Laws of Motion applied to two interacting objects. For all of the mathematical relationships/representations described in this subtopic (see "Instruments, Measurement, and Representations" on previous page), students having a qualitative or semi-quantitative understanding (e.g., mathematical relationships such as proportionality) is more important than calculating particular quantities. See clarification statements under content boundaries for "Changes in Matter." Note connections between this subtopic and content statements P12.2, P12.4, P12.6. See Appendix F for elaboration of P12.19.

Life Science

Life science principles are essential for understanding the functioning of living organisms and their interactions with their environment. In addition, life science principles are crucial for understanding advances in science and technology and appreciating their implications for social and personal decisions. Take, for instance, the following discoveries of the past 25 years, all of which rely on understanding basic ideas in life science: the publication of the human genome and genomes of other organisms, the ability to monitor the oxygen level of specific regions of the brain, and the depletion of the ozone layer by human activities. The media regularly ask questions related to health and disease, such as what constitutes a healthy lifestyle and how to deal with the mutability of bacteria, viruses, and parasites that thwart efforts to develop antibiotics and vaccines. While science does not currently provide complete answers to questions like these, it provides the tools for understanding and addressing them.

Understanding principles in Life Science is inextricably linked with understanding principles in Physical Science and Earth and Space Science. “Living organisms are made of the same components as all other matter, involve the same kind of transformations of energy, and move using the same basic kinds of forces” (AAAS, 1994, p. 59).

Understanding living systems and their interactions with their environment requires not only understanding various levels of biological organization—molecules, cells, tissues/organs, organisms, populations, ecosystems—but also understanding interactions (including the transfer of information) within and across these levels and how they can change over time. For example, understanding how populations of organisms change over time is greatly facilitated by understanding the changes that occur in DNA molecules. These changes are manifest in an organism’s traits and may affect its ability to survive and reproduce, which can lead to changing proportions of traits in populations over time.

As summarized in Table 8, the Life Science content statements are sorted into topics and subtopics that, collectively, address structure, function, and patterns of change in living systems. However, any attempt to organize Life Science by a linear set of topics and subtopics, such as those listed below, is somewhat arbitrary. The overlap is evident in Table 9, Life Science Content Statements for Grades 4, 8, and 12.

Table 8. Life Science Content Topics and Subtopics

Structures and Functions of Living Systems
Organization and Development Matter and Energy Transformations Interdependence
Changes in Living Systems
Heredity and Reproduction Evolution and Diversity

Structures and Functions of Living Systems

This topic comprises the ways that living systems are organized and how living systems carry out their life functions. Reasoning about living systems often involves relating different levels of organization, from the molecule to the biosphere, and understanding how living systems are structured at each level. The functions of living systems at these levels, particularly how they transform matter and energy, are included, as are the interactions among living systems and how they depend on one another to carry out their functions.

Organization and Development

As was pointed out early in the 20th century, "Long ago it became evident that the key to every biological problem must finally be sought in the cell; for every living organism is, or at some time has been, a cell" (Wilson, 1928, p. 1). All living things are made up of cells whose work is carried out by many different types of molecules. Cellular and molecular biology has the power to explain a wide variety of phenomena related to the organization and development of living systems, such as synthesis and reproduction, the extraction of energy from food, and regulation. Living organisms have a variety of observable features that enable them to obtain food and reproduce (grade 4). The functions of living organisms are carried out at different levels of organization. In multicellular organisms, cells form organs and organ systems (grade 8). Organisms are subsystems of populations, communities, ecosystems, and the biosphere. Cellular processes are carried out by molecules, particularly proteins. These processes are regulated, both internally and externally, by the environments in which cells exist, including local environments that lead to cell differentiation during the development of multicellular organisms (grade 12).

Matter and Energy Transformations

Matter and energy transformations are involved in all life processes, such as photosynthesis, growth and repair, cellular respiration, and the need of living systems for continual input of energy.

All single-celled and multicellular organisms have the same basic needs: water, air, a source of energy and materials for growth and repair, waste disposal, and conditions for growth and reproduction (grade 4). In terms of matter and energy transformations, the source of food is the distinguishing difference between plants and animals (see textbox below).

Clarification: "Food"

Both plants and animals require a source of energy and materials for growth and repair, and both plants and animals use high-energy compounds as a source of fuel and building material. Plants are distinguished from animals by the fact that plants have the capability (through photosynthesis) to take energy from light to form higher energy molecules containing carbon, hydrogen, and oxygen (carbohydrates) from lower energy molecules.

Plants are similar to animals in that, to make other molecules for their growth and reproduction, they use the energy that is released as carbohydrates react with oxygen. In making these other molecules, plants use breakdown products of carbohydrates, along with minerals from the soil and from fertilizers (known colloquially as "plant foods"), as building blocks. Plants also synthesize substances (carbohydrates, fats, proteins, vitamins) that are components of foods eaten by animals.

So, while synthesis and breakdown are common to both plants and animals, photosynthesis (the conversion of light energy into stored chemical energy) is unique to plants, making them the primary source of energy for all animals.

Basic needs are connected with the processes of growth and metabolism. Organisms are made up of carbon-containing molecules; these molecules originate in molecules that plants assemble from carbon dioxide and water. In converting carbon-containing molecules back to water and carbon dioxide, organisms release energy, making some of it available to support life functions (grade 8). Matter and energy transformations in cells, organisms, and ecosystems have a chemical basis (grade 12). The flow of energy through an ecosystem illustrates principles that cut across the content areas (see discussion of crosscutting content on p. 106).

<u>Interdependence</u>

The species interaction in an ecosystem, the dynamics of population growth and decline, the use of resources by multiple species, their impact on their environment, and the complex interactions among all of these have enormous consequences to the survival of all species, including humans.

All animals and most plants depend on both other organisms and their environments for their basic needs (grade 4). Organisms interact with one another in a variety of ways, such as producer/consumer, predator/prey, and parasite/host. In addition to competition among organisms, the size of populations depends on environmental conditions, such as the availability of water, light, and other suitable conditions (grade 8). Ecosystems are characterized by both stability and change, on which human populations can have an impact (grade 12).

Changes in Living Systems

This topic comprises how organisms reproduce, how they pass genetic information to their offspring, and how genetic information can change as it passes from one generation to the next. Over time, these changes can affect the size, diversity, and genetic composition of populations (i.e., the process of biological evolution).

Heredity and Reproduction

Organisms closely resemble their parents; their slight variations can accumulate over many generations and result in more obvious differences between organisms and their ancestors. Recent advances in biochemistry and cell biology have increased understanding of the mechanisms of inheritance and enabled the detection of disease-related genes. Such knowledge is making it possible to design and produce large quantities of substances to treat disease and, in years to come, may lead to cures.

All plants and animals (and one-celled organisms) develop and have the capacity to reproduce (grade 4). Reproduction, whether sexual or asexual, is a requirement for the survival of species. Characteristics of organisms are influenced by heredity and environment (grade 8). Genetic differences among individuals and species are fundamentally chemical. Different organisms are made up of somewhat different proteins. Reproduction involves passing the DNA with instructions for making these proteins from one generation to the next with occasional modifications (grade 12).

Evolution and Diversity

Earth's present-day life forms have evolved from common ancestors reaching back to the simplest one-celled organisms almost four billion years ago. Modern ideas about evolution provide a scientific explanation for three main sets of observable facts about life on Earth: the enormous number of different life forms that exist; the systematic similarities in anatomy and molecular chemistry seen within that diversity; and the sequences of changes in fossils found in successive layers of rock that have been formed over more than a billion years. The modern concept of evolution, including natural selection and common descent, provides a unifying principle for understanding the history of life on Earth, relationships among all living things, and the dependence of life on the physical environment. The concept is so well established that it provides a framework for organizing most of biological knowledge into a coherent picture.

All organisms are similar to and different from other organisms, and some kinds of organisms and individuals have advantages in particular environments (grade 4).

Preferential survival means that differences among individuals in a population affect their ability to survive and reproduce. Classification reflects degrees of relatedness among species (grade 8). Evolution is the consequence of natural selection and differential reproduction. Natural selection and common descent provide the scientific explanation for the history of life on Earth as depicted in the fossil record and as indicated by anatomical and chemical similarities evident within the diversity of existing organisms (grade 12).

Table 9. Life Science Content Statements for Grades 4, 8, and 12

Grade 4	Grade 8	Grade 12
Structures and Functions of Living Systems		
Organization and Development: From basic needs of organisms (4) to the levels of organization of living systems (8) to the chemical basis of living systems (12).		
L4.1: Organisms need food, water, and air; a way to dispose of waste; and an environment in which they can live.[13]	**L8.1:** All organisms are composed of cells, from just one cell to many cells. About two-thirds of the weight of cells is accounted for by water, which gives cells many of their properties. In multicellular organisms, specialized cells perform specialized functions. Organs and organ systems are composed of cells and function to serve the needs of cells for food, air, and waste removal. The way in which cells function is similar in all living organisms.[14] **L8.2:** Following fertilization, cell division produces a small cluster of cells that then differentiate by appearance and function to form the basic tissues of an embryo.	**L12.1:** Living systems are made of complex molecules (including carbohydrates, fats, proteins, and nucleic acids) that consist mostly of a few elements, especially carbon, hydrogen, oxygen, nitrogen, and phosphorous. **L12.2:** Cellular processes are carried out by many different types of molecules, mostly proteins. Protein molecules are long, usually folded chains made from combinations of amino-acid molecules. Protein molecules assemble fats and carbohydrates and carry out other cellular functions. The function of each protein molecule depends on its specific sequence of amino acids and the shape of the molecule. **L12.3:** Cellular processes are regulated both internally and externally by environments in which cells exist, including local environments that lead to cell differentiation during the development of multicellular organisms. During the development of complex multicellular organisms, cell differentiation is regulated through the expression of different genes.

[13] See p. 53 for a textbox on "Food."
[14] Human organs and organ systems are subsumed under this content statement. See details in Table 10, Life Science Content Boundaries.

Table 9. Life Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Structures and Functions of Living Systems		
Matter and Energy Transformations: From the basic needs of organisms for growth (4) to the role of carbon compounds in growth and metabolism (8) to the chemical basis of matter and energy transformation in living systems (12).		
L4.2: Organisms have basic needs. Animals require air, water, and a source of energy and building material for growth and repair. Plants also require light.	**L8.3:** Cells carry out the many functions needed to sustain life. They grow and divide, thereby producing more cells. Food is used to provide energy for the work that cells do and is a source of the molecular building blocks from which needed materials are assembled. **L8.4:** Plants are producers—they use the energy from light to make sugar molecules from the atoms of carbon dioxide and water.[15] Plants use these sugars along with minerals from the soil to form fats, proteins, and carbohydrates. These products can be used immediately, incorporated into the plant's cells as the plant grows, or stored for later use. **L8.5:** All animals, including humans, are consumers that meet their energy needs by eating other organisms or their products. Consumers break down the structures of the organisms they eat to make the materials they need to grow and function. Decomposers, including bacteria and fungi, use dead organisms or their products to meet their energy needs.	**L12.4:** Plants have the capability (through photosynthesis) to take energy from light to form higher energy sugar molecules containing carbon, hydrogen, and oxygen from lower energy molecules. These sugar molecules can be used to make amino acids and other carbon-containing (organic) molecules and assembled into larger molecules with biological activity (including proteins, DNA, carbohydrates, and fats). **L12.5:** The chemical elements that make up the molecules of living things pass through food webs and are combined and recombined in different ways. At each link in an ecosystem, some energy is stored in newly made structures, but much is dissipated into the environment as heat. Continual input of energy from sunlight keeps the process going. **L12.6:** As matter cycles and energy flows through different levels of organization of living systems—cells, organs, organisms, communities—and between living systems and the physical environment, chemical elements are recombined in different ways. Each recombination results in storage and dissipation of energy into the environment as heat. Matter and energy are conserved in each change.

[15] The statement "they use the energy from light" does not imply that energy is converted into matter or that energy is lost. See pp. 105-106 for more on energy as Crosscutting Content.

Table 9. Life Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Structures and Functions of Living Systems		
Interdependence: From the interdependence of organisms (4) to specific types of interdependence (8) to consequences of interdependence (12).		
L4.3: Organisms interact and are interdependent in various ways including providing food and shelter to one another. Organisms can survive only in environments in which their needs are met. Some interactions are beneficial; others are detrimental to the organism and other organisms. **L4.4:** When the environment changes, some plants and animals survive and reproduce; others die or move to new locations.	**L8.6:** Two types of organisms may interact with one another in several ways: They may be in a producer/consumer, predator/prey, or parasite/host relationship. Or, one organism may scavenge or decompose another. Relationships may be competitive or mutually beneficial. Some species have become so adapted to each other that neither could survive without the other. **L8.7:** The number of organisms and populations an ecosystem can support depends on the biotic resources available and abiotic factors, such as quantity of light and water, range of temperatures, and soil composition. **L8.8:** All organisms cause changes in the environment where they live. Some of these changes are detrimental to the organisms or other organisms, whereas others are beneficial.	**L12.7:** Although the interrelationships and interdependence of organisms may generate biological communities in ecosystems that are stable for hundreds or thousands of years, ecosystems always change when climate changes or when one or more new species appear as a result of migration or local evolution. The impact of the human species has major consequences for other species.

Table 9. Life Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Changes in Living Systems		
Heredity and Reproduction: From life cycles (4) to reproduction and the influence of heredity and the environment on an offspring's characteristics (8) to the molecular basis of heredity (12).		
L4.5: Plants and animals have life cycles. Both plants and animals begin life and develop into adults, reproduce, and eventually die. The details of this life cycle are different for different organisms. **L4.6:** Plants and animals closely resemble their parents.	**L8.9:** Reproduction is a characteristic of all living systems**;** because no individual organism lives forever, reproduction is essential to the continuation of every species. Some organisms reproduce asexually. Other organisms reproduce sexually. **L8.10:** The characteristics of organisms are influenced by heredity and environment. For some characteristics, inheritance is more important; for other characteristics, interactions with the environment are more important.	**L12.8:** Hereditary information is contained in genes, located in the chromosomes of each cell. A human cell contains many thousands of different genes. One or many genes can determine an inherited trait of an individual, and a single gene can influence more than one trait. **L12.9:** The genetic information encoded in DNA molecules provides instructions for assembling protein molecules. Genes are segments of DNA molecules. Inserting, deleting, or substituting DNA segments can alter genes. An altered gene may be passed on to every cell that develops from it. The resulting features may help, harm, or have little or no effect on the offspring's success in its environment. **L12.10:** Sorting and recombination of genes in sexual reproduction results in a great variety of possible gene combinations from the offspring of any two parents.

Table 9. Life Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Changes in Living Systems		
Evolution and Diversity: From differences and adaptations of organisms (4) to preferential survival and relatedness of organisms (8) to the mechanisms of evolutionary change and the history of life on Earth (12).		
L4.7: Different kinds of organisms have characteristics that enable them to survive in different environments. Individuals of the same kind differ in their characteristics, and sometimes the differences give individuals an advantage in surviving and reproducing.	**L8.11:** Individual organisms with certain traits in particular environments are more likely than others to survive and have offspring. When an environment changes, the advantage or disadvantage of characteristics can change. Extinction of a species occurs when the environment changes and the characteristics of a species are insufficient to allow survival. Fossils indicate that many organisms that lived long ago are extinct. Extinction of species is common; most of the species that have lived on the Earth no longer exist. **L8.12:** Similarities among organisms are found in anatomical features, which can be used to infer the degree of relatedness among organisms. In classifying organisms, biologists consider details of internal and external structures to be more important than behavior or general appearance.	**L12.11:** Modern ideas about evolution (including natural selection and common descent) provide a scientific explanation for the history of life on Earth as depicted in the fossil record and in the similarities evident within the diversity of existing organisms. **L12.12:** Molecular evidence substantiates the anatomical evidence for evolution and provides additional detail about the sequence in which various lines of descent branched. **L12.13:** Evolution is the consequence of the interactions of (1) the potential for a species to increase its numbers, (2) the genetic variability of offspring due to mutation and recombination of genes, (3) a finite supply of the resources required for life, and (4) the ensuing selection from environmental pressure of those organisms better able to survive and leave offspring.

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12

Grade 4	Grade 8	Grade 12
Organization and Development: Examples, Observations, and Phenomena		
Common plants and animals that live on land or in water, including vertebrates and insects Food sources for common animals Need for organisms to eliminate waste, waste including undigested food in animals Ecosystems as they provide sources of food and water and air for animals and water and air for plants	Cellular composition of animals (including vertebrates and insects); terrestrial and aquatic plants; and fungi, bacteria, protists Specific organs and organ systems of vertebrates: digestive system (esophagus, stomach, intestines, colon); respiratory system (lungs, gills); circulatory system (heart, arteries, veins, capillaries); urinary system (kidneys, bladder); skin Parts of terrestrial plants: roots, stems, leaves, flowers, seeds Carbon dioxide as an example of waste product in vertebrates, insects, terrestrial and aquatic plants, and many microorganisms Patterns of development in vertebrates, including humans <u>Exclusions:</u> • Specific organs and organ systems of invertebrates • Nitrogen as a component of animal wastes	Complex molecules (listed in the content statement) and their subunits: simple sugars, fatty acids, amino acids Nucleotides as subunits of nucleic acids Cells making the molecules they need: different proteins and their different work (without specific proteins, the processes may not be carried out at all or as well) Cellular processes (e.g., transport of materials, energy capture and release, protein building, waste disposal) in plants and animals <u>Exclusions:</u> • Detailed structure of fats (glycerol, etc.)

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

<table>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td colspan="3">Organization and Development: Instruments, Measurement, and Representations</td></tr>
<tr><td>Pictures and drawings of common plants and animals

Pictures and drawings of common habitats</td><td>Microscopes to see cells in plant and animal tissues and cell division

Relative scale of organs, tissues, cells (not directly measured)

Drawings and photographs of organs and cells of vertebrates

Drawings and photographs of developing vertebrate embryos

Drawings showing movement of food, water, oxygen, carbon dioxide, and digestive waste through land plants and vertebrates

<u>Exclusions:</u>
• Drawings of parts of cells (cellular organelles)
• Drawings showing movement of nitrogen waste through organisms
• Cell differentiation into three tissue layers and blastula formation</td><td>Relative scale of organs, tissues, cells, biomolecules (not directly measured)

Representations of large biomolecules as polymers of simpler subunits

Structural formulas of monomers of fats, proteins, carbohydrates (fatty acids, amino acids, simple sugars)</td></tr>
<tr><td colspan="3">Organization and Development: Technical Vocabulary</td></tr>
<tr><td><u>Exclusions:</u>
• Cells and cellular functions
• Vertebrates
• Ecosystems</td><td>Vocabulary associated with organ systems and terrestrial plants in above “Examples, Observations, and Phenomena” category

<u>Exclusions:</u>
• Seed parts
• Organs not listed above such as pancreas, liver, gall bladder
• Cellular organelles; stages in development of embryos</td><td><u>Exclusions:</u>
• Names of cellular organelles
• Primary, secondary, tertiary structure of proteins; names, functional groups, and structural formulas of amino acids
• Differentiation among types of amino acids: hydrophobic, hydrophilic
• The term “nucleotide” if used in items should be followed by “(a subunit of DNA).”</td></tr>
</table>

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Organization and Development: Clarification		
Analysis of examples of plants and animals that are familiar to them, with consideration of four key needs: • Food • Air • Waste disposal • Environmental conditions See textbox on p. 53 for detail about the use of the word "food." Note connection between this subtopic and content statement L4.7. <u>Exclusions:</u> • Internal structures and functions in plants and animals • Cells, cell organelles	Key idea: How the structures of an organism work together to fulfill its needs and the differentiation of cells that occurs during growth and development Key idea: How organ systems work together physiologically to support the needs of the entire organism and the cells of which it is composed See Appendix G for elaboration of L8.1. <u>Exclusions:</u> • Structure and function of specific tissues and cells within organs (e.g., different types of blood cells or muscle cells)	Subunits (hence complex molecules and living things themselves) are made up of the same elements—carbon, hydrogen, oxygen, nitrogen, phosphorous. Cell processes—replicating genetic information, repairing cell structures, helping other molecules to get in or out of the cell, and generally catalyzing and regulating molecular interactions—are carried out mainly by proteins. Proteins are long and usually folded chains and are made up of amino acids (smaller molecules that are the building blocks of proteins). The function of each protein depends on the amino acid sequence and resulting shape the protein takes when it folds. Note connection between this subtopic and content statement L12.9. See Appendix G for elaboration of L12.2. <u>Exclusions:</u> • How proteins are made (RNA, codons, etc.), how proteins function (e.g., molecular basis of enzyme catalysis, muscle contraction, membrane transport)

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Matter and Energy Transformations: Examples, Observations, and Phenomena		
Common plants and animals that live on land or in the water, including vertebrates and insects	Producers, consumers, and decomposers including animals, plants, protists, fungi, bacteria Gas exchange in plants, animals, decomposers Cells in plants and animals Internal structures and organs of plants and animals responsible for movement of materials <u>Exclusions:</u> • Cellular organelles of plants and animals	Producers, consumers, and decomposers including animals, plants, protists, fungi, bacteria Representative ecosystems (e.g., wetlands, rainforests) that are widely represented in textbooks and of importance to all students, even those not living near them
Matter and Energy Transformations: Instruments, Measurement, and Representations		
Drawings or photographs showing plant parts responsible for taking in water (roots) and sunlight (green leaves) and animals' external organs that take in air, water, and food	Scale of multicellular organisms, cells, molecules Structural formulas of carbon dioxide, oxygen, and water Drawings with arrows showing movements of matter and energy within plants and animals (If arrows are used, their meanings must be clearly explained.) Energy pyramids showing trophic levels and energy transfer in an ecosystem <u>Exclusions:</u> • Structural formulas of biomolecules (sugars, fats, amino acids) • Drawings with arrows showing movements of substances inside plant and animal cells	Representations of large biomolecules as polymers of simpler subunits Labeled structural formulas of monomers of fats, proteins, carbohydrates (amino acids, fatty acids, simple sugars) Chemical equations for overall processes of cellular respiration and photosynthesis General representations of synthesis and breakdown of large biomolecules Diagrams of matter cycling and energy flow in ecosystems <u>Exclusions:</u> • Representations of specific detailed steps of synthesis and decomposition (e.g., intermediate steps and molecules, details of dehydration synthesis)

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Matter and Energy Transformations: Technical Vocabulary		
[No special vocabulary]	<u>Exclusions:</u> • Although the content statements describe the general processes of photosynthesis and cellular respiration, these terms are not included, nor is the word chlorophyll.	Cellular respiration Key elements in biomolecules: carbon, hydrogen, oxygen, nitrogen, phosphorus, sulfur, potassium, sodium <u>Exclusions:</u> • Calvin cycle, Krebs cycle, glycolysis, or intermediate products in respiration and photosynthesis
Matter and Energy Transformations: Clarification		
Materials or sources of energy (light) that plants and animals take in. Legitimate questions include identifying the sources of these essential materials for different plants and animals. See pp. 105-106, "Energy Sources and Transfer" and "Uses, Transformations, and Conservation of Energy," for more on the crosscutting nature of energy. See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of matter in living systems. <u>Exclusions:</u> • *How* plants take in light	Tracing the sequence of transformations of matter and energy that enable plants and animals to grow and function Transformations that use or produce gases are especially significant and difficult for students. Note connections between this subtopic and content statements P8.13 and E8.11. See pp. 105-106, "Energy Sources and Transfer" and "Uses, Transformations, and Conservation of Energy," for more on the crosscutting nature of energy. See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of matter in living systems. See Appendix G for elaboration of L8.4.	Tracking particular atoms or elements (especially carbon) through the processes of photosynthesis, growth, food webs, and cellular respiration at multiple levels: cellular, organismal, and ecological Key idea: Conservation of matter and energy in living systems Note connections between this subtopic and content statements P12.14, P12.15, P12.16, E12.9, E12.12. See pp. 105-106, "Energy Sources and Transfer" and "Uses, Transformations, and Conservation of Energy," for more on the crosscutting nature of energy. See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of matter in living systems. See Appendix G for elaboration of L12.6.

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Interdependence: Examples, Observations, and Phenomena		
Common plants and animals that live on land or in the water, including vertebrates and insects Environmental conditions and changes in environmental conditions that affect survival of plants and animals, including temperature, shelter, light, food sources, water Beneficial and detrimental interactions among organisms that affect survival	Populations of producers, consumers, and decomposers in all five kingdoms Examples of relationships of the types listed in the content statement Examples of disturbances to ecosystems that cause changes in size of populations (e.g., changes in factors listed in content statement, invasive species, human interventions)	Anthropogenic disturbances in ecosystems or environments, including local and global climate change, uses of tilling and pesticides to favor human crops, human land use, harvesting of fish stocks, pollution, invasive species, etc. Natural disturbances in ecosystems (e.g., fires, sedimentation, volcanism) Responses of ecosystems to changes, including succession patterns, alterations due to natural or anthropogenic disturbances Representative ecosystems (e.g., wetlands, rainforests) that are widely represented in textbooks and of importance to all students, even those not living near them
Interdependence: Instruments, Measurement, and Representations		
Identifying and counting organisms in plots Identifying evidence of interactions (e.g., parasites on plants or animals, evidence of predation) Temperature on the Celsius scale Drawings or photographs of plants, animals, and their interactions	Measures of population size and density: range, number of individuals per plot Graphic or geographic representations of size, range, density of populations Food webs identifying relationships among producers, consumers, and decomposers in an ecosystem	Qualitative descriptions of biodiversity in different ecosystems Qualitative descriptions of fluctuations in populations due to natural and anthropogenic disturbances Graphic or geographic representations of size, diversity, key species in ecosystems Analysis of quantitative data on biodiversity
Interdependence: Technical Vocabulary		
Descriptions of environmental conditions and interactions in non-technical language <u>Exclusions:</u> • Predation, symbiosis, parasitism	[No special vocabulary]	See above descriptions of natural and anthropogenic disturbances. <u>Exclusions:</u> • Anthropogenic (use “human-generated” or similar term)

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Interdependence: Clarification		
Key idea: Individual plants and animals interact with their environments or with plants and animals of other species, and those interactions affect their survival or ability to reproduce.	Key idea: Populations, their interactions with other populations and the non-living environment, and how changes in the environment or interactions can affect the size and composition of populations See Appendix G for elaboration of L8.7.	Key idea: Ecosystems (including coupled human and natural systems such as farms and cities) are dynamic systems that are constantly changing, especially in response to the large and increasing scale of human impact on ecosystems. See Appendix G for elaboration of L12.7.

Grade 4	Grade 8	Grade 12
Heredity and Reproduction: Examples, Observations, and Phenomena		
Variety of plant and animal life cycles Descriptions of organisms in terms of characteristics (e.g., height, eye color, hair color) and traits that express variation of these characteristics (e.g., brown eyes, black hair) Different external features of organisms that may resemble their parents (e.g., height, color of hair and eyes)	Sexual and asexual reproduction as part of a life cycle of animals, plants, protists, bacteria	Changes in DNA (mutations) Genetic variation in cells arising from gamete formation and sexual reproduction—inherited traits, mutations Effects of mutations on proteins and protein functions Basic bioinformatics (DNA fingerprinting, reproductive and therapeutic cloning) used as context for items (information on bioinformatics would need to be supplied in the item) <u>Exclusions:</u> • Details of transcription and translation • Mendelian genetics • Details of meiosis and mitosis (content to be assessed limited to the following: the result of DNA duplication in cell division is the copying of all DNA for the descendent cells whereas the result of gamete formation is eggs/sperm with only half of the DNA) • Nature of bonding between DNA strands

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

<table>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td colspan="3">Heredity and Reproduction: Instruments, Measurement, and Representations</td></tr>
<tr><td>Drawings and photographs of plants and animals and their external features</td><td>Representations showing degree of relatedness (e.g., pedigree charts) used to decide which pairs of organisms are more/less closely related</td><td>Using microscopes to identify chromosomes

Electrophoresis as a technique for separating molecules based on differences in their properties (e.g., mass)

Dyes, binding of radioactive labels as a way to visualize molecules (most of which are colorless)

Double helix as a representation to include understanding that it is the sequence of nucleotides that gets passed from one generation to the next and that is responsible for DNA functions

Graphic results of electrophoresis

<u>Exclusions:</u>
PCR
How electrophoresis works
Structure of nucleotides
Graphic demonstration of mitosis or meiosis</td></tr>
<tr><td colspan="3">Heredity and Reproduction: Technical Vocabulary</td></tr>
<tr><td>Descriptions of organisms in commonly understood terms</td><td>Common terms related to reproduction including egg, sperm, fertilization

<u>Exclusions:</u>
Mendelian genetics (e.g., dominant, recessive)
Details of reproduction and development (e.g., gamete, zygote, alternation of generations)</td><td><u>Exclusions:</u>
Names and structures of nucleotides</td></tr>
</table>

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Heredity and Reproduction: Clarification		
[No clarification needed]	Key idea: Every organism has a set of instructions for specifying its traits. Heredity is the passage of these instructions from one generation to another. When cells divide, they are reproducing asexually. In asexual reproduction, the new cell (organism) is identical to the parent. In sexual reproduction, a sperm and egg unite (fertilization); half of an individual's traits come from one parent and half from the other.	Understanding the molecular basis of heredity requires a coherent understanding of two main functions of DNA: (1) determining the characteristics of organisms, and (2) passing information from one generation to the next. Clarification of the role of DNA in determining the characteristics of an organism (function 1): An organism's characteristics reflect the actions of its proteins; proteins carry out the work of cells (and organs/organ systems); and DNA directs the production of proteins. Clarification of DNA carrying hereditary information from one generation to the next (function 2): DNA is what gets passed from parents to offspring. Genes and chromosomes are then defined in terms of DNA. See DNA's role in function 1. Double helix as a representation includes understanding of its significance to heredity. Note connections between this subtopic and content statements L12.1, L12.2, and L12.3. See Appendix G for elaboration of L12.8 and L12.9.

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Evolution and Diversity: Examples, Observations, and Phenomena		
Different plants and animals in the context of their populations and environments Adaptations of populations of plants and animals to their environments Differences among plants and animals in a population	Examples of more and less closely related organisms, including primates Examples of fossils, remains of recently extinct organisms, major extinctions during Earth's history Examples of modern changes in populations (e.g., development of insect resistance to pesticides) <u>Exclusions:</u> • Phylogenetic taxonomy (e.g., kingdoms, phyla)	Examples of modern diversity and evolutionary changes in populations (e.g., development of insect resistance to pesticides, bacterial resistance to antibiotics, viral strains). Such examples illustrate the four numbered factors in L12.13. Examples of fossil records and existing organisms with evident similarities, including humans Examples of lines of descent accompanied by molecular evidence, including humans
Evolution and Diversity: Instruments, Measurement, and Representations		
Measurements or descriptions of characteristics of plants and animals: height, weight, leaf size, color, etc. Photographs or pictures that show differences among individuals within and between populations (can be used to hypothesize relationships between organisms and their environments)	Measurements or descriptions of similarities and differences in internal and external structures of organisms Photographs or pictures that show differences among individuals in a population in internal or external characteristics Simple timelines depicting fossil records and extinction of species	Graphs or tables showing distribution of traits in populations Representations of degree of biochemical similarity in terms of percentages (e.g., 85% of DNA in common)
Evolution and Diversity: Technical Vocabulary		
Words in common language that describe adaptive characteristics and characteristics that differentiate individuals <u>Exclusions:</u> • Trait • Species	Adaptation, fitness	[No special vocabulary]

Table 10. Life Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Evolution and Diversity: Clarification		
Note that this includes *both* diversity of species (different species have different traits) and diversity within populations (different individuals in a population have different traits). Describing and comparing individual organisms and reasoning about how their traits might help or hinder their survival Note connection between this subtopic and content statement L4.1. See Appendix G for elaboration of L4.7. <u>Exclusions:</u> • Distinguishing heritable from non-heritable characteristics	Fossils and internal and external anatomical evidence of appearance and disappearance of species, including extinction Fossils and internal and external anatomical evidence of relatedness of species Note connections between this subtopic and the Earth and Space Science subtopic, "History of Earth." <u>Exclusions:</u> • Chemical evidence of change in populations over time • Chemical evidence of relatedness of species	Producing accounts of changes in populations that include all four of the numbered factors in L12.13 Key idea: Process of evolutionary change in populations through natural selection Linking evidence (including molecular and anatomical) of similarity within diversity of organisms to common descent Note connections between this subtopic and the Earth and Space Science subtopic, "History of Earth."

Earth and Space Science

The past few decades have brought rapid changes in the character of Earth and Space Science. The study of Earth has shifted from surface geology and mining toward global change and Earth systems; and research methods have changed from human observations and mapping to remote sensing and computer modeling.[16] This concept of Earth as a complex and dynamic entity of interrelated subsystems implies that there is no process or phenomenon within the Earth system that occurs in complete isolation from other elements of the system. There has also been a shift in goals, as advances in theory have made it possible to more accurately predict changes in weather and climate, to provide life-saving warnings of floods, hurricanes, earthquakes, and volcanic eruptions, and to understand how human activities influence ecosystem and climate changes across the globe.

In Space Science, similar changes have taken place as a result of new technologies. Successful probes to Mars, Jupiter, and Saturn have vastly expanded knowledge of the solar neighborhood. The discovery of more than 100 planets outside the solar system has raised new questions about the origin of life. Furthermore, advances in ground and space-based telescopes capable of observing many different parts of the electromagnetic spectrum with unprecedented detail have revolutionized understanding of the structure and evolution of the universe itself. In brief, descriptive methods of Earth and Space Science have given way to theory-based inquiry and problem-solving approaches that have far-reaching consequences with regard to understanding the universe and stewardship of Planet Earth.

Changes in Earth and Space Science education are beginning to catch up with advances in research. The *National Standards* emphasizes a systems approach to studying Earth, especially at the high school level. Some of today's textbooks pay less attention to describing Earth features and focus instead on a systems perspective in which Earth is viewed as a physical system of interrelated phenomena, processes, and cycles. Some high school curricula have integrated the traditional Earth science disciplines of geology, meteorology, and oceanography with aspects of biology, chemistry, and physics to introduce students to a more holistic study of Earth.

The tools available to students for learning about Earth and space have changed as well, although not all of these resources are available to all students. Visualization tools, such as Geographical Information System (GIS) software, have made it possible for Earth Science students to have direct access to the raw data and models used by scientists. Other web-based programs allow students to view and process satellite images of Earth, to direct a camera on board the Space Shuttle, and to access professional telescopes around the world to carry out science projects. In other words, the core concepts, subject matter, and tools used by students have undergone profound changes in recent decades that mirror many of the advances in Earth and Space Science. The data and images

[16] "Earth" is capitalized, rather than referred to as "the earth," in order to recognize it as one of the planets in the solar system.

gathered by these tools could be used as source materials for assessment items. For example, students could examine data on sea surface temperatures and upper atmospheric winds, derived from satellite observations, to predict the intensity and track of a hurricane.

To reflect the importance of this content area, NAEP will include questions about Earth and Space Science at the 4^{th}, 8^{th}, and 12^{th} grade levels. The content statements have been divided into topics and subtopics as summarized in Table 11.

Table 11. Earth and Space Science Content Topics and Subtopics

Earth in Space and Time
Objects in the Universe History of Earth
Earth Structures
Properties of Earth Materials Tectonics
Earth Systems
Energy in Earth Systems Climate and Weather Biogeochemical Cycles

Earth in Space and Time

Earth in Space and Time is divided into two subtopics: Objects in the Universe and History of Earth. The idea that "the universe is large and ancient, on scales staggering to the human mind" (AAAS, 1994, p. 40) connects these subtopics.

One of the earliest discoveries of the scientific age was that Earth is not the center of the universe. It is now known that Earth is a planet in space, one of a family of planets and other bodies that circle a yellow star in a vast galaxy of other stars. Like countless other worlds that are known to exist, Earth has a beginning and a history. That history can be read by carefully and thoughtfully observing the world and the universe.

<u>Objects in the Universe</u>

Objects in the sky, such as the sun and the moon, have patterns of movement. These patterns can be observed through changes in shape or placement in the sky based on time of day or season (grade 4). By recognizing these patterns, people have developed calendars and clocks and explained such phenomena as moon phases, eclipses, and seasons (grade 8).

It was previously thought that Earth was the center of the universe, but it is now known that the sun is the central and largest body in the solar system, which includes Earth and other planets and their moons as well as other objects such as asteroids and comets. Objects in the solar system are kept in predictable motion by the force of gravity (grade 8).

According to the "big bang" theory, the entire contents of the known universe expanded explosively into existence from a hot, dense state 13.7 billion years ago. Early in the history of the universe, stars coalesced out of clouds of hydrogen and helium and clumped together by gravitational attraction into billions of galaxies. When heated to a sufficiently high temperature by gravitational attraction, stars begin nuclear reactions, which convert matter to energy and fuse light elements into heavier ones (grade 12).

<u>History of Earth</u>

Theories of planet formation and radioactive dating of meteorites have led to the conclusion that the sun, Earth, and the rest of the solar system formed from a nebular cloud of dust and gas 4.6 billion years ago. Early Earth was very different from today's planet. Initially, there was no life and no molecular oxygen in the atmosphere. There is evidence that one-celled organisms—the bacteria—were the first forms of life on our planet, appearing about 3.5 billion years ago. These bacteria are thought to be responsible for adding oxygen to Earth's atmosphere, making it possible for a wider diversity of life forms to evolve (grade 12).

Earth processes seen today, such as erosion and mountain building, have made possible the measurement of geologic time though methods such as observing rock sequences and using fossils to correlate the sequences at various locations. Fossils also provide evidence of how life and environmental conditions have changed (grade 8). Early methods of determining geological time, such as the use of index fossils and stratigraphic sequences, allowed for the relative dating of geologic events. However, absolute dating was impossible until the discovery that certain radioactive isotopes in rocks have known decay rates, making it possible to determine how many years ago a given rock sample formed (grade 12).

Earth's surface changes over time. Some changes are due to slow processes, such as erosion and weathering; and others are due to rapid processes such as volcanic eruptions, landslides, and earthquakes (grade 4). Changes caused by violent earthquakes and volcanic eruptions can be observed on a human time scale, but many geological processes, such as the building of mountain chains and shifting of entire continents, take place over hundreds of millions of years. Water, ice, waves, and wind sculpt Earth's surface to produce distinctive landforms (grade 12).

Earth Structures

Content statements related to Earth Structures fall into two subtopics: Properties of Earth Materials and Tectonics. The study of Earth materials has contributed to understanding dynamic processes, which are, in turn, driven by the movement of vast tectonic plates. Conversely, the development of tectonic theory has made it possible to locate and extract Earth materials for a wide variety of human uses.

Properties of Earth Materials

Earth materials that occur in nature include rocks, minerals, soils, water, and the gases of the atmosphere. Natural materials have different properties, which sustain plant and animal life (grade 4). Soil consists of weathered rocks and decomposed organic material from dead plants, animals, and bacteria. Soils are often found in layers, with each having a different chemical composition and texture (grade 8). Some Earth materials have properties that make them useful either in their present form or designed and modified to solve human problems and enhance the quality of life (grade 4).

Rocks and rock formations bear evidence of the conditions and forces that created them, ranging from the violent conditions of volcanic eruptions to the slow deposition of sediments. The atmosphere is a mixture of nitrogen, oxygen, and trace gases that include water vapor. The atmosphere has a different physical and chemical composition at different elevations (grade 8).

Tectonics

A basic understanding of geological history, described above, forms the foundation for later understanding of tectonics. Earth's internal structure is layered with a lithosphere, hot convecting mantle, and dense metallic core. Lithospheric plates, on the scale of continents and oceans, constantly move at rates of centimeters per year in response to movements in the mantle. Major geological events such as earthquakes, volcanic eruptions, and mountain building result from these plate motions (grade 8).

Although continental drift was first suggested as early as the late 16th century and further developed in the early 1900s, it was not widely accepted until more convincing evidence emerged as a result of extensive exploration of the sea floor. By the late 1960s, mapping of the Mid-Atlantic Ridge, evidence of sea floor spreading, and subduction led to the more general theory of plate tectonics. The current explanation is that the outward transfer of Earth's internal heat propels the plates comprising Earth's surface across the face of the globe, pushing the plates apart where magma rises to form mid-ocean ridges, and pulling the edges of plates back down where Earth materials sink into the crust at deep trenches (grade 12).

Earth as a whole has a magnetic field that is detectable at the surface with a compass. Earth's magnetic field is similar to the field of a natural or human-made magnet with north and south poles and lines of force. For thousands of years, people have used compasses to aid navigation on land and sea (grade 8). Crucial evidence in support of tectonic theory came from studies of the magnetic properties of rocks on the ocean floor (grade 12).

Earth Systems

Earth Systems is organized according to three subtopics: Energy in Earth Systems, Climate and Weather, and Biogeochemical Cycles. The explorers of the 16th century who circumnavigated the planet were the first to become aware of global weather and climate patterns. As science began to mature and diversify in the 19th and 20th centuries, those who scientifically studied the planet did so from the perspective of the traditional disciplines, such as geology, oceanography, and meteorology. Currently, working with vastly improved technologies, most scientists take an Earth systems perspective, including the study of how energy moves through Earth systems and the integration of disciplines to better understand Earth's biogeochemical cycles.

Energy in Earth Systems

The sun warms the land, air, and water and helps plants grow (grade 4). The sun is the major source of energy for phenomena on Earth's surface. The sun drives convection within the atmosphere and oceans, producing winds, ocean currents, and the water cycle. Seasons result from annual variations in the intensity of sunlight and length of day due to the tilt of Earth's rotation axis relative to the plane of its yearly orbit around the sun (grade 8).

Earth's systems have internal and external sources of energy, both of which create heat. The sun is the major external source of energy. Two primary sources of internal energy are the decay of radioactive isotopes and the gravitational/thermal energy from Earth's original formation (grade 12).

Climate and Weather

Weather changes from day to day and over the seasons. Scientists use tools for recording and predicting weather changes (grade 4). Global patterns of atmospheric movement influence local weather (grade 8).

Climate is determined by energy transfer from the sun at and near Earth's surface. This energy transfer is influenced by dynamic processes such as cloud cover, atmospheric gases, and Earth's rotation, as well as static conditions such as the position of mountain ranges and oceans, seas, and lakes (grade 12). Oceans have a major effect on climate because water in the oceans holds a large amount of heat (grade 8).

Biogeochemical Cycles

Earth is a system containing essentially a fixed amount of each stable chemical atom or element. Each element can exist in several different chemical forms. Elements move within and between the lithosphere, atmosphere, hydrosphere, and biosphere as part of biogeochemical cycles (see discussion of crosscutting content on p. 107). Movement of matter through Earth’s systems is driven by Earth's internal and external sources of energy. These movements are often accompanied by a change in the physical and chemical properties of the matter. Carbon, for example, occurs in carbonate rocks such as limestone, in coal and other fossil fuels, in the atmosphere as carbon dioxide gas, in water as dissolved carbon dioxide, and in all organisms as complex molecules that control the chemistry of life (grade 12).

Water, which covers the majority of Earth’s surface, circulates through the crust, oceans, and atmosphere in what is known as the “water cycle.” Water evaporates from Earth’s surface, rises and cools as it moves to higher elevations, condenses as clouds, falls as rain or snow, and collects in lakes, oceans, soil, and underground (grade 8).

Natural ecosystems provide an array of basic processes that affect humans. These processes include maintenance of the quality of the atmosphere, generation of soils, control of the hydrologic cycle, disposal of wastes, and recycling of nutrients (grade 12). The supply of many Earth resources such as fuels, metals, fresh water, and farmland is limited.

Humans change environments in ways that can either be beneficial or detrimental for themselves and other organisms (grade 4). Humans have devised methods for extending the use of Earth resources through recycling, reuse, and renewal (grade 4). However, other activities—reducing the amount of forest cover, increasing the amount and variety of chemicals released into the atmosphere, and intensive farming—have changed Earth’s land, oceans, and atmosphere. Studies of plant and animal populations have shown that such activities can reduce the number and variety of wild plants and animals and sometimes result in the extinction of species (grade 8).

Table 12. Earth and Space Science Content Statements for Grades 4, 8, and 12

Grade 4	Grade 8	Grade 12
Earth in Space and Time		
Objects in the Universe: From patterns in the sky (4) to a model of the solar system (8) to a vision of the universe (12).		
E4.1: Objects in the sky have patterns of movement. The sun, for example, appears to move across the sky in the same way every day, but its path changes slowly over the seasons. The moon appears to move across the sky on a daily basis much like the sun. **E4.2:** The observable shape of the moon changes from day to day in a cycle that lasts about a month.	**E8.1:** In contrast to an earlier theory that Earth is the center of the universe, it is now known that the sun, an average star, is the central and largest body in the solar system. Earth is the third planet from the sun in a system that includes seven other planets and their moons, as well as smaller objects, such as asteroids and comets. **E8.2:** Gravity is the force that keeps most objects in the solar system in regular and predictable motion. Those motions explain such phenomena as the day, the year, phases of the moon, and eclipses.	**E12.1:** The origin of the universe remains one of the greatest questions in science. The “big bang” theory places the origin approximately 13.7 billion years ago when the universe began in a hot, dense state. According to this theory, the universe has been expanding ever since. **E12.2:** Early in the history of the universe, matter, primarily the light atoms hydrogen and helium, clumped together by gravitational attraction to form countless trillions of stars and billions of galaxies. **E12.3:** Stars, like the sun, transform matter into energy in nuclear reactions. When hydrogen nuclei fuse to form helium, a small amount of matter is converted to energy. These and other processes in stars have led to the formation of all the other elements.

Table 12. Earth and Space Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Earth in Space and Time		
History of Earth: From evidence of change (4) to estimating the timing and sequence of geologic events (8) to theories about Earth's history (12).		
E4.3: The surface of Earth changes. Some changes are due to slow processes, such as erosion and weathering, and some changes are due to rapid processes, such as landslides, volcanic eruptions, and earthquakes.	**E8.3:** Fossils provide important evidence of how life and environmental conditions have changed in a given location. **E8.4:** Earth processes seen today, such as erosion and mountain building, made possible the measurement of geologic time through methods such as observing rock sequences and using fossils to correlate the sequences at various locations.	**E12.4:** Early methods of determining geologic time, such as the use of index fossils and stratigraphic sequences, allowed for the relative dating of geological events. However, absolute dating was impossible until the discovery that certain radioactive isotopes in rocks have known decay rates, making it possible to determine how many years ago a given rock sample formed. **E12.5:** Theories of planet formation and radioactive dating of meteorites and lunar samples have led to the conclusion that the sun, Earth, and the rest of the solar system formed from a nebular cloud of dust and gas 4.6 billion years ago. **E12.6:** Early Earth was very different from today's planet. Evidence for one-celled forms of life—the bacteria—extends back more than 3.5 billion years. The evolution of life caused dramatic changes in the composition of Earth's atmosphere, which did not originally contain molecular oxygen. **E12.7:** Earth's current structure has been influenced by both sporadic and gradual events. Changes caused by violent earthquakes and volcanic eruptions can be observed on a human time scale, but many geological processes, such as the building of mountain chains and shifting of entire continents, take place over hundreds of millions of years.

Table 12. Earth and Space Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Earth Structures		
Properties of Earth Materials: From natural and human-made materials (4) to soil analysis and layers of the atmosphere (8).		
E4.4: Earth materials that occur in nature include rocks, minerals, soils, water, and the gases of the atmosphere. **E4.5:** Natural materials have different properties, which sustain plant and animal life. **E4.6:** Some Earth materials have properties that make them useful either in their present form or designed and modified to solve human problems and enhance the quality of life, as in the case of materials used for building or fuels used for heating and transportation.	**E8.5:** Rocks and rock formations bear evidence of the minerals, materials, temperature/pressure conditions, and forces that created them. Some formations show evidence that they were deposited by volcanic eruptions. Others are composed of sand and smaller particles buried and cemented by dissolved minerals to form solid rock again. Still others show evidence that they were once earlier rock types that were exposed to heat and pressure until they changed shape and in some cases melted and recrystallized. **E8.6:** Soil consists of weathered rocks and decomposed organic material from dead plants, animals, and bacteria. Soils are often found in layers with each having a different chemical composition and texture. **E8.7:** The atmosphere is a mixture of nitrogen, oxygen, and trace gases that include water vapor. The atmosphere has a different physical and chemical composition at different elevations.	

Table 12. Earth and Space Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Earth Structures		
Tectonics: From the basics of tectonic theory and Earth magnetism (8) to the physical mechanism that drives tectonics and its supporting evidence (12).		
	E8.8: The Earth is layered with a lithosphere; hot, convecting mantle; and dense, metallic core. **E8.9:** Lithospheric plates on the scale of continents and oceans constantly move at rates of centimeters per year in response to movements in the mantle. Major geological events, such as earthquakes, volcanic eruptions, and mountain building, result from these plate motions. **E8.10:** Earth as a whole has a magnetic field that is detectable at the surface with a compass. Earth's magnetic field is similar to the field of a natural or human-made magnet with north and south poles and lines of force. For thousands of years, people have used compasses to aid in navigation on land and sea.	**E12.8:** Mapping of the Mid-Atlantic Ridge, evidence of sea floor spreading, and subduction provided crucial evidence in support of the theory of plate tectonics. The theory currently explains plate motion as follows: the outward transfer of Earth's internal heat propels the plates comprising Earth's surface across the face of the globe. Plates are pushed apart where magma rises to form mid-ocean ridges, and the edges of plates are pulled back down where Earth materials sink into the crust at deep trenches.

Table 12. Earth and Space Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Earth Systems		
Energy in Earth Systems: From role of the sun (4) to the sun's observable effects (8) to internal and external sources of energy in Earth systems (12).		
E4.7: The sun warms the land, air, and water and helps plants grow.	**E8.11:** The sun is the major source of energy for phenomena on Earth's surface. The sun provides energy for plants to grow and drives convection within the atmosphere and oceans, producing winds, ocean currents, and the water cycle. **E8.12:** Seasons result from annual variations in the intensity of sunlight and length of day, due to the tilt of Earth's rotation axis relative to the plane of its yearly orbit around the sun.	**E12.9:** Earth systems have internal and external sources of energy, both of which create heat. The sun is the major external source of energy. Two primary sources of internal energy are the decay of radioactive isotopes and the gravitational energy from Earth's original formation.
Climate and Weather: From local weather (4) to global weather patterns (8) to systems that influence climate (12).		
E4.8: Weather changes from day to day and over the seasons. **E4.9:** Scientists use tools for observing, recording, and predicting weather changes from day to day and over the seasons.	**E8.13:** Global patterns of atmospheric movement influence local weather. Oceans have a major effect on climate because water in the oceans holds a large amount of heat.	**E12.10:** Climate is determined by energy transfer from the sun at and near Earth's surface. This energy transfer is influenced by dynamic processes such as cloud cover, atmospheric gases, and Earth's rotation, as well as static conditions such as the positions of mountain ranges and of oceans, seas, and lakes.

Table 12. Earth and Space Science Content Statements for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Earth Systems		
Biogeochemical Cycles: From uses of Earth resources (4) to natural and human-induced changes in Earth materials and systems (8) to biogeochemical cycles in Earth systems (12).		
E4.10: The supply of many Earth resources such as fuels, metals, fresh water, and farmland is limited. Humans have devised methods for extending the use of Earth resources through recycling, reuse, and renewal. **E4.11:** Humans depend on their natural and constructed environment. Humans change environments in ways that can either be beneficial or detrimental for themselves and other organisms.	**E8.14:** Water, which covers the majority of Earth's surface, circulates through the crust, oceans, and atmosphere in what is known as the "water cycle." Water evaporates from Earth's surface, rises and cools as it moves to higher elevations, condenses as clouds, falls as rain or snow, and collects in lakes, oceans, soil, and underground. **E8.15:** Human activities, such as reducing the amount of forest cover, increasing the amount and variety of chemicals released into the atmosphere, and intensive farming, have changed Earth's land, oceans, and atmosphere. Studies of plant and animal populations have shown that such activities can reduce the number and variety of wild plants and animals and sometimes result in the extinction of species.	**E12.11:** Earth is a system containing essentially a fixed amount of each stable chemical atom or element. Most elements can exist in several different chemical forms. Earth elements move within and between the lithosphere, atmosphere, hydrosphere, and biosphere as part of biogeochemical cycles. **E12.12:** Movement of matter through Earth's systems is driven by Earth's internal and external sources of energy. These movements are often accompanied by a change in the physical and chemical properties of the matter. Carbon, for example, occurs in carbonate rocks such as limestone, in coal and other fossil fuels, in the atmosphere as carbon dioxide gas, in water as dissolved carbon dioxide, and in all organisms as complex molecules that control the chemistry of life. **E12.13:** Natural ecosystems provide an array of basic processes that affect humans. These processes include maintenance of the quality of the atmosphere, generation of soils, control of the hydrologic cycle, disposal of wastes, and recycling of nutrients.

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12

Grade 4	Grade 8	Grade 12
Objects in the Universe: Examples, Observations, and Phenomena		
Apparent positions of the sun and moon in the sky at different times of day Differences in the appearance of the moon during different times of the month Sun's apparent path across the sky in different seasons Real-world problems involving knowledge of the sun's daily path (e.g., placement of drapes for sleeping during the day, position of solar panels to take full advantage of solar radiation) Exclusions: • Clouds, birds, or airplanes (since the intention of these statements is to introduce students to "Objects in the Universe") • Observations of stars and galaxies	Moon phases Solar and lunar eclipses Gravity acting between all objects and operating to keep planets and satellites in their orbits Launching rockets and placing satellites in orbit Astronauts in orbit experiencing gravity as they "fall" around Earth Exclusions: • Mathematical formulation of Kepler's laws or the equations for elliptical orbits or parabolic trajectories	Evolution of universe: major features of "big bang" theory Structure of universe: Location of our galaxy among others and role of gravitation in galaxy formation Origin of elements in general: formation of H and He in early universe and formation of heavier elements by fusion in stars Processes in stars: Qualitative understanding about nuclear fusion in stars that release energy and produce heavier elements Exclusions: • Details of the "big bang" such as "inflation" or "decoupling" • Reaction rates in stars • How galaxies are organized into larger structures • Details of nuclear reactions

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Objects in the Universe: Instruments, Measurement, and Representations		
Sundial-type devices to find patterns in the sun's movement with respect to Earth: length and position of shadows of a vertical stick placed in the ground (e.g., flagpole) observed at regular intervals during the day (understanding the relationship between position of an object, its shadow, and the sun) Photographs or simple sketches of monthly cycle of moon phases (full, more than half full, about half full, less than half full, "a sliver") Period length of the sun's path across the sky (one day) and moon's phases (about one month) Simple sketches showing the movement of a gnomon's shadow, the path of the sun, moon phases, or changing position and appearance of the moon <u>Exclusions:</u> • Telescope	Unaided eye (e.g., for daytime and early evening observations of sun, moon, and planets) Telescopes and binoculars (e.g., to see stars; planets; and lunar craters, mountains, and large dark areas on the moon) Drawings to represent spatial relationships (using space perspective) Drawings or photographs of the sun and planets in the solar system and their motions Illustrations of mathematical concepts (ellipses, angular measures, and scale) using conventional notation Sketches and diagrams of sun's path, moon phases, planetary movements, rocket launches Qualitative relationships among distance, mass, and gravitational forces <u>Exclusions:</u> • Inverse square relationship of gravitation and distance measurements	Telescopes and binoculars to see stars, nebulae, and galaxies Spectroscopes to determine the composition of stars as well as the speed of stars and galaxies towards or away from us Computer simulations of processes in stars, the formation of galaxies and the "big bang" Technologies (e.g., telescopes, spectroscopes) that provide evidence to test theories about stars and the universe Images taken by large or space-based telescopes and spectra of stars and galaxies Interpretation of graphs in terms of relationships (e.g., brightness versus temperature, distance to galaxies versus redshift) <u>Exclusions:</u> • How to operate spectroscopes or calculate redshift • Naming laws represented by graphs (e.g., Hubble's Law, HR Diagram)
Objects in the Universe: Technical Vocabulary		
<u>Exclusions:</u> • Northern Hemisphere • Technical terms related to moon phases (e.g., waning, waxing, gibbous) • Ecliptic (use term such as "path of the sun") • Gnomon (use term such as "vertical stick")	Common terms in astronomy included, such as trajectory, orbit, and eclipse, and angular measurement in degrees <u>Exclusions:</u> • Technical terms nonessential for conceptual understanding, such as apogee and perigee	Technical vocabulary from physics and chemistry questions as well as commonly used astronomical terms such as galaxy, nebula, and supernova <u>Exclusions:</u> • Exotic objects (e.g., neutron stars, black holes)

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

<table>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td colspan="3">Objects in the Universe: Clarification</td></tr>
<tr>
<td>Changes and patterns in the sky as seen from Earth—the Earth-centric view

In the Northern Hemisphere, in the morning, the sun is generally to the east; at noon, it is toward the south; and during the afternoon, it is to the west.

The moon can be seen some of the time during the day and some of the time during the night. It seems to make a daily path across the sky. When the moon can be seen, the lighted side gets its light from the sun.

The moon looks a little different every day, but it has the same appearance again in about a month. Proper sequencing of pictures of moon phases during the month.

See Appendix H for elaboration of E4.1 and E4.2.

<u>Exclusions:</u>
• The space perspective in explaining these patterns (moon phases, sun’s path across the sky, seasons)
• Explanations for sun and moon eclipses; relationships between eclipses and moon phases
• Relative distances of sun, moon, Earth</td>
<td>The space perspective of the solar system to explain phenomena such as moon phases

Appearance of the sun, moon, and stars going around Earth once a day is due to Earth’s rotation on its axis.

The moon revolves around Earth about once a month, and Earth revolves around the sun once a year. The space perspective can explain the day-night cycle, moon phases, and solar and lunar eclipses.

Gravity pulls things “down” towards the center of Earth. A common naïve conception is that there is no gravity in space because there is no air there.

See Appendix H for elaboration of E8.2.

<u>Exclusions:</u>
• Underlying physics principles of above phenomena</td>
<td>Observational evidence and physics concepts that support theories about how stars shine and the origin of the universe

<u>Exclusions:</u>
• Phenomena such as neutron stars or black holes</td>
</tr>
</table>

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
History of Earth: Examples, Observations, and Phenomena		
Simple changes such as rust on metals, wood rot on buildings, and mold growth on food Examples of processes that break down Earth structures, such as erosion and landslides Examples of processes that build up Earth structures and materials, such as the slow formation of soils and sedimentary rocks when rivers and streams deposit sand and mud; or the rapid buildup (or destruction) of mountains during volcanic eruptions, or uplift of large land masses during earthquakes <u>Exclusions:</u> • Names of the processes (focus should be on the processes themselves). For example, it is not important that students distinguish between weathering and erosion. • Slow mountain building resulting from plate tectonics	Order of events based on images or drawings of road cuts, sediment cores, or other evidence; evidence of more recent events found closer to the surface (Law of Superposition) Data on current rate of deposition used to determine how long it took a thick layer of sediment to form (keeping numbers simple) Fossil evidence showing that environmental conditions changed over time (a cluster of items may tap this well) Determining if two rock strata that look alike date from the same period of time by examining the fossils within them	Using index fossils to determine time sequences and half-life of radioactive isotopes to determine time spans Key phases in the geologic evolution of Earth (e.g., formation of the planet more than 4 billion years ago, evolution of single-celled animals more than 3 billion years ago, development of an atmosphere with oxygen about 2 billion years ago, multi-celled animals less than 1 billion years ago) Simple quantitative problems based on measured rates of geologic processes, such as mountain-building, deposition, or erosion; calculations of how long it took for major geologic changes to occur <u>Exclusions:</u> • Detailed mechanism of radioactive decay (emphasis should instead be on its use to solve geologic puzzles)

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
History of Earth: Instruments, Measurement, and Representations		
Laboratory experiments or computer simulations of processes that change Earth structures and materials (e.g., stream tables showing modeling consequences of erosion and deposition; rotted log becoming soft and looking more like soil than wood; metal body of a car rusting; comparisons of erosion of slopes with bare soil and slopes with grassy soil) Photographs that capture images at different points in time Images showing how Earth structures have changed over time (e.g., Mt. St. Helens before eruption, just after eruption, one year later, and 20 years later); hypothesize the simple processes that were likely to have caused those changes	Sample cores of a lakebed or other column, revealing the history of deposition over time Simple calculations of time spans given data on rate of deposition Interpretation of fossil replicas (or images of fossils) to reconstruct extinct animals and past environments Drawings, images, charts, and graphs that can be used for paper-and-pencil items Rock samples and fossil replicas that can be used for hands-on tasks Computer simulations showing changes in Earth structures over time that can be used as a part of interactive computer tasks (ICTs)	Data on radioactive isotope decay presented in graph form and used as a tool to determine the age of a rock sample given percentages of different elements in the rock Drawing of rock layers with the location of fossils from an extinct prehistoric animal, along with dates of volcanic rock layers—determined by radioisotope methods above and below the fossils—used to give upper and lower bounds of time during which the extinct animal lived Illustrations of rock sequences, drawings of index fossils, and computer simulations of radioactive decay (can be used to assess understanding of how to measure geologic time) Charts of long time spans (can be used to assess understanding of the sequence and approximate timeframe for changes in Earth history)

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
History of Earth: Technical Vocabulary		
[No special vocabulary]	Common terms in geology (e.g., sedimentation) and paleontology (e.g., fossils) Exclusions: • Names of extinct animals identified by their fossils, except for the most common such as trilobites and better-known dinosaurs • Names of geologic eras, periods, and epochs • Names of the processes (focus should be on the processes themselves). For example, students should be expected to understand and apply but not name "Law of Superposition." • Names of physical mechanisms that caused past environmental changes	Technical vocabulary associated with geology (e.g., index fossils) as well as relevant terms from life sciences (e.g., bacteria, algae, extinct) and chemistry (e.g., radioisotopes, half-life) Names of eras (Pre-Cambrian, Paleozoic, Mesozoic, and Cenozoic) but students are not expected to recall these from memory Exclusions: • Names of specific life forms (e.g., stromatolites) • Geologic periods (e.g., Cambrian, Carboniferous) and epochs • Naming the processes of radioactive decay
History of Earth: Clarification		
Simple changes are noted easily in students' daily lives. Other changes are less noticeable or are observed only occasionally but are still appropriate for assessment at this grade level. (See "Examples" category above.) Some changes in Earth's surface take place rapidly and others take place slowly in terms of a human lifetime. (Note that emphasis should be on changes that take place within a human lifespan, such as decay of wood buildings, erosion, landslides, volcanoes.) See Appendix H for elaboration of E4.3. Exclusions: • Changes not easily observable that take place over very long stretches of time (e.g., Grand Canyon formation, tectonic mountain building, soil development)	Students better able to grasp large spans of time Evidence (e.g., fossils, rock sequences) that enable scientists to "read" Earth history Note connections between this subtopic and the Life Science subtopic, "Evolution and Diversity."	Key idea: Evidence for measuring geologic time and determining a sequence of events. Revisit mechanisms that build up and tear down Earth structures, sources of evidence, and vast span of geologic time. Note connections between this subtopic and the Life Science subtopic, "Evolution and Diversity."

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Properties of Earth Materials: Examples, Observations, and Phenomena		
Common Earth materials related to fuels (e.g., oil, wood, coal); building materials (e.g., granite, marble); materials related to plant and animal nutrition (e.g., air, water, minerals); metals (e.g., rocks, ores) Shelter for animals (e.g., trees, bushes, soil, caves) General temperature conditions under which water is solid (ice) or gas (water vapor) Soil as a medium for plant growth <u>Exclusions:</u> • Carbon dioxide and other specific component gases of the atmosphere • Precise temperatures for when liquid water freezes or becomes water vapor	Evidence of how rocks were formed: layers in sedimentary rocks, small crystals in granite indicate slow cooling and crystallization, glassy texture of obsidian indicates rapid cooling, and fossils frequently found in coal indicate formation from decay of once-living matter Igneous (e.g., granite, basalt), metamorphic (e.g., marble, slate), and sedimentary (e.g., sandstone, limestone) rocks Metamorphic rocks identified by evidence that one type of rock was subjected to heat and pressure, such as layers of sedimentary rock that have been folded and twisted Gases in the atmosphere: mainly nitrogen, some oxygen, small amounts of water vapor, and traces of other gases Layers in the atmosphere: bottom layer (troposphere) has all of the clouds, weather, and most of the atmosphere's matter; air is cold and still in next layer up (stratosphere) Trends in temperature and pressure: with increasing elevation in atmosphere, temperature decreases and air pressure decreases Components of soil: rock fragments, sand, clay, decaying plant parts and other organic material, water, and/or other materials <u>Exclusions:</u> • Names of intrusive magma bodies (e.g., batholith) • Names of extrusive lava bodies (e.g., sills, dikes) • Names of atmosphere layers in addition to troposphere and stratosphere (e.g., thermosphere, ionosphere)	[No content statements at this grade]

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Properties of Earth Materials: Instruments, Measurement, and Representations		
Classifying materials as either Earth or human-made materials (Note that this may be complex since some examples could fall in both categories. Materials used in items should fall clearly in one category or the other.) Qualitative descriptions of how Earth materials (e.g., bricks, wood, glass) are used in houses Simple bar graphs and tables, pictures, drawings, and diagrams Simple maps with keys, legends	Quantitative measures of mineral hardness using Moh's Hardness Scale Quantitative measures of soil porosity and permeability Quantitative measures of sediment size using sieves Litmus paper test of soil acidity/alkalinity Qualitative descriptions of instruments that can be used to separate Earth materials (e.g., sieves to separate soil particles of various sizes) Topographic, geologic and physical geographic maps Pie charts, bar graphs Drawings of stratigraphic sequences	[No content statements at this grade]
Properties of Earth Materials: Technical Vocabulary		
Generally limited to common natural materials (e.g., sand, clay, pebbles, oil, stone, wood, grasses, rocks, minerals) Generally limited to common human-made materials (e.g., bricks, tiles, cement, aluminum poles, nails, wire)	Sedimentary rock (sandstone, limestone, shale, coal, salt); igneous rock (granite, basalt, pumice, ash); metamorphic rock (marble, gneiss, schist, coal) Ores: iron, zinc, uranium, aluminum, copper, coal Soil layers, weathering, erosion, climate Soil textures (silt/loam/sand); soil acidity (acidic/alkaline) <u>Exclusions:</u> • Ores: rare earth (e.g., titanium), alloys (e.g., bronze, steel, pewter, ceramics) • Soils: names of soils	[No content statements at this grade]

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Properties of Earth Materials: Clarification		
Distinguishing Earth materials and their properties from human-made materials How humans and other animals use Earth materials for food, shelter, and tools The living and non-living parts of an ecosystem working together See Appendix H for elaboration of E4.4, E4.5, and E4.6.	More in-depth physical and chemical properties of Earth materials and the atmosphere. Evidence in rocks, minerals, and soils can be used to "tell a story" about the environment of formation. Maps are tools for interpreting "rock stories." Knowing that layered rock is formed under conditions of lower heat and pressure (which exist closer to the surface) than rock that contains crystals; existence of crystalline rock at the surface is considered evidence that such rock was pushed up to the surface after being formed far below. Exclusions: • Core physics and chemistry concepts (e.g., chemical reactions, stress and strain)	[No content statements at this grade]

Grade 4	Grade 8	Grade 12
Tectonics: Examples, Observations, and Phenomena		
[No content statements at this grade]	Common geographic structures related to plate boundaries (e.g., Pacific ring of fire, deep-sea trenches, volcanoes with subduction plate boundaries) Earth surface features that represent evidence for Continental Drift (e.g., matching continent shapes, mid-ocean ridges, matching rock formations on different continents or on either side of spreading zones) Locations of major mountain chains on geographic maps revealing subduction zones Qualitative descriptions of how mountains are uplifted when plates collide or when volcanoes result where one plate slides under another Alignment of compass needle with Earth's magnetic field	Distinction between less dense continental crust and more dense oceanic crust, and what happens when they collide Historical events, evidence, and arguments for continental movement ("fit" of continents on either side of the Atlantic, continuity of rock and fossil formations across continents) and sea floor spreading (mid-ocean ridges, similar rock depositions on either side of the ridge) Magnetic properties of rocks on the ocean floor as strong evidence in support of sea-floor spreading Mountain-building at the collision of two continental plates Discrepancy between oldest age of ocean floor rocks and continental rocks

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Tectonics: Instruments, Measurement, and Representations		
[No content statements at this grade]	Use of seismometers and quantitative measures of earthquake intensities (e.g., Mercalli and Richter scales) Maps of ocean bottom and continental features associated with plate boundaries Earthquake maps to observe patterns of earthquakes Labeled models/representations of Earth's interior and Earth's magnetic field Maps of locations of active volcanoes	Calculating rate of plate motion using age of rock and distance from spreading ridge Topographic maps to construct elevation profile Drawings, models, or map projections to show earthquake foci and their relationship to plate boundaries Maps, models, or map projections to locate tectonic plates Geologic maps to locate and interpret rock formations and rock sequences resulting from tectonic plate interactions Maps and models to align ores with tectonic activity Maps of the ocean floor revealing the mid-ocean ridge and magnetic orientation of rocks on both sides of the mid-ocean ridge Maps showing movement of major plates since Pangaea
Tectonics: Technical Vocabulary		
[No content statements at this grade]	Names of major continents and oceans; names of plate boundaries (e.g., mid-ocean ridge, subduction zone) Richter and Mercalli scales	Convection, sea floor spreading, continental drift, tetonic theory

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Tectonics: Clarification		
[No content statements at this grade]	Correlation among plate boundaries, earthquakes and volcanoes, and mountain ranges Reading and interpreting physical, topographic, and geologic maps See Appendix H for elaboration of E8.9.	Convection currents in the mantle drive plate movements. Tectonic theory explains both sea-floor spreading and continental drift. Key idea: Historical and contemporary models, mechanisms, and theories associated with tectonics; evidence for the support of the models, mechanisms, and theories See Appendix H for elaboration of E12.8.

Grade 4	Grade 8	Grade 12
Energy in Earth Systems: Examples, Observations, and Phenomena		
Effects of sunlight on Earth: sunlight causes evaporation of water; sunlight melts snow and ice; sunlight is needed for plant growth Earth being warm at equator and cold at poles	Sunlight penetrating the upper layer of the ocean where it provides energy for plant and algae growth Warmer air/water rises; colder air sinks; air colder at higher elevations Sea breezes and land breezes, driven by differences in how water and land absorb energy from the sun Changes associated with seasons: seeing different constellations at different times of year; changes in length of day and average temperature during the year	Phenomena on the surface of Earth driven by the sun's energy, including ocean currents and winds, and the growth of plants and animals Decay of radioactive materials and gravitational energy left over from formation of Earth as sources of Earth's internal energy Phenomena on the surface of Earth driven by Earth's internal energy, including the movement of tectonic plates, which, in turn, causes earthquakes and volcanoes and builds mountains

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Energy in Earth Systems: Instruments, Measurement, and Representations		
Temperature measures (e.g., of bottled water—one left in sunlight, one left in shade) Location on map projections of hot and cold regions of Earth Pictures and diagrams of simple experiments (e.g., relationship between energy from the sun and temperature of soil)	Measurement of the length of shadows over intervals of months Measurement of the location of sunrise and sunset on the horizon Drawings of convection cells: warm air and water rise, cold air and water sink Drawings of the Earth–sun model: revolution, rotation, tilt-of-axis Drawing showing the position of Earth at the summer and winter solstices and spring and fall equinoxes Maps of major atmospheric and oceanic currents Graphs of mean monthly temperatures showing how temperature varies during the year at different latitudes and in different hemispheres <u>Exclusions:</u> • Precession of the equinoxes • Coriolis effect	Models and diagrams showing the flow of energy from the sun through Earth systems and resulting phenomena Models and diagrams showing the movement of Earth's internal convection currents and resulting phenomena.
Energy in Earth Systems: Technical Vocabulary		
Evaporation, light, heat, warm, cold, temperature, thermometer, absorb	Equinoxes, solstices, incoming solar radiation, mean annual temperatures	Related terms in the "Tectonics" subtopic

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Energy in Earth Systems: Clarification		
Key idea: Energy from the sun makes possible life on Earth. See pp. 105-106, "Energy Sources and Transfer" and "Uses, Transformations, and Conservation of Energy," for more on the crosscutting nature of energy.	Energy from the sun drives all surface phenomena, including winds and ocean currents. Seasons arise due to the orientation of Earth's axis with respect to the sun. Earth's axis is tilted 23.5^{o} from the plane of its orbit around the sun. Since the axis stays pointing in the same direction as it circles the sun, the Northern Hemisphere leans toward the sun during one half of the year and away from the sun during the other half of the year. When the Northern Hemisphere leans towards the sun, longer and hotter days (summer) are experienced than when it leans away from the sun when shorter and cooler days (winter) are experienced. Uneven heating causes a convection current to develop. The resulting moving air is experienced as wind. Sea and land breezes can be used as context to assess understanding of forces that drive winds. Note connections between this subtopic and content statements P8.13 and L8.4. See pp. 105-106, "Energy Sources and Transfer" and "Uses, Transformations, and Conservation of Energy," for more on the crosscutting nature of energy. See Appendix H for elaboration of E8.11 and E8.12.	Earth has both internal and external energy sources; each of these sources drives different phenomena. Note connections between this subtopic and content statements P12.14, P12.15, P12.16, L12.4, L12.5, L12.6, and E12.12. See pp. 105-106, "Energy Sources and Transfer" and "Uses, Transformations, and Conservation of Energy," for more on the crosscutting nature of energy.

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

<table>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td colspan="3">Climate and Weather: Examples, Observations, and Phenomena</td></tr>
<tr>
<td>Data collected with simple weather instruments

Changes in weather conditions during the day, from day to day, and over seasons</td>
<td>Global wind patterns

Large scale systems: storms, hurricanes, warm fronts, cold fronts

Changes in local weather caused by fronts and storms

<u>Exclusions:</u>
Names of convection processes such as Hadley Cells, Coriolis Effect, or Conveyor Belt theory</td>
<td>Application of energy transfer concepts to weather and climate phenomena (e.g., how local topography affects weather; how greenhouse gases “trap” infrared energy; how Earth’s rotation affects global wind patterns and storm systems through Coriolis Effect)

Major weather and climate phenomena: the jet stream, high and low pressure areas, and phenomena related to weather fronts

<u>Exclusions:</u>
Going beyond application of physics principles (i.e., students should not be expected to be meteorologists)</td>
</tr>
</table>

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Climate and Weather: Instruments, Measurement, and Representations		
Images of or actual weather instruments (thermometers, weather vanes, anemometers, rain gauges); use of these simple instruments to measure weather conditions Measurable quantities to describe weather (e.g., temperature, wind direction, wind speed, rainfall) Interpretation of bar graphs (histograms), where the x-axis is hours or days and the y-axis is temperature or rainfall Illustrations or pictures of different weather conditions over hours, days, or months (e.g., cloud cover, observable changes in wind speed) Charts or bar graphs reflecting variations in annual temperature measurements to show seasons <u>Exclusions:</u> • Use of barometers to measure pressure or instruments to measure humidity (both are weather conditions that are harder for 4th graders to visualize and experience) • How weather differs in different climate zones	Computer animations illustrating global patterns (animations can be used to stimulate questions) Reference to weather instruments (even though not explicitly referenced in the 8th grade content statement), including those for 4th grade, plus barometers to measure atmospheric pressure Diagrams of weather systems and/or satellite images of Earth (e.g., in relation to storm systems) Interpretation of weather maps, which generally show weather patterns intermediate between global and local scales	Data on past climates and levels of greenhouse gases provided by ice cores Manipulation of computer simulations to demonstrate how physics principles (and especially energy transfer concepts) help to explain climate (e.g., manipulating factors such as fossil fuel use to mitigate future global warming) Computer simulations to show how geological features result in microclimates Energy flow from the sun through the atmosphere (possible context for assessing understanding of both weather and climate) Interpretation of graphs, charts, and maps, both in print and computer displays
Climate and Weather: Technical Vocabulary		
Common vocabulary related to weather instruments, weather conditions, and units, such as thermometer (temperature in °C), rain gauge (amount of rain in cm), weather vane (direction of wind using compass points) <u>Exclusions:</u> • Terms related to pressure, humidity, and long-term climate trends	Terms relevant to this subtopic, including various names of weather instruments, conditions, and phenomena (e.g., high and low pressure areas, warm and cold fronts); use of the most common terms <u>Exclusions:</u> • Regarding drivers of weather and its effects, avoid technical terms where possible	Technical terms related to both the physics principles and weather and climate phenomena, although emphasis should be on understanding the processes rather than knowing the names of the processes

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Climate and Weather: Clarification		
Observations and direct measurements of weather conditions locally <u>Exclusions:</u> • Relating weather to long-term climate changes	Global weather patterns and how these affect local weather Uneven heating causes global wind patterns as air rises over the tropics and falls closer to the poles, forming huge convection (Hadley) cells. Movements of winds across the surface of the globe cause rotating weather systems to develop. These large-scale systems bring changes in local weather. Extensive worldwide ocean currents (Conveyor Belt), driven by convection due to changes in temperature and salinity play an important role in climate. <u>Exclusions:</u> • Quantitative analysis of energy flow through Earth systems	Application of principles from Life Science and Physical Science to the understanding of weather and long-term climate changes Key idea: How energy moves through Earth systems

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

<table>
<tr><th>Grade 4</th><th>Grade 8</th><th>Grade 12</th></tr>
<tr><td colspan="3">Biogeochemical Cycles: Examples, Observations, and Phenomena</td></tr>
<tr>
<td>Limitations of specific resources: fuels, metals, fresh water, and farmland

Resources recycled by humans (e.g., newspapers, glass, cans, rubber, plastics, most metals)

Resources that humans can reuse rather than discard (e.g., jars, bottles, computer components, used cars, other manufactured goods)

Some materials can be renewed (e.g., allowing farmland to lie fallow)

Sources of fresh water: precipitation, rivers, lakes, wells (ground water)

Reservoirs and waste treatment plants</td>
<td>Conditions that result in evaporation and condensation of water, both in the lab and in the field

Evaporation of salt water: leaving behind salt, therefore rain falling as fresh water

Sources of fresh water: precipitation, rivers, lakes, wells (ground water)

Tracts of land set aside as parks and preserves for wildlife and water conservation areas

Water pollutants: domestic, industrial, and farm sewage, which can affect pH of water, dissolved oxygen, microbes, bacteria

Human and natural effects on water flow:
• Dams alter river channels and floodplains, which can have adverse effects on plants and animals.
• Trees, bushes, and grasses help prevent erosion. Clearing land may have adverse effects on plants and animals as well as downstream where sediment is deposited.
• River and stream channeling and the construction of levees and dikes can have adverse effects on floodplains and deltas.

Examples of human activity on flood plains, wetlands, forests, watersheds, estuaries, beaches, etc. and human effects on plant and animal populations

<u>Exclusions:</u>
• Chemistry or physics of absorption of infrared radiation</td>
<td>Examples of cycles of different elements and the energy sources that drive them

Different chemical forms of common elements (e.g., carbon is a gas when combined with oxygen to form carbon dioxide, a liquid in carbonic acid or petroleum, or a solid in plant tissue or limestone)

How elements move through Earth systems (e.g., when carbon dioxide gas combines with rainwater to form natural acid rain, ending up in rivers and oceans)

Carbon and other elements injected into the air by volcanoes

Solar energy evaporating water, which condenses into rain that removes carbon dioxide and other substances from the air

Solar energy evaporating water from oceans, increasing salinity, which drives ocean currents

Short-term carbon cycle: recycling of atmospheric gases needed by plants and animals to survive

Natural environments disposing of plant and animal wastes through decomposition

Physical processes (e.g., weathering, erosion, deposition) that provide inorganic constituents of soils

Decomposition processes that provide organic constituents of soils</td>
</tr>
</table>

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Biogeochemical Cycles: Instruments, Measurement, and Representations		
Scales, rulers, and other measuring devices to determine how much of something Thermometers to monitor change in weather conditions Observing and recording individual animal behavior to see change over time Representations of changes as cyclic (circular) or linear (with a beginning and end) Images showing how people use Earth materials and how they may be recycled, reused, or renewed Distinguishing natural from human-made materials Images of human activities changing natural conditions that may lead to loss of habitats (lakes) or loss of species (fish, birds) Matching organisms to ecosystem, habitat, or niche (e.g., worms with soil, fish with lakes, ponds; frogs with ponds, land, trees)	Rain gauges, thermometers, etc. to measure and record attributes of an environment (e.g., patterns and changes in mean minimum and maximum monthly temperatures, rainfall, cloud coverage) Sampling methods to monitor and describe changes in animal populations and correlate with changes in the environment caused by human activities Diagrams illustrating differences between cyclic and linear changes Maps, diagrams, or charts showing populations of selected plants and animals	Appropriate chemical tests and instruments to monitor quality of soil, water, and air and the health of ecosystems (e.g., wetlands, deserts, plains, rainforests) Illustrations of chemical and physical transformations of matter and material as they move through Earth systems Diagrams or animations of natural processes that produce useful substances (e.g., fossil fuel formation) Models of matter (e.g., carbon, oxygen, nitrogen) and energy (e.g., thermal, chemical) moving through and across boundaries of Earth systems Maps of population boundaries
Biogeochemical Cycles: Technical Vocabulary		
Fallow	Precipitation, hail, physical properties, pollutants, pH, dissolved oxygen	Decomposition

Table 13. Earth and Space Science Content Boundaries for Grades 4, 8, and 12 (cont.)

Grade 4	Grade 8	Grade 12
Biogeochemical Cycles: Clarification		
Human dependence on natural and constructed environments Key idea: Renewable and non-renewable resources and materials Through the use of various technologies over the centuries, humans have impacted environments. Humans have learned and continue to learn how to coexist with natural environments. See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of the principles in this subtopic.	Key idea: Water systems and their relation with other Earth systems (biosphere, atmosphere, lithosphere) Human activities impacting the Earth systems on local, regional and global scales; some impacts can be reversed; others result in permanent changes (e.g., extinction of species, soil clogging of river harbors, soil contamination from hazardous waste dumping) See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of the principles in this subtopic. See Appendix H for elaboration of E8.15.	Key idea: Interaction of Earth systems on a global scale with emphasis on a hierarchy of systems and tracing processes within systems (e.g., carbon moving through rock, soils, plants, animals, liquids, and gases) Natural ecosystems provide "services" essential for human life to exist, such as drinkable water, breathable air, and fertile soil for growth of plants. Oceans and continents are dynamic interacting systems that help to maintain Earth's climate by transporting thermal energy and dissolved gases. Note connections between this subtopic and content statements P12.14, P12.15, P12.16, L12.4, L12.5, L12.6, and E12.9. See p. 107, "Biogeochemical Cycles," for more on the crosscutting nature of the principles in this subtopic. See Appendix H for elaboration of E12.13.

Distribution of Items by Content Area

The distribution of items at each grade level by the three science content areas, as measured by percentage of student response time, is shown in Table 14. The relative emphasis of science content should be represented on the assessment as a whole. In the interest of maintaining valid and reliable measures of trends over time, the assessment should be built so that each content area is represented from one administration to the next; that is, the developer should adequately sample the content areas to avoid item pools that under- or over-emphasize particular content.

Table 14. Distribution of Items by Content Area and Grade

	Grade 4 (% student response time)	Grade 8 (% student response time)	Grade 12[17] (% student response time)
Physical	33.3	30.0	37.5
Life	33.3	30.0	37.5
Earth and Space	33.3	40.0	25.0

Crosscutting Content

Scientists define their specializations narrowly (e.g., astronomy, molecular biology, organic chemistry) to organize their research communities; and the categories of Physical Science, Life Science, and Earth and Space Science are helpful for organizing school science. These categorizations mask the fact that some science principles cut across the content areas. In the *Framework* and *Specifications,* crosscutting content is not represented by abstractions such as “models,” “constancy and change,” or “form and function,” but is anchored in the content statements themselves. Some examples of crosscutting content are described in this section: “Energy Sources and Transfer”; “Uses, Transformations, and Conservation of Energy”; and “Biogeochemical Cycles.” Such examples illustrate opportunities for assessing specific content in greater depth.

The *Specifications* does not prescribe the proportion of the assessment to be devoted to crosscutting content. Nor is every potential crosscutting content specified here; hence, assessment developers have latitude in identifying and defining additional examples of crosscutting content. Crosscutting content is intended to generate sets (clusters) of items. Each item in a given crosscutting content cluster should be assigned a single content area classification and be reported on only one of the content subscales. Thus, a crosscutting content cluster would be comprised of items that, taken together, are classified under more than one content area. The intent is not to generate individual items, each of which has a score that is assigned to more than one content subscale.

---

[17] These recommendations are based on NAEP data regarding students’ 12th grade course-taking patterns. If these patterns change materially after 2009, these recommendations should be reconsidered.

Since some content statements are themselves crosscutting in nature, that is, they may fall into more than one content area, it is expected that items developed from such content statements may also be crosscutting in nature. However, each item should be assigned a primary content area classification and be reported on only one of the content subscales.

There should be at least one cluster of crosscutting content items at each grade level in each assessment year. However, the nature of the crosscutting content may vary among grade levels and from one administration year to the next.

Energy Sources and Transfer

When a sufficiently high temperature due to gravitational attraction occurs in the sun, nuclear reactions take place. These reactions result in the transfer of energy from nuclei to their surroundings. At the same time, those high temperatures cause the sun to radiate visible light and many other forms of electromagnetic waves. A small fraction of this light energy reaches Earth, heating the land, air, and water. Some of this energy causes some water to evaporate. The water vapor travels high into the atmosphere, thus increasing its gravitational potential energy. There it cools and condenses, some of it falling into reservoirs behind dams. At many dams, some of this water is directed downhill through tubes, resulting in the transfer of gravitational potential energy to the descending water as kinetic energy. This water is then used to turn turbines, and energy is thus transferred from the moving water to electrical appliances through circuits and power lines. Accordingly, the energy used to power something as commonplace as a light bulb, TV, radio, or stereo can be traced back to nuclear reactions deep inside the sun.

Consider also the transfer of energy that occurs as a diver falls through air into water. When the diver is initially poised on a cliff high above a lake's surface, one says that the diver has potential energy with respect to the air and water below. As the diver falls, her speed (kinetic energy) increases as her potential energy decreases. Her body transfers energy to the medium through which she falls, that is, the diver's body rubs against the air and water (heating them) and exerts force on the air and water (moving them aside). When the diver finally comes to rest in the lake, some or most of her potential energy has been transferred to heating and setting into motion the air and water through which she fell.

The following grade 12 content statements illustrate the crosscutting nature of energy sources and transfer. They are not intended to represent an exhaustive catalog of all statements related to this crosscutting content.

Physical Science	Earth and Space Science
P12.11: Fission and fusion are reactions involving changes in the nuclei of atoms. Fission is the splitting of a large nucleus into smaller nuclei and particles. Fusion involves joining of two relatively light nuclei at extremely high temperature and pressure. <u>Fusion is the process responsible for the energy of the sun and other stars.</u>[18]	**E12.9:** Earth systems have internal and external sources of energy, both of which create heat. <u>The sun is the major external source of energy.</u> Two primary sources of internal energy are the decay of radioactive isotopes and the gravitational energy from Earth's original formation.

[18] Underlining is used here and on pp. 106-107 to link segments of the content statements across content areas.

Uses, Transformations, and Conservation of Energy

The principles of energy uses, transformations, and conservation hold true across different types of systems. These systems include biological organisms, Earth systems, ecosystems (combining both life forms and their physical environment), the solar system, other systems in the universe, and human-designed systems.

From *Science for All Americans* (AAAS, 1994, p. 66):

> However complex the workings of living organisms, they share with all other natural systems the same physical principles of the conservation and transformation of matter and energy. Over long spans of time, matter and energy are transformed among living things, and between them and the physical environment. In these grand-scale cycles, the total amount of matter and energy remains constant, even though their form and location undergo continual change.
>
> Almost all life on earth is ultimately maintained by transformations of energy from the sun. Plants capture the sun's energy and use it to synthesize complex, energy-rich molecules (chiefly sugars) from molecules of carbon dioxide and water. These synthesized molecules then serve, directly or indirectly, as the source of energy for the plants themselves and ultimately for all animals and decomposer organisms (such as bacteria and fungi). This is the food web: The organisms that consume the plants derive energy and materials from breaking down the plant molecules, use them to synthesize their own structures, and then are themselves consumed by other organisms. At each stage in the food web, some energy is stored in newly synthesized structures and some is dissipated into the environment as heat produced by the energy-releasing chemical processes in cells. A similar energy cycle begins in the oceans with the capture of the sun's energy by tiny, plant-like organisms. Each successive stage in a food web captures only a small fraction of the energy content of organisms it feeds on.

The flow of energy in an ecosystem (such as that described above) can be compared to the flow of energy illustrated on the previous page. They are both identical (the principle) and different (the context). In each case, energy is transformed from one form to another; and while some is no longer available for human use, it is not lost to the system.

The following grade 4 content statements illustrate the crosscutting nature of uses, transformations, and conservation of energy. They are not intended to represent an exhaustive catalog of all statements related to this crosscutting content.

Physical Science	Life Science	Earth and Space Science
P4.7: Heat (thermal energy), electricity, <u>light,</u> and sound are forms of <u>energy.</u>	**L4.2:** Organisms have basic needs. Animals require air, water, and a source of energy and building material for growth and repair. <u>Plants also require light.</u>	**E4.7:** <u>The sun</u> warms the land, air, and water and <u>helps plants grow.</u>

Biogeochemical Cycles

To demonstrate an understanding of biogeochemical cycles, students must draw on their knowledge of Matter and Energy (Physical Science), Structures and Functions of Living Systems (Life Science) and Earth Systems (Earth and Space Science).

Essentially fixed amounts of chemical atoms or elements cycle within the Earth system, and energy drives their translocation and transformation. Examples of biogeochemical cycles include water, carbon, and nitrogen. The basic processes underlying the translocation of matter (e.g., changes of state, gravity) and transformations involving the rearrangement of atoms in chemical reactions are described in Physical Science (p. 32) and the role of living organisms in cycling atoms between inorganic and organic forms is described in Life Science (p. 57).

Biogeochemical cycles are described more fully in the "Earth Systems" section of Table 12, Earth and Space Science Content Statements for Grades 4, 8, and 12 (p. 83).

The following grade 12 content statements illustrate the crosscutting nature of biogeochemical cycles. They are not intended to represent an exhaustive catalog of all statements related to this crosscutting content.

Physical Science	Life Science	Earth and Space Science
P12.7: A large number of important reactions involve the transfer of either electrons (oxidation/reduction reactions) or hydrogen ions (acid/base reactions) between reacting ions, molecules, or atoms. In other chemical reactions, atoms interact with one another by sharing electrons to create a bond. <u>An important example is carbon atoms, which can bond to one another in chains, rings, and branching networks to form, along with other kinds of atoms—hydrogen, oxygen, nitrogen, and sulfur— a variety of structures, including synthetic polymers, oils, and the large molecules essential to life.</u>	**L12.5:** <u>The chemical elements that make up the molecules of living things pass through food webs and are combined and recombined in different ways.</u> At each link in an ecosystem, some energy is stored in newly made structures, but much is dissipated into the environment as heat. Continual input of energy from sunlight keeps the process going.	**E12.12:** Movement of matter through Earth's systems is driven by Earth's internal and external sources of energy. These movements are often accompanied by a change in the physical and chemical properties of the matter. <u>Carbon, for example, occurs in carbonate rocks such as limestone, in coal and other fossil fuels, in the atmosphere as carbon dioxide gas, in water as dissolved carbon dioxide, and in all organisms as complex molecules that control the chemistry of life.</u>

Additional Suggestions

Other suggestions for crosscutting content are biochemistry for grade 12 and evolution and plate tectonics for grade 8 or 12. These are summarized below.

Biochemistry

In “Changes in Matter” (p. 32) in Physical Science at 12th grade, there is mention of large molecules essential to life. At the same grade in Life Science, there is mention of proteins in “Organization and Development of Living Systems” (p. 56) and proteins, DNA, and carbohydrates in “Matter and Energy Transformations” (p. 57). The chemical nature of biomolecules is important to biological systems and may represent crosscutting content. In the 21st century, the area of biochemistry will be one of the driving forces for our economy and for our understanding of how biological systems work. Using biochemistry as an opportunity to address science concepts in an interdisciplinary manner would be not only productive but also reflective of scientific issues germane to the research community and to society.

Evolution and Plate Tectonics

One of the central features of speciation is the division of an original population into subpopulations by geographic barriers. Where do these barriers come from? Mountain ranges, valleys, and oceans are the direct result of tectonic events, such as plate separation or collision. These produce not only separate but also distinct changes in environment and the gradients between each environment. For example, divergent margins, such as the east coast of North America, have broad, shallow sloped continental shelves, producing not only a wider range of environments but also energy regimes for coastal processes. The latitude of land masses and oceans, as well as the proximity of a given location to the ocean or land mass, directly determine the type of physical environment. Hence, understanding the position of the continents today allows us to define environments to which species adapt. Knowledge of past plate positions allows a description of the past environment in a location and hence the environmental pressures on a species. Fossil evidence correlated across ocean basins, barriers, or large distances allows the determination of the timing of species divergence.

Fossils and paleoenvironmental factors (i.e., factors associated with ancient environments in the geologic past) also allow for the identification and prediction of transitional forms of species, many of which have been verified by new fossil finds. For example, evolutionary theory predicts that present-day Cetaceans (whales) are descendants of early large land mammals that spent much time in shallow, coastal waters. Transitional forms of proto-whales, possessing characteristics of both whales and land mammals, should be found in shallow marine, coastal sediments of a particular age. Such fossils, predicted by theory, have in fact been found within the last 15 years, in exactly the age of sediment predicted and in the region where such sediments were deposited at the time of the transitional form’s existence. When correlated with geochemical data (e.g., isotopic ratios of oxygen) and geophysical data (e.g., paleomagnetic orientations), there are multiple points of convergence exactly where both evolution and plate tectonics predict them to be.[19] Thus, the evolution of the surface of the Earth and the evolution of life are inextricably linked.

From Science Content to Science Practices

This chapter has presented the science content that defines the NAEP Science Assessment content domain. The content statements, as presented in this chapter, do not describe students' performances in observable terms. The next chapter, Chapter Three, will describe science practices and cognitive demands. Chapter Four will bring Chapters Two and Three together by showing how science content statements can be combined ("crossed") with science practices to generate performance expectations (i.e., descriptions of students' expected and observable performances on the NAEP Science Assessment). Based on these performance expectations, assessment items can be developed, and then, finally, inferences can be derived from student responses about what students know and can do in science. Chapter Four will provide illustrative examples of this process.

[19] Minerals bearing iron and nickel take on the signature of the Earth's magnetic field at the time the rocks were formed. As the Earth's magnetic field changes intensity and polarity, the rock retains the signature of the ancient field. As the position of the rock changes (i.e., the plate drifts), the north-pole direction imprinted in the rock continues to "point" to the ancient north magnetic pole. This is a primary geophysical indicator of past, ancient plate locations.

CHAPTER THREE: SCIENCE PRACTICES

Introduction

Chapter Two presented content statements that define the key science principles (as well as the facts, concepts, laws, and theories) to be assessed by NAEP in 2009. However, NAEP will assess not only science content statements but also the ways in which knowledge is used. This chapter defines what students should be able to do with the science content statements by articulating key science practices to be assessed by NAEP—Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design. These practices are useful for generating science-rich assessment items.

To assist assessment developers, the science practices can be associated with the cognitive demands that they place on students. This chapter employs a set of four cognitive demands: (1) “knowing that,” (2) “knowing how,” (3) “knowing why,” and (4) “knowing when and where to apply knowledge.” These cognitive demands help ensure NAEP assessment items are developed so as to elicit the kinds of knowledge and thinking that underlie the performance expectations of the *Framework* and *Specifications*. They also provide a tool for interpreting students’ responses on the assessment items.

Three types of textboxes are used throughout this chapter. *Clarification* textboxes provide details on potentially confusing topics, such as the distinction between Identifying Science Principles and Using Science Principles. *Illustrative Item* textboxes provide assessment items that exemplify recommendations discussed in the text. Answers to selected-response illustrative items are indicated within the textbox; scoring guides for constructed-response illustrative items are provided in Appendix D.[20] *Item Suggestion* textboxes also serve to illustrate points made in the text, but these textboxes present item ideas that require further development in that they are neither published nor field-tested items. Answers to selected-response item suggestions are indicated within the textbox; however, scoring guides have not been developed for constructed-response item suggestions. Although items in both Illustrative Item and Item Suggestion textboxes may assess more than one content statement or practice, only the primary content and practice designations are provided. This follows NAEP practice, which uses only primary designations for items in the analysis and reporting of student responses.

Overview of Practices

Over the course of human history, people have developed many interconnected and validated ideas about the physical and biological world. These ideas have enabled successive generations to achieve an increasingly comprehensive and reliable understanding of the natural world. Scientific ideas are generated and verified by observing natural phenomena, finding patterns in these observations, and constructing

[20] See Table 27 in Appendix D for a complete list of illustrative items appearing in the *Specifications*.

theoretical models to explain these patterns.[21] These patterns and models can in turn be used to describe, measure, classify, explain, and predict other observations. Science knowledge is used to reason about the natural world and to improve the quality of scientific thought and action. Hence, NAEP will assess how well 4th, 8th, and 12th grade students can engage in the following broadly organized science practices:

- Identifying Science Principles
- Using Science Principles
- Using Scientific Inquiry
- Using Technological Design

Because these practices are closely related, these categories are not distinct and some overlap is expected. For example, in solving a problem involving a series circuit, students will need to *identify* a science principle (i.e., what a series circuit is) and *use* this principle to model such a circuit. If science practices are not conceptually independent, then assessment items written to one practice may actually tap more than one practice. This said, each item can be linked to a "predominant practice." At a particular grade level, it is expected that the item will tap primarily one practice for most students, even though this may not be the case for some (e.g., advanced) students.

Sources for the Development of Practices

The *Framework* and *Specifications* developers examined a number of sources to develop the short list of practices to be assessed in the NAEP Science Assessment. The most important were the "Science as Inquiry" sections of the *National Standards* and "Chapter 12: Habits of Mind" in *Benchmarks*. The committee also consulted the *National Standards* and *Benchmarks* sections on "Science and Technology" and "The Designed World," and the Validities of Science Inquiry Assessments project (Quellmalz, Haertel, DeBarger, & Kreikemeier, 2005). Conducted by SRI International during 2001-05, this project classified assessment items according to the inquiry standards discussed in the *National Standards*. The practices described below are found in most of the above sources. Cognitive research on science learning, international frameworks, and state standards were also used as reference points.

Communication

The ability to communicate accurately and effectively is essential in science, and this expectation is a strand that runs across the practices. Accurate and effective communication may include (but is not limited to) writing clear instructions that others can follow to carry out an investigation; reading and organizing data in tables and graphs; locating information in computer databases; using audio, video, multimedia, and other technologies to access, process, and integrate scientific findings; using language and

---

[21] Because natural phenomena are understood and described based on collected observations, the terms "phenomena" and "observations" are intricately intertwined. For ease of communication, the *Framework* and *Specifications* use the term "observations" to represent both specific observations of a natural phenomenon and the phenomenon itself.

scientific terms appropriately; drawing pictures or schematics to aid in descriptions of observations; summarizing the results of scientific investigations; and reporting to various audiences about facts, explanations, investigations, and data-based alternative explanations and designs (AAAS, 1993; NRC, 1996; Partnership for 21st Century Skills, 2004).

Quantitative Reasoning in Science

Quantitative reasoning is fundamental to science. Quantitative reasoning is the capacity not only to calculate (e.g., determine density given an object's mass and dimensions, calculate emissions from radioactive decay) but also to model a system (e.g., determine the energy released from a chemical reaction) (see, for example, Ayala, Shavelson, Yin, & Schultz, 2002; Hamilton, Nussbaum, Kupermintz, Kerkhoven, & Snow, 1995). Several of the science content statements inherently require quantitative reasoning. For example, content statement P12.20 describes acceleration in terms of the quantitative relationship $a=F_{net}/m$; content statement L8.7 assumes measures of population size and density; content statements E12.4 and E12.7 necessitate an understanding of the vast span of geologic time, that is, an understanding of order of magnitude. The boundary conditions set forth to clarify the content statements and performance expectations (Chapter Two) identify important opportunities for creating assessment items that tap quantitative reasoning in science.

Two caveats are in order. First, the quantitative reasoning and calculations assessed by NAEP Science should be directly relevant to the science content and performance expectation being tapped, and not simply a mathematical exercise. Second, students' time should be spent reasoning or applying an algorithm or formula. Their time should not be spent computing solutions to problems with large numbers and decimals, as hand-held calculators are not used in the NAEP Science Assessment. See pp. 48-50 for an example of priority given to qualitative or semi-quantitative understanding (e.g., understanding of mathematical relationships such as proportionality rather than formulaic use of equations) in boundary statements elaborating grade 12 content in "Forces Affecting Motion."

See p. 120 for an item that illustrates the use of quantitative reasoning in a context related to the practice of Using Scientific Inquiry.

Identifying Science Principles

This category focuses on students' ability to recognize, recall, define, relate, and represent basic science principles specified in the Physical Science, Life Science, and Earth and Space Science content statements presented in Chapter Two. The content statements themselves are often closely related to one another conceptually. Moreover, the science principles included in the content statements can be represented in a variety of forms, such as words, pictures, graphs, tables, formulas, and diagrams (AAAS, 1993; NRC, 1996). NAEP will assess students' ability to describe, measure, or classify observations; state or recognize principles included in the content statements; connect closely related content statements; and relate different representations of science knowledge. The practices assessed in this category draw on "declarative knowledge," or "knowing that," which is described in the "Cognitive Demands" section later in this chapter. Identifying Science Principles comprises the following general types of performance expectations:

- Describe, measure, or classify observations (e.g., describe the position and motion of objects; measure temperature; classify relationships between organisms as being predator/prey, parasite/host, producer/consumer).
- State or recognize correct science principles (e.g., "mass is conserved when substances undergo changes of state"; "all organisms are composed of cells"; "the atmosphere is a mixture of nitrogen, oxygen, and trace gases that include water vapor").
- Demonstrate relationships among closely related science principles (e.g., connect statements of Newton's three laws of motion, relate energy transfer with the water cycle).
- Demonstrate relationships among different representations of principles (e.g., verbal, symbolic, diagrammatic) and data patterns (e.g., tables, equations, graphs).

Identifying Science Principles is integral to all of the other science practices.

The following two items illustrate the expectation that students recognize correct science principles. The first item assesses students' ability to correctly identify simple information about the location of bodies within the solar system ("declarative knowledge"). More than half of the 8th graders answered it incorrectly. Thirty-five percent of the students thought that the moon is sometimes closer to the sun than to the Earth.

Illustrative Item

The Earth's Moon is

A. always much closer to the Sun than it is to the Earth.
B. always much closer to the Earth than it is to the Sun.
C. about the same distance from the Sun as it is from the Earth.
D. sometimes closer to the Sun than it is to the Earth and sometimes closer to the Earth than it is to the Sun.

Key: B

E8.1, Identifying Science Principles
Source: NAEP 2000, Grade 8

Illustrative Item

Animals and plants are made up of a number of different chemical elements. What happens to all of these elements when animals and plants die?

A. They die with the animal or plant.
B. They evaporate into the atmosphere.
C. They are recycled back into the environment.
D. They change into different elements.

Key: C

L8.5, Identifying Science Principles
Source: TIMSS 2003, Grade 8[22]

Using Science Principles

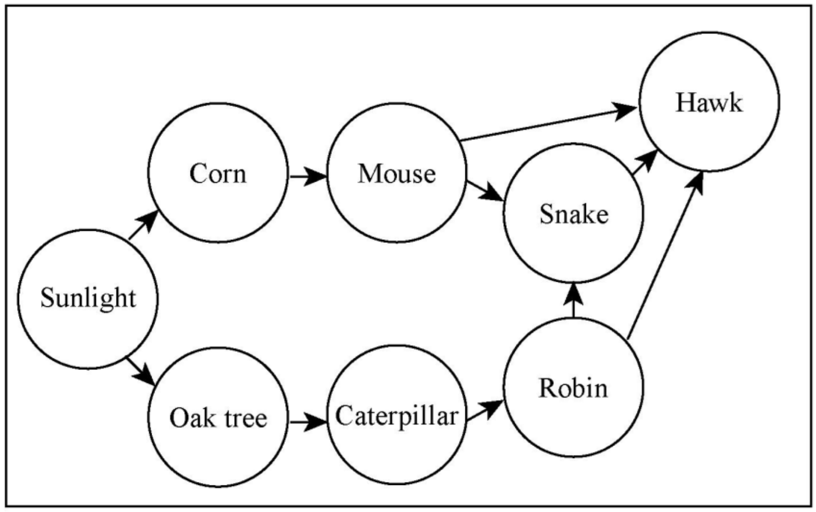
Scientific knowledge is useful for making sense of the natural world. Both scientists and informed citizens can use patterns in observations and theoretical models to predict and explain observations that they make now or that they will make in the future. The practices assessed in this category draw primarily on “schematic knowledge,” or “knowing why,” in addition to “declarative knowledge,” which are described in the “Cognitive Demands” section later in this chapter. Using Science Principles comprises the following general types of performance expectations:

- Explain observations of phenomena (using science principles from the content statements).
- Predict observations of phenomena (using science principles from the content statements, including quantitative predictions based on science principles that specify quantitative relationships among variables).
- Suggest examples of observations that illustrate a science principle (e.g., identify examples where the net force on an object is zero; provide examples of observations explained by the movement of tectonic plates; given partial DNA sequences of organisms, identify likely sequences of close relatives).
- Propose, analyze, and/or evaluate alternative explanations or predictions.

[22] TIMSS items appearing in the *Specifications* are copyrighted © by the International Association for the Evaluation of Educational Achievement (IEA).

The following item illustrates the expectation that students predict phenomena.

Illustrative Item



Look at the food web above. If the corn crop failed one year what would most likely happen to the robin population? Explain your answer.

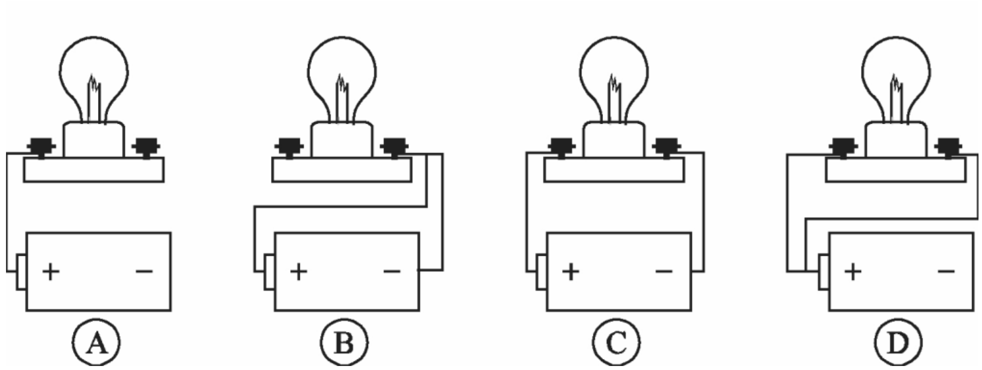
(See Appendix D for item scoring guides.)

L8.6, Using Science Principles
Source: TIMSS 1999, Grade 8

The first two science practice categories—Identifying Science Principles and Using Science Principles—both require students to correctly state or recognize the science principles contained in the content statements. A difference between the categories is that Using Science Principles focuses on what makes science knowledge valuable or in other words, its usefulness in making accurate predictions about phenomena and in explaining observations of the natural world in coherent ways (i.e., "knowing why"). Distinguishing between these two categories draws attention to differences in the depth and richness of individuals' knowledge of the content statements. Certain actions on the part of students lead to an inference of Identifying Science Principles, while other actions lead to an inference of Using Science Principles. Assuming a continuum from "just knowing the facts" to "using science principles," there is considerable overlap at the boundary. The line between the Identifying and Using categories is not distinct. Consider the following item, which illustrates the expectation that students connect different representations. In this case, the student must identify the correct pictorial representation of a complete electrical circuit.

Illustrative Item

The pictures show a lightbulb connected to a battery.
Which bulb will light?



Key: C

P4.11, Using Science Principles
Source: TIMSS 2003, Grade 4

Student responses to this item are open to two interpretations. If students have had a great deal of exposure to these types of circuit representations, their responses would fall under Identifying Science Principles. If, however, these circuit representations are relatively novel for students, then they would need to do more reasoning and their responses would fall under Using Science Principles. Reasoning with novel representations also brings to bear “strategic knowledge” or “knowing when and where to apply knowledge,” which is described later in the “Cognitive Demands” section of this chapter.

The following textbox provides further illustration of the distinction between *identifying* the boiling point of water (a fact) and *using* the relationship between boiling point and pressure (altitude) to explain or predict.

Clarification: Distinguishing between Identifying Science Principles and Using Science Principles—A Boiling Point Example

Grade 8 Content Statement: Matter—Properties of Matter:
P8.4: . . . Each element and compound has physical and chemical properties, such as boiling point, density, color, and conductivity, which are independent of the amount of the sample.

Distinguishing between the two categories of Identifying and Using Science Principles is a function of actions or performances. Using boiling point as an example, one might observe different responses to the question, “What is the boiling point of water?” Behaviors or actions might include:

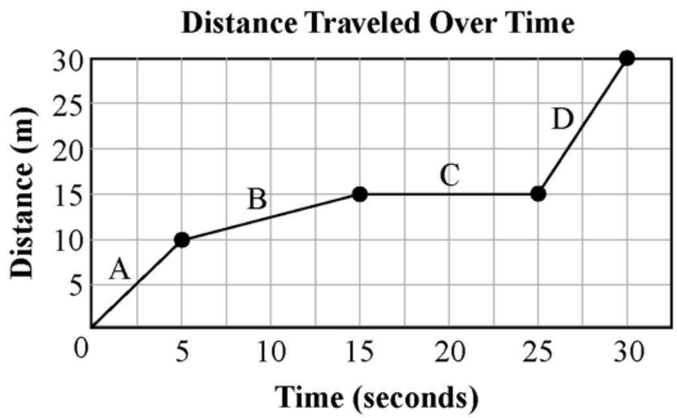
- Penciling in the oval corresponding to 100^0C in a selected-response item.
- Writing: “The boiling point of water is 100^0C at sea level.”
- Writing: “The boiling point of water changes as pressure changes. So, even though water boils at 100^0C at sea level (1 atm pressure), it might boil at a lower temperature on top of a mountain because pressure is lower up there.”

The above responses evoke different inferences about the science understanding of the individual responding. Both the first and second responses suggest that the question is only assessing knowledge of facts or the ability to identify a science principle; however, they illustrate the difference between recognizing a correct answer and retrieving that correct answer from memory. The third response contains even more sophisticated information, suggesting that the student can use a science principle to make predictions. Distinctions between these two categories can be clarified by examining student responses and conducting cognitive labs.

The following item set provides yet another example. The first item in the set (Item 7) illustrates Identifying Science Principles, since students are expected simply to read a distance/time graph. The second item in the set (Item 8), however, illustrates Using Science Principles, as students are required to apply the principle of speed to interpret the distance/time graph. Item 8 addresses the expectation that students interpret data presented in a graph and use the data to perform a mathematical calculation. (See p. 112 for more on quantitative reasoning in science.)

Illustrative Item

The graph below shows the distance traveled over time by a student walking down a hall. Use the information shown on the graph to do Numbers 7 and 8.



7. During which time interval was the student moving the fastest?

   - A
   - B
   - C
   - D

Key: D

P8.14, Identifying Science Principles

8. What was the average speed of the student from 0 seconds to 5 seconds?

   Average speed: ________________________________

(See Appendix D for item scoring guides.)

P8.14, Using Science Principles

Source: Colorado Department of Education, 2002, Grade 8

Using Scientific Inquiry

Scientists make observations about the natural world, identify patterns in data, and propose explanations to account for the patterns. While scientists differ greatly from one another in what phenomena they study and in how they go about their work, scientific inquiry involves the collection of relevant data, the use of logical reasoning, and the application of imagination and evidence in devising hypotheses to explain patterns in data. Scientific inquiry is a complex and time-intensive process that is iterative rather than linear. Scientists are also expected to exhibit, indeed to model, the habits of mind—curiosity, openness to new ideas, informed skepticism—that are part of science literacy. This includes reading or listening critically to assertions in the media, deciding what evidence to pay attention to and what to dismiss, and distinguishing careful arguments from shoddy ones. These critical thinking and systems thinking skills are the basis for exercising sound reasoning, making complex choices, and understanding the interconnections among systems (Partnership for 21st Century Skills, 2004). Thus, Using Scientific Inquiry depends on the practices described above—Identifying Science Principles and Using Science Principles. Moreover, in addition to involving “declarative” and “schematic knowledge,” Using Scientific Inquiry draws heavily on “procedural knowledge”—“knowing how” (e.g., knowing how to determine the mass of an object). NAEP Science focuses on a few key inquiry practices that are practical to measure in the NAEP Science Assessment. Using Scientific Inquiry comprises the following general types of performance expectations:[23]

- Design or critique aspects of scientific investigations (e.g., involvement of control groups, adequacy of sample).
- Conduct scientific investigations using appropriate tools and techniques (e.g., selecting an instrument that measures the desired quantity—length, volume, weight, time interval, temperature—with the appropriate level of precision).
- Identify patterns in data and/or relate patterns in data to theoretical models.
- Use empirical evidence to validate or criticize conclusions about explanations and predictions (e.g., check to see that the premises of the argument are explicit, notice when the conclusions do not follow logically from the evidence given).

Scientific inquiry is more complex than simply making, summarizing, and explaining observations; and it is more flexible than the rigid set of steps often referred to as the “scientific method.” The *National Standards* makes clear that inquiry goes beyond “science as a process” to include an understanding of the nature of science (p. 105) and further states the following:

> It is part of scientific inquiry to evaluate the results of scientific investigations, experiments, observations, theoretical models, and the explanations proposed by other scientists. Evaluation includes reviewing the experimental procedures,

[23] Additionally, 12th graders at the Advanced level are expected to be able to identify a scientific question for investigation. See Appendix C for Preliminary Achievement Level Descriptions.

> examining the evidence, identifying faulty reasoning, pointing out statements that go beyond the evidence, and suggesting alternative explanations for the same observations (p. 171).

In the NAEP Science Assessment, when students Use Scientific Inquiry, they are drawing on their understanding about the nature of science, including the following ideas (see *Benchmarks*):

- Arguments are flawed when fact and opinion are intermingled, or the conclusions do not follow logically from the evidence given.
- A single example can never support the inference that something is always true, but sometimes a single example can support the inference that something is not always true.
- If more than one variable changes at the same time in an experiment, the outcome of the experiment may not be clearly attributable to any one of the variables.
- The way in which a sample is drawn affects how well it represents the population of interest. The larger the sample, the smaller the error in inference to the population. But, large samples do not necessarily guarantee representation, especially in the absence of random sampling.

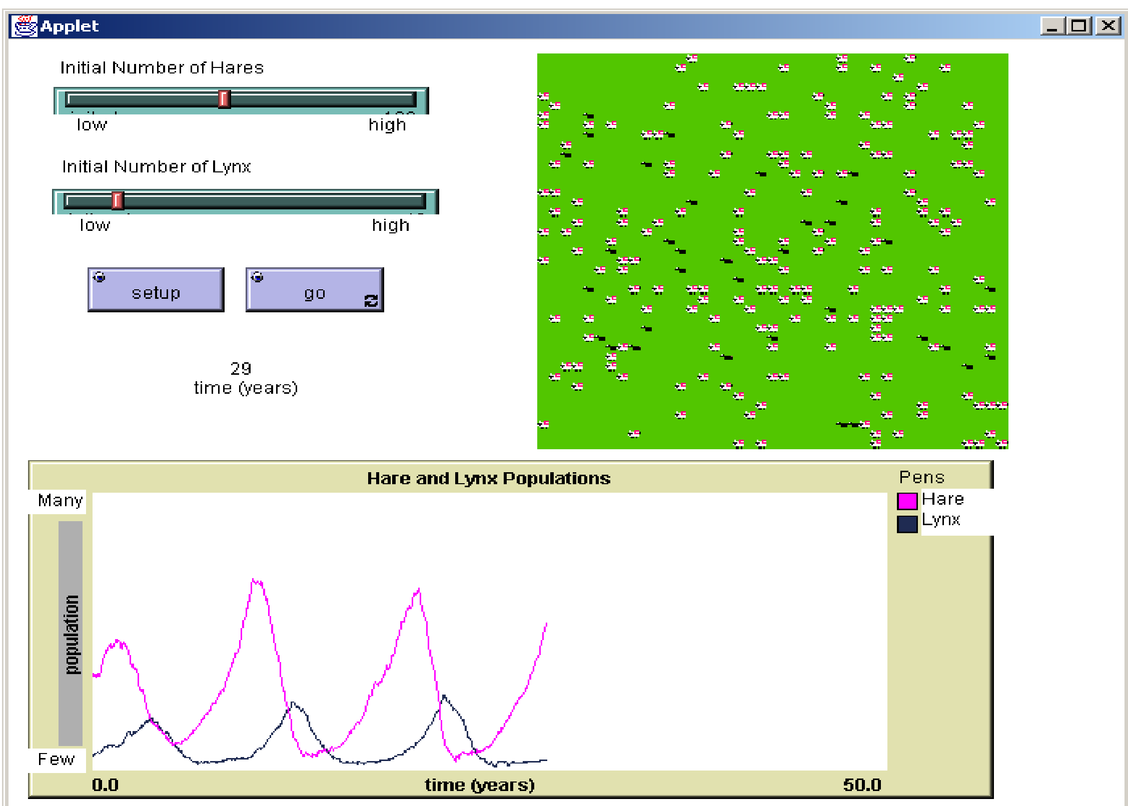
NAEP will assess students' abilities to Use Scientific Inquiry in two ways: students will be required to *do* the practices specified above, and students will *critique examples* of scientific inquiry. It is incorrect to assume that assessment of Using Scientific Inquiry is best or only achieved through hands-on performance tasks and ICTs. In both cases of *doing* and *critiquing*, some assessment tasks will also be presented as paper-and-pencil items. In *doing*, tasks may present data tables and ask students which conclusions are consistent with the data. Other tasks will be presented as hands-on performance and/or interactive computer tasks (e.g., where students collect data and present their results or where students specify experimental conditions on computer simulations and observe the outcomes). As to *critiquing*, students might be asked to identify flaws in a poorly designed investigation or suggest changes in the design in order to produce more reliable data. Tasks may be based on print or electronic media (e.g., items may ask students to suggest alternative interpretations of data described in a newspaper article). For more information on types of items, see Chapter Five.

The following middle school (grade 8) item illustrates the expectation that students conduct scientific investigations. By manipulating the simulation, students gather data and solve the problems given.

Illustrative Item

This interactive computer task is one module in an extended assessment of students' abilities to use a range of technologies to investigate a complex problem, "Should lynx be re-introduced into a national park?" Students accessed, organized, and analyzed data on the number of hares in the park over a 25-year period, researched factors that would impact the population, and created a graph to analyze the trend. (See Appendix D for description of the full task.)

This module allows students to interact with a simulated predator/prey (lynx/hare) population model. Students use the modeling tool to observe population trends that result from different parameter values for the lynx and hare populations. The screen shot below is an example of what students see after they have selected parameters and run the simulation. Note that it is a single screen shot and represents only a small subset of the many screens actually seen by students engaged in this interactive computer task. After students have run the modeling software, they are asked a series of questions (e.g., size of the hare population over time).



L8.6, Using Scientific Inquiry
Source: Quellmalz, Griffin, Hurst, Kreikemeier, Rosenquist, and Zalles (2004)

Using Technological Design[24]

In both the *National Standards* and *Benchmarks*, the term "technological design" refers to the process that underlies the development of all technologies, from paperclips to space stations. As pointed out in the *National Standards,* this meaning "is not to be confused with 'instructional technology,' which provides students and teachers with exciting tools—such as computers—to conduct inquiry and to understand science" (p. 24).

In the *Framework* and *Specifications*, Using Technological Design describes the systematic process of applying science knowledge and skills to solve problems in a real-world context. The reason for including technological design in the science curriculum is clearly stated in the *National Standards*: "Although these are science education standards, the relationship between science and technology is so close that any presentation of science without developing an understanding of technology would portray an inaccurate picture of science" (p. 190). The *National Standards* defines technology and its relationship to science as follows:

> As used in the *Standards*, the central distinguishing characteristic between science and technology is a difference in goal: The goal of science is to understand the natural world, and the goal of technology is to make modifications in the world to meet human needs. Technology as design is included in the *Standards* as parallel to science as inquiry (p. 24).

As it is in scientific inquiry, the professional practice of technological design (also called *engineering design*) is complex and time-intensive. Because NAEP addresses the subject area of science, the use of technological design components in the 2009 NAEP Science Assessment will be limited to those that reveal students' abilities to apply science principles in the context of technological design. Students' abilities to Identify and Use Science Principles should provide the opportunities as well as the limits for assessment tasks related to Using Technological Design. For example, if students are asked to design a town's energy plan, they may be expected to consider the environmental effects of using natural gas versus using coal, but they would not be expected to consider the economic, political, or social ramifications of such a plan.

The *Framework* and *Specifications* sample key components of Using Technological Design from the more complete descriptions found in the *National Standards* and *Benchmarks*. Using Technological Design comprises the following general types of performance expectations, all of which entail students using science knowledge to accomplish the following:

- Propose or critique solutions to problems, given criteria and scientific constraints.
- Identify scientific trade-offs in design decisions and choose among alternative solutions.
- Apply science principles or data to anticipate effects of technological design decisions.

[24] This practice is elaborated in some detail since it is new in NAEP Science Assessments.

The three components of Using Technological Design are elaborated below.

First, the technological design process is rooted in the definition of a *problem* that can be solved through a technological design process. The problem generally describes a human need or want and specifies criteria and constraints for an acceptable solution (International Technology Education Association, 2000). Only constraints that reflect the science content statements in the *Framework* and *Specifications* will be considered in developing relevant NAEP assessment items. The engineer who designs a bridge, for example, must take into account the effects of wind and water currents by using relevant physics principles to simulate these effects on possible structures before the bridge is built.

Second, even if limited to the application of science principles, choosing between alternative solutions almost always involves *trade-offs*. As stated in *Benchmarks*:

> There is no perfect design. Designs that are best in one respect . . . may be inferior in other ways . . . . Usually some features must be sacrificed to get others. How such trade-offs are received depends upon which features are emphasized and which are down-played (p. 49).

The application of science principles may be used to compare alternative technological solutions to see which will better solve the problem and accomplish the goals of the project.

Finally, while the chosen solution may be intended to solve a human problem or meet a human need, the *effects* are not always as planned. When the automobile was invented, no one could have predicted the environmental and human health impacts of vehicle emissions. However, it is the job of scientists and engineers working together to apply their knowledge of the natural world to make such predictions. According to the *National Standards*, students in grades K-4 should know about the effects of design solutions:

> People continue inventing new ways of doing things, solving problems, and getting work done. New ideas and inventions often affect other people; sometimes the effects are good and sometimes they are bad. It is helpful to try to determine in advance how ideas and inventions will affect other people (p. 140).

In terms of cognitive demands, both “declarative (knowing that) knowledge” and “schematic (knowing why) knowledge” come into play for the three components of Using Technological Design, as does “strategic knowledge—knowing when and where to apply knowledge.”

The role of technological design in U.S. science classrooms currently varies widely, and it is not possible to predict the extent to which it will be integrated into the school curriculum in the future. The role of technological design in NAEP Science will need to be revisited regularly, in response to its evolving role in school science.

Since Using Technological Design in the NAEP Assessment needs to have direct relevance to science, it is assumed that students have some understanding about the relationship between science and technology. The science-technology relationship is further discussed in Chapter Five as providing context for assessment items.

Several types of items are appropriate to measuring this practice. The open nature of constructed-response items requires special effort by item writers to focus students on the scientific aspects of technological design. The items should not allow for vague responses that include non-scientific factors, such as social, economic, political, or aesthetic considerations. In addition to constructed-response items, selected-response items can also probe students' ability to use technological design (i.e., to apply science knowledge and skills to solve problems in real-world contexts). Because this is a relatively new area to be included in science assessments, the *Specifications* includes only a few items that exemplify the assessment of Using Technological Design. The work of NAEP item writers in this area will set a standard for these types of assessment items.

The following items illustrate the expectation that students apply science principles to technological design problems.

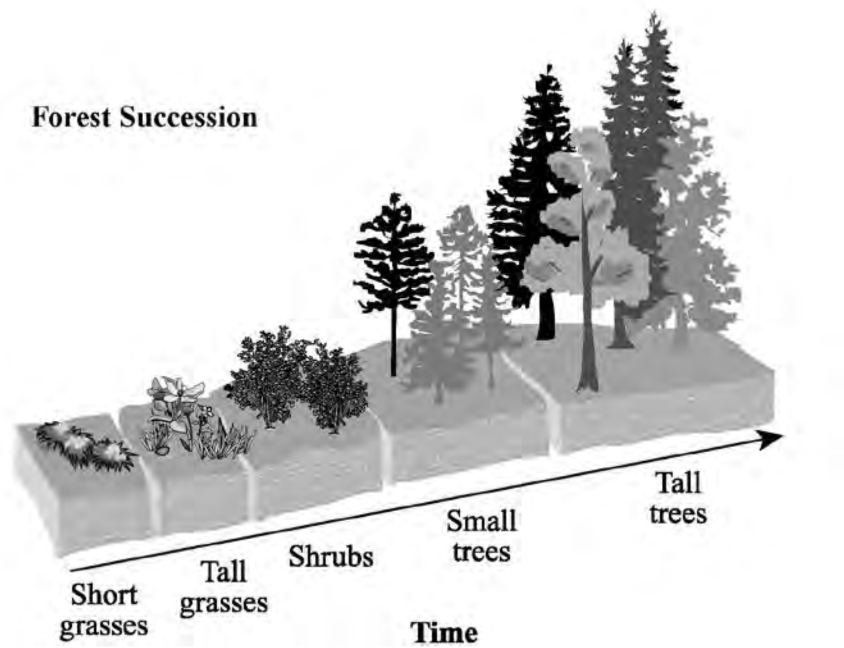
Item Suggestion

A grade 12 item could be devised that contains text as well as a series of maps showing air currents that carry industrial pollution materials from remote places over vast distances. Students could be asked to read the text, interpret the maps, and offer their own science-based solutions with respect to the problem of acid rain, either generally or for a specified area. Students' answers should not reflect personal opinions but should rely exclusively on the scientific data presented.

E8.15, Using Technological Design
Source: Adapted from the 1996-2005 *Science Assessment and Exercise Specifications for the National Assessment of Educational Progress*

Illustrative Item

Occasionally, a fire will destroy a forest, burning down trees and pushing wildlife out of their forest homes. However, the forest will grow back. Eventually, through the process of forest succession as shown below, short grasses and flowers begin to grow and animals make new homes.



Over time, shrubs and trees begin to grow. The forest returns to a lush habitat for the wildlife listed in the chart below.

Forest Wildlife

Ground-dwelling	Worms, beetles
Reptiles and amphibians	American toads, wood frogs, snakes, Eastern box turtles
Small animals	Squirrels, chipmunks
Medium to large animals	Opossums, raccoons, white-tailed deer, black bears
Airborne	Butterflies, moths, bees, wild turkeys, red-tailed hawks, bald eagles

A power company owns part of a forest that was destroyed by a fire. The forest could take decades to rebuild on its own. The company's department of environmental studies suggests planting new trees to help the forest rebuild.

Using the information in the scenario:

- Explain how planting trees could **benefit** the natural ecosystem.
- Explain how planting trees could **harm** the natural ecosystem.

(See Appendix D for item scoring guides.)

E8.15, Using Technological Design
Source: Washington Assessment of Student Learning, 2004, Grade 8

Summary of Practices

The general performance expectations for each of the four practices are summarized in Table 15. Dashed lines indicate that the boundaries between these categories are not distinct, and some overlap is to be expected.

The *Specifications* does not provide suggested developmental progressions for the four science practices, that is, the implication of specific practice components is not detailed for each grade level. For example, "explain observations of phenomena" and "conduct scientific investigations" indicate different student performances at grades 4, 8, and 12. The nature of science practices at a given grade level depends on the science content with which it is crossed; when the same science practice component is combined with different science content, different student performance expectations will result. The section on "Setting Boundaries" and the tables of content boundaries in Chapter Two should inform the student performance implications of specific practice components at grades 4, 8, and 12.

Table 15. General Performance Expectations for Science Practices

<table>
<tr><td rowspan="4">← Communicate accurately and effectively →</td><td>Identifying Science Principles</td><td>Describe, measure, or classify observations</td><td>State or recognize correct science principles</td><td colspan="2">Demonstrate relationships among closely related science principles</td><td>Demonstrate relationships among different representations of principles</td></tr>
<tr><td>Using Science Principles</td><td>Explain observations of phenomena</td><td>Predict observations of phenomena</td><td colspan="2">Suggest examples of observations that illustrate a science principle</td><td>Propose, analyze, and/or evaluate alternative explanations or predictions</td></tr>
<tr><td>Using Scientific Inquiry</td><td>Design or critique aspects of scientific investigations</td><td>Conduct scientific investigations using appropriate tools and techniques</td><td colspan="2">Identify patterns in data and/or relate patterns in data to theoretical models</td><td>Use empirical evidence to validate or criticize conclusions about explanations and predictions</td></tr>
<tr><td>Using Technological Design</td><td colspan="2">Propose or critique solutions to problems given criteria and scientific constraints</td><td colspan="2">Identify scientific trade-offs in design decisions and choose among alternative solutions</td><td>Apply science principles or data to anticipate effects of technological design decisions</td></tr>
</table>

Clarification: Sample Performance Expectations for a Life Science Content Statement

The examples below are all related to the following Grade 8 Life Science content statement:

L8.4: Plants are producers—they use the energy from light to make sugar molecules from the atoms of carbon dioxide and water. Plants use these sugars along with minerals from the soil to form fats, proteins, and carbohydrates. These products can be used immediately, incorporated into the plant's cells as the plant grows, or stored for later use.

All examples are also related to a specific situation:

Two different varieties of grass—one better adapted to full sunlight and one better adapted to shade—are each grown in sunlight and in shade.

The results of a controlled experiment along these lines might resemble the following:

Condition	**Grass Type A**	**Grass Type B**
Sunlight	"Better growth"*	"Less good growth"*
Shade	"Less good growth"*	"Better growth"*

* Several variables could be used to indicate growth: mass or dry mass of plants, thickness of stems, number of new sprouts, etc.

<u>Identifying Science Principles</u>

1. State from where a plant's food originates.
2. Classify the grass plants as producers or consumers.

The first performance calls for students to repeat information found in the content statement with little or no modification. The second performance asks students to use the definition of producers given in the content statement to classify or identify the plants.

<u>Using Science Principles</u>

1. Predict whether sugar will move up or down the stems of the grass plants and explain your prediction.
2. Explain where the mass of the growing grass originates.

These performances require students to use principles in the content statement to predict or explain specific observations (growing grass in this case). The content statement itself does not provide the answers to the questions.

Using Scientific Inquiry

1. Given a data table showing the mass of grass plants of each type grown in the sunlight and shade, draw conclusions about which variety of grass is better adapted to each condition.
2. List other variables that should be controlled in order to feel confident about your conclusions.

The first performance is related to the content statement in that the importance of light for plant growth is useful background information for students. However, the performance requires interpretation of new information (the data table) that has to do with differences among types of plants, while the content statement contains generalizations about all plants. Thus, the performance requires students to use the data to develop new knowledge that they had not had before. The second performance is in part an assessment of the students' understanding of experimental design. However, good answers would also require knowledge of this and related content statements to identify variables that are relevant to plant growth.

Using Technological Design

1. Given experimental results on the growth of different varieties of grass plants under sunlight and shade conditions, develop a plan for using different types of grass seed in different parts of a partially shaded park.

This performance requires students to use knowledge of the content statement and the experimental results in order to accomplish a practical goal, in this case, a park with grass growing well in areas that receive varying amounts of sunlight.

Distribution of Items by Science Practice

The item distribution for the four science practices, as measured by percentage of student response time at each grade level, should have the following approximate allocations:

Table 16. Distribution of Items by Science Practice and Grade

	Grade 4 (% student response time)	Grade 8 (% student response time)	Grade 12 (% student response time)
Identifying Science Principles	30	25	20
Using Science Principles	30	35	40
Using Scientific Inquiry	30	30	30
Using Technological Design	10	10	10

Moving from grades 4 to 8 to 12, the emphasis on Using Science Principles increases, while the emphasis on Identifying Science Principles decreases. The expectation is that, as students move up through the grades, their critical response skills and methodological and analytical capabilities will increase.

Cognitive Demands

The four science practices—Identifying Science Principles; Using Science Principles; Using Scientific Inquiry; and Using Technological Design—articulate what students should know and be able to do with the science principles presented in Chapter Two. Certain ways of knowing and reasoning—cognitive demands—underpin these four science practices. Here, the four cognitive demands—"knowing that," "knowing how," "knowing why," and "knowing when and where to apply knowledge"—are elaborated. The goal is to further elucidate the descriptions of the science practices, to facilitate item specifications and item writing, and to provide a framework for interpreting students' responses. That is, the set of four cognitive demands can be used as a lens to facilitate item development and to analyze student responses (Li, Ruiz-Primo, & Shavelson, in press; Li & Shavelson, 2001; Shavelson, 2006; Shavelson, Ruiz-Primo, & Wiley, 2005), thereby checking expectations regarding what science content and practice(s) are being tapped by a given assessment item.[25]

Knowing That

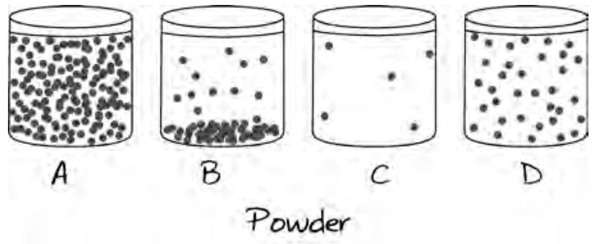
This cognitive demand, often called "declarative knowledge," sets up the expectation that students should know and reason with basic science principles (e.g., density is mass per unit volume) and that they should be able to recall, define, represent (schematically, pictorially, verbally, mathematically—e.g., $density = mass / volume$), use, and relate these basic principles as appropriate. Following a learning trajectory, these science principles become increasingly more sophisticated over time.

For example, consider a student investigating saturation. The student knows that saturation is "the maximum amount something can hold" and needs to *recall* this construct in order to investigate which of four powders saturates 20 mL of water the most. Additionally, the student needs to mentally represent what a saturated solution looks like as more substance is added.

[25] Note that more than one cognitive demand can be associated with the more complex science practices. These associations may shift according to the knowledge that students at different grade levels bring to an assessment task.

Item Suggestion

To find out which one of four powders is most soluble in water, Maria put 10 spoonfuls of each powder into beakers filled with 20 mL of water. Based on the evidence from the beakers, which powder is most soluble in water?



A. Powder A
B. Powder B
C. Powder C
D. Powder D

Key: C

P8.4, Identifying Science Principles

To go further, following the work of Li et al. (in press) and Li and Shavelson (2001), the next item taps "declarative knowledge." It asks students to identify the digestive substance in the mouth and its function. First, the response is expected to be in the form of terms, vocabulary (e.g., saliva), and factual statements. Second, the cognition involved is likely to be recall of a term from memory. (The item is not only similar to school-type problems but similar to the way texts are written and students are taught.) The cognitive process involved in answering the item is to directly retrieve information or do a minimum of scientific reasoning to organize the relevant information. The response format (i.e., constructed-response) is more cognitively demanding than multiple-choice. That is, instead of recognizing correct answers from options given, students must recall and organize their responses.

Illustrative Item

What digestive substance is found in the mouth? What does it do?

(See Appendix D for item scoring guides.)

L8.1, Identifying Science Principles
Source: TIMSS 1995, Grade 8

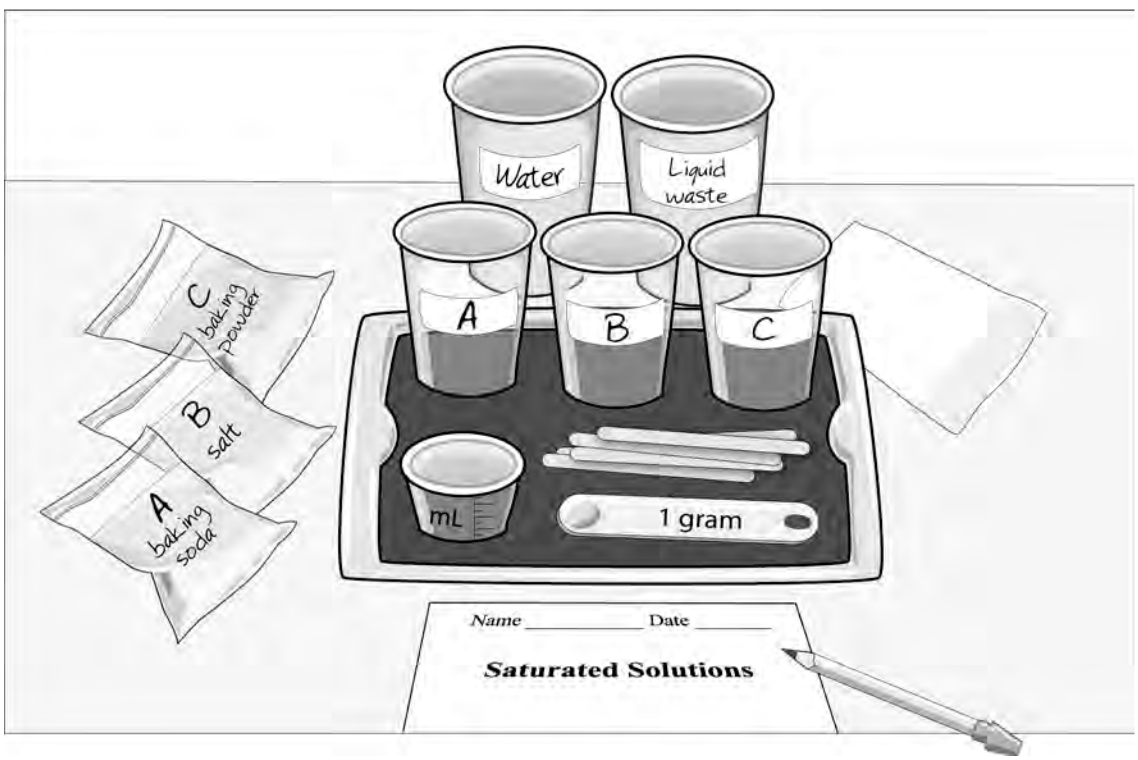
Knowing How

This cognitive demand, often called "procedural knowledge," sets up the expectation that students can apply the science principles in doing science. Students know how to perform and reason with simple (routine) and complex procedures. Simple or routine procedures involve systematically observing and recording which objects sink and float in water, following well-established procedures for using a balance scale, measuring an object's mass, calculating an object's density, finding patterns in data, or graphing the depth of sinking of objects varying in mass (holding volume constant). More complex procedures are involved in designing and interpreting the results of an investigation (e.g., manipulating one variable and holding others constant), or creating a tool for some practical purpose (e.g., moving heavy objects).

Consider again the saturation investigation. The student knows how to hold certain variables constant, how to consistently use appropriate materials and methods to measure out 20mL of water for each of three powders, and so on.

Illustrative Item

You have three bags of white powder in front of you: Bag A contains baking soda, Bag B contains salt, and Bag C contains baking powder. How many grams of each powder are needed to saturate 20 mL of water? Of the three powders, which one do you have to use the most to saturate the liquid? Which one do you have to use the least? Record your procedures, observations, and findings in the notebook provided.



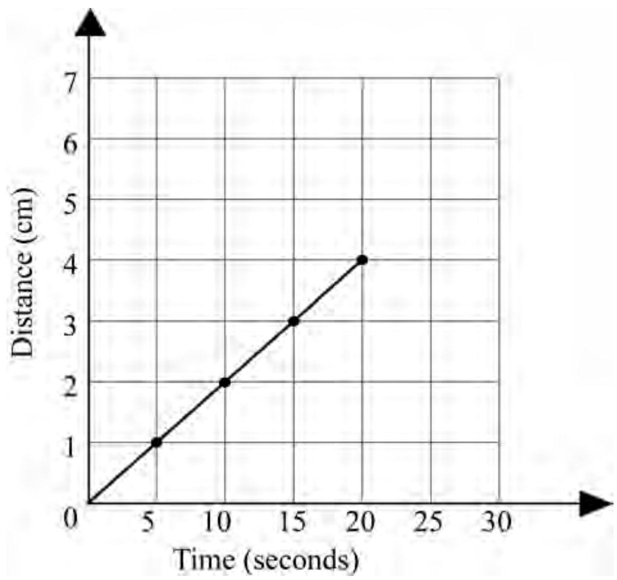
(See Appendix D for further description of the task and scoring system.)

P8.4, Using Scientific Inquiry
Source: Ruiz-Primo, Shavelson, Hamilton, and Klein (2002)

Again, to go further into the notion of "procedural knowledge," the following item provides students with a graph showing the distance that an ant moves over 20 seconds and asks the distance that the ant will travel at 30 seconds. The item requires students to respond with an algorithm to calculate speed and/or to interpret the graph to obtain the number for distance. Second, the cognitive process students probably engage in is either applying the calculation algorithm of speed by dividing distance with time or extending the line in the graph to 30 seconds and simply reading the distance on the vertical axis. (However, the item structure may limit students in applying the procedure, and some students may arrive at the correct answer by working backwards from options.)

Illustrative Item

The graph shows the progress made by an ant moving along a straight line.



If the ant keeps moving at the same speed, how far will it have traveled at the end of 30 seconds?

A. 5cm
B. 6cm
C. 20cm
D. 30cm

Key: B

P8.14, Using Science Principles
Source: TIMSS 1995, Grade 8

Knowing Why

This cognitive demand, sometimes called "schematic knowledge," sets up the expectation that students can explain and predict natural phenomena, as well as account for how and why scientific claims are evaluated, argued and justified, or warranted (explaining and reasoning with principles and models). That is, this cognitive demand deals with students' understanding of how the natural world works, such as why some things sink and others float in water, why light is essential to the propagation of most plants, or why the moon changes phases. This cognitive demand applies equally to knowing how and why science proceeds as it does, by observation, hypothesis/counter-hypothesis, prediction, empirical testing, modeling, argumentation, justification, and replication. Accordingly, students can, for example, distinguish observation from model, account for uncertainty, and be alert to the need to control extraneous variables. In the saturation investigation described earlier, a ("mental") model of why extraneous variables need to be controlled would alert the student to place the same amount of each of three powders in the same amount of water (20mL). Stated another way, this cognitive demand requires students to argue from a ("mental") model of the way in which nature and science works.

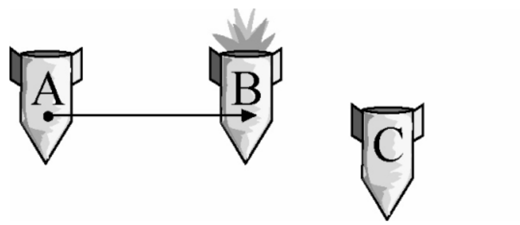
Unless known by rote after a great deal of practice, "knowing why" would permit a student to accurately predict the effects of changes in system variables (e.g., a light bulb is added to a series circuit). Likewise, unless memorized through prior practice, "knowing why" would permit students to explain why Washington, DC experiences a change of seasons using a model of the relationship between Earth's orbit and the sun. They would be expected to justify why, in studying the paths of planets in the night sky, modeling of a time-series of observations and not a randomized comparative trial would be most appropriate to building an understanding of planetary movement. As another example, students would be able to predict what would happen if the volume of an object increased while the mass remained constant (e.g., designing a boat). "Knowing why" would be brought to bear in explaining the complex phenomenon of buoyancy.

Note that the student's model may not be scientifically acceptable. Naïve models of natural and human systems abound; a major goal of science education is to move students along in their learning toward the scientifically accepted model.

In summary, this cognitive demand is tapped when students are asked to predict and/or explain natural and human-made phenomena. This knowledge and reasoning with "why" or "how" things work—i.e., the model underlying why—might be expressed verbally in a set of propositions, schematically in a model, mathematically in a model, and the like. This cognitive demand is the focus of a great deal of research on students' "mental models" and change in conceptual understanding.

The following item, originally used with freshman engineering students, assesses "schematic knowledge" by probing students' conceptions in introductory mechanics. A typical incorrect answer shows the rocket returning to a horizontal path after the engine is shut off at C. This reveals a common naïve conception—that "continuing motion implies the presence of a continuing force in the same direction as the motion" (Clement, 1982, p. 67). In other words, students may assume that continuous motion from A to B means that a continuous force is moving in the same direction, and this force will once again "take over" when the rocket engine is shut off. Furthermore, a typical incorrect response shows instantaneous changes in the rocket's motion (straight lines), which is in contrast to the continuous motion (curved lines) represented in a physicist's more sophisticated answer. These are naïve conceptions that some students continue to hold even in the face of their high school physics courses.

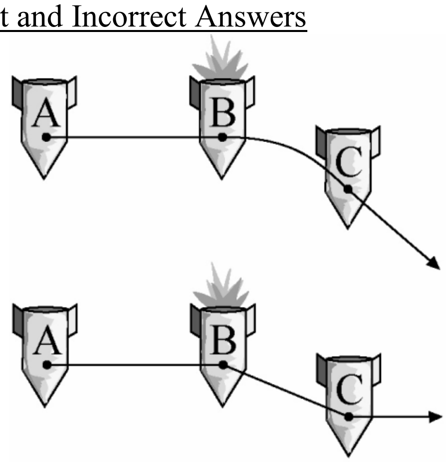
Illustrative Item



A. A rocket is moving along sideways in deep space, with its engine off, from point A to point B. It is not near any planets or other outside forces. Its engine is fired at point B and left on for 2 sec while the rocket travels from point B to point C. Draw in the shape of the path from B to C. (Show your best guess for this problem even if you are unsure of the answer.)

B. Show the path from C after the engine is turned off on the same drawing.

<u>Correct and Incorrect Answers</u>



Physicist's Answer

Typical Incorrect Answer

P12.19, Using Science Principles
Source: Clement (1982)

The following item also taps students' "schematic knowledge." This item asks students to recognize the explanation for why there is daylight and darkness on Earth. Note, the item is intended to assess students' knowledge about the cause of daylight and darkness because it invites students to select conceptual explanations by using the phrase "explain why" in the stem. The knowledge that students use to answer this item might be models or theories. Some students might use visualization or diagrams to represent and infer the process and consequence of the movement. However, depending on school curriculum or prior experience, some students may have memorized the correct answer and therefore may be using declarative knowledge (see "Comments on Cognitive Demands" later in this chapter).

The dominant cognitive process expected from students is reasoning with models or theories, perhaps with the aid of visual representations and/or spatial manipulation of objects mentally. Students have to determine the correct answer from four models by reasoning about the movement of the Earth and sun (e.g., rotating on axis and revolving). Indeed, a critical element in answering the question is to bring additional underlying principles and knowledge related to Earth and sun movement instead of merely reading the options. For example, while reading the options, students might also think about and figure out that the option C partially explains the seasons and D partially explains the year.

Illustrative Item

Which statement explains why daylight and darkness occur on Earth?

A. The Earth rotates on its axis.
B. The Sun rotates on its axis.
C. The Earth's axis is tilted.
D. The Earth revolves around the Sun.

Key: A

E8.2, Using Science Principles
Source: TIMSS 1995, Grade 8

Knowing When and Where to Apply Knowledge

This cognitive demand, also called "strategic knowledge," is commonly talked about as "transfer" of current knowledge to new situations (tasks or problems). "Strategic knowledge" involves both knowing when and where to use science knowledge in a new situation *and* reasoning through ("assembling cognitive operations for") the novel task to reach a goal. "Strategic knowledge" sets up the expectation that students can take their current knowledge and apply it to a somewhat novel situation. In doing so, they need to understand the dimensions underlying the new situation, see how these dimensions are related to what they know, and draw upon that knowledge. In the process, they need to monitor whether their knowledge is leading toward the task goal or sub-goal and evaluate their plan for addressing the task along the way.

Expert scientists exemplify “strategic knowledge” to a high degree. Darwin, for instance, did not just observe the finches, turtles, and other animals of the Galapagos, but he pondered on why minor differences existed between inhabitants of different islands. He searched for an explanation. The explanation he created became available to others as a piece of “knowing why (schematic)” knowledge. *Knowing how to arrive at the explanation* was “strategic knowledge.” Of course, one does not expect school students to possess “strategic knowledge” of such an order. However, they can demonstrate “strategic knowledge” to some extent, for example, by recognizing that an answer is absurd and finding out why or setting sub-goals for a laboratory investigation into a novel problem.

“Strategic knowledge” arises throughout the process of taking an assessment that the student has not seen before. Certain items, such as recall items (“declarative knowledge”), place little pressure on “strategic knowledge,” especially if an association is immediately available. (If “test-wiseness” is needed in the absence of clear recall, an increased demand is placed on “strategic knowledge.”) Distinctions between “strategic knowledge” and other knowledge types must be based on knowledge of the curriculum to which students have been exposed as well as of their learning development. Knowing how to control variables in studying pendulum periodicity, for example, might be a routine application of “knowing how” as this topic is covered extensively in many elementary curricula; applying control of variables to a new situation (e.g., studying the impact of scarification of hard shell seeds on germination) would involve considerable demand on “strategic knowledge” (and “knowing why” or “schematic knowledge”). As another example, items that are designed to have students demonstrate their understanding of, say, photosynthesis in the context of historical information about its discovery draw substantially upon “strategic knowledge.” Here is an example of an item illustrating “strategic knowledge”:

Item Suggestion

Suppose that Earth is positioned at a 90 degree angle in its orbit around the sun instead of being positioned at about a 23 degree angle (23.4) offset from perpendicular. Explain how this would affect the change of seasons in North America.

E8.12, Using Science Principles

Comments on Cognitive Demands

Three comments about these cognitive demands are in order. The first comment deals with their interrelated nature. Simply stated, they are not independent demands but, rather, related. That is, when explaining “why,” a student will need to call on “knowing that” and, at times, in justifying “why,” may have to call on “knowing how.” And, depending on the novelty of the task, “strategic knowledge (knowing when and where to apply knowledge)” may be called into play. Nevertheless, these related cognitive demands can be distinguished, and it is useful and practical to do so.

The second comment is that there is not a one-to-one correspondence between the cognitive demand and the type of assessment item used. Research has shown that multiple-choice items assess “knowing that” effectively and efficiently. Yet, they also can assess “knowing why,” especially when “clusters” of multiple-choice items are used to probe students’ mental models (e.g., Sadler, 1998). In addition, performance tasks have been shown to tap “knowing how” reasonably well, but computer simulations and some paper-and-pencil items can also do this. In the end, in order to verify that a particular item is tapping the intended cognitive demand, cognitive labs that probe students’ thinking will be needed (see Chapter Five).

Third, the cognitive demand exerted by any particular item depends on what a student brings to the assessment situation. Consider an open-ended item that asks students to explain why some things sink and other things float in water. The item calls for an explanation based on a model of buoyancy. Most students learning about buoyancy will evoke a “mental model” and use this model to provide an explanation—in this case, the item is working as intended. However, students well along in their learning of buoyancy might simply recall from memory the answer to the question, as they have reasoned through this kind of problem and practiced the answer in various forms over and over. For these students, the item has tapped “declarative knowledge (knowing that).” Again, cognitive labs that probe students’ thinking will be needed to determine what cognitive demand is being probed by assessment items for the particular population of students taking the assessment, recognizing that not all students are at the same level of competency.

As pointed out earlier in this chapter, one or more cognitive demands can be related to each of the science practices. The practice of Identifying Science Principles largely taps students’ “declarative knowledge” or “knowing that.” Clearly, the other three science practices also draw on students’ “declarative knowledge”; however, they ought to be used primarily as opportunities to tap the other cognitive demands. The practice of Using Science Principles draws heavily on students’ “schematic knowledge” or “knowing why,” but it also involves the other types of knowledge. The practices of Using Scientific Inquiry and Using Technological Design both draw heavily on “procedural knowledge” or “knowing how,” but they also involve the other types of knowledge, especially “strategic knowledge” or “knowing when and where to apply knowledge.” These relationships between the science practices and cognitive demands imply that the distribution of items according to the four science practices corresponds reasonably well to that of the cognitive demands.

Generating and Interpreting Items

The next chapter will show how science content statements can be combined or “crossed” with practices to generate performance expectations, which then guide the development of assessment items. By comparing student responses to the particular science content and practice being assessed, inferences about what students know (about particular science principles) and can do (with respect to particular science practices) are made.

CHAPTER FOUR: GENERATING AND INTERPRETING ITEMS

Introduction

The NAEP Science Assessment will focus on how students bring science content (as described in Chapter Two) to bear as they engage in the practices described in Chapter Three. That is, science practices are not content-free skills; they require knowledge of the Physical, Life, and Earth and Space Sciences as well as knowledge about scientific inquiry and the nature of science (e.g., drawing conclusions from investigations).

Performance Expectations

Performance expectations are derived from the intersection of content statements and science practices. If the content statements from the Physical, Life, and Earth and Space Sciences are the columns of a table and the practices (Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, Using Technological Design) are the rows, the cells of the table are inhabited by performance expectations. Examples are provided in Table 17, which is based on Figure 1 in Chapter One. Note that performance expectation cells may overlap, since the content and practice categories themselves are not distinct (as indicated by dashed lines).

Table 17. Generating Examples of Grade 8 Performance Expectations

		Science Content		
		Physical Science content statements	**Life Science** content statements	**Earth and Space Science** content statements
Science Practices	**Identifying Science Principles**	Identify the units that might be used to measure the speed of an ant and the speed of an airplane. (See **P8.14**.)[26]	Identify the raw materials that plants use to make sugars. (See **L8.4**.)	Identify wind as the movement of air from higher to lower pressure regions. (See **E8.11**.)
	Using Science Principles	An object (e.g., a toy car) moves with a constant speed along a straight line. Predict (with justification) what might happen to this object's speed as it rolls downhill. (See **P8.16**.)	Explain why sugars are found to move primarily down the stem of a growing plant (e.g., potato, carrot). (See **L8.4**.)	Explain why mountain soils are generally thinner than floodplain soils. (See **E8.6**.)
	Using Scientific Inquiry	Design an experiment to determine how the speed of a battery-operated toy car changes as a result of added mass. (See **P8.16**.)	Criticize conclusions about likely consequences of consuming various diets based on flawed premises or flaws in reasoning. (See **L8.5**.)	Given data (indexed by month) on annual trends of incoming solar radiation for five cities, determine whether the location is in the Northern or Southern Hemisphere. (See **E8.12**.)
	Using Technological Design	Evaluate the following car designs to determine which one is most likely to maintain a constant speed as it goes down a hill. (See **P8.16**.)	Identify possible ecological side effects of agricultural fertilizer runoff into a lake. (See **L8.7**.)	Describe the consequences (e.g., erosion) of undercutting a steep slope for a road cut. (See **E8.4**.)

[26] In order to identify the science content statement on which each performance expectation is based, the content statement's unique code (from Tables 6, 9, and 12 in Chapter Two) is provided.

The content statements from Chapter Two on which these performance expectations are based are written in general terms. The process of creating performance expectations requires further clarification of the content statements themselves. As described in Chapter Two, this involves "detailing" the meanings of the content statements and setting boundaries on the content to be assessed at a given grade level. Moreover, if the crossing of content statements with practices were done for every science content statement and practice, the number of performance expectations generated could be unmanageably large. Selected example performance expectations for some content statements are provided later in this chapter and in Appendices F, G, and H. These examples are illustrative, not exhaustive. It is expected that assessment developers will continue this process for all content and practices sampled for a particular NAEP Science Assessment.

Performance expectations are written with particular verbs indicating the desired performance expected of the student. The action verbs associated with each practice are not firmly fixed, and the use of any action verb must be contextualized. For example, when the science practice component, "*conduct* scientific investigations," is crossed with a states-of-matter content statement, this can generate a performance expectation that employs a different action verb, "*heats* as a way to evaporate liquids."

Examples of Generating and Interpreting Items

Neither the content statements from Chapter Two nor the practices discussed in Chapter Three will be assessed in isolation. All assessment items will be derived from a combination of the two (i.e., from performance expectations). Assessment developers have latitude in determining which item type is most appropriate for any given performance expectation (see Chapter Five for an explication of types of items). Observed student responses to these items can then be compared with expected student responses in order to make inferences about what students know and can do.

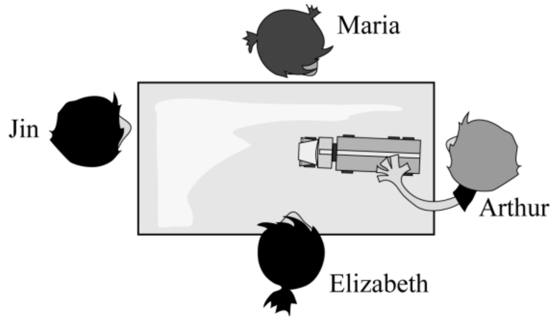
Examples of generating and interpreting items have been developed for grades 4, 8, and 12 within each content area. Examples differ in level of completeness. Following are three examples of this process of generating and interpreting items: one 4th grade Physical Science example (Table 18), one 12th grade Life Science example (Table 19), and one 8th grade Earth and Space Science example (Table 20). Complete examples, such as the one contained in Table 20, start with a content statement or a set of related content statements, followed by commentary on the content statement(s) and examples of performance expectations. These performance expectations are elucidated by *item suggestions* (descriptions of items to be developed) or *illustrative items* (released items from various large-scale assessments). For additional examples, see Appendices F, G, and H.

Table 18. Grade 4 Physical Science Example of Generating and Interpreting Items

Grade 4: Motion—Motion at the Macroscopic Level (The following example also appears in Appendix F.)
Content statement **P4.12:** An object's position can be described by locating the object relative to other objects or a background. The description of an object's motion from one observer's view may be different from that reported from a different observer's view.
Commentary The above content statement has the following implications: it is confusing when individuals observing the same motion from different positions describe it differently; using an external frame of reference results in observers describing the motion in the same way; and there is scientific value in consistent descriptions of an object's motion. A person observing the motion of a car from one side of the road might say that the car is moving from left to right, while a person on the other side of the road might say that the car is moving from right to left. Each person is using his/her body as the reference point. If the same observers each use a compass, they will observe that the car is moving from north to south. Their observations will be the same because they are using the same frame of reference, the cardinal points of the compass. An external reference point needs to be made clear when describing motion.
Examples of Performance Expectations <u>Identifying Science Principles.</u> Students can: • Recognize that in order to consistently describe an object's motion, an external reference point (e.g., North, South, East, West) is needed. <u>Using Science Principles.</u> Students can: • Interpret a situation in which more than one person observes an object in motion from different positions and describe or select descriptions of observations made by individuals from different positions. *Examples of performance expectations for Using Scientific Inquiry and Using Technological Design are not provided.*

Items to Assess Using Science Principles

Item Suggestion 1



Arthur, Maria, Jin, and Elizabeth are sitting around a table. Arthur rolls a toy truck toward Jin. Arthur observes that the truck moves toward Jin. Maria observes that the truck moves from left to right.

a. What does Jin observe about the motion of the truck?
b. What does Elizabeth observe about the motion of the truck?
c. How do you explain differences in what each person observed about the truck's motion?

Item Suggestion 2

If Maria observes that the truck moves from left to right, what should Elizabeth observe about the motion of the truck?

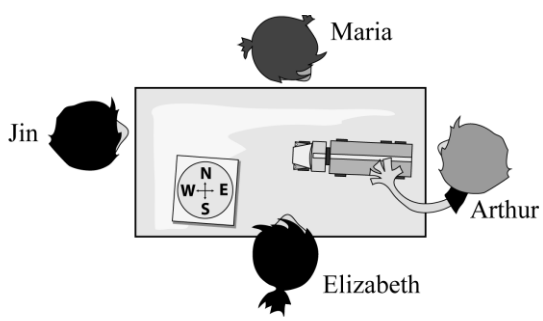
Elizabeth should observe that the truck moves

A. left to right.
B. right to left.
C. toward Arthur.
D. away from Jin.

Key: B

Item Suggestion 3

After the students have made their observations of the truck's motion, Ms. Fu, the science teacher, places a compass on the table and asks the students to use the compass to make their observations.



Items to Assess Using Science Principles (cont.)

The students use the compass to describe the truck's motion after Arthur rolls it toward Jin. Now, what will each student observe about the motion of the truck?

Student	Observation
Arthur	
Maria	
Jin	
Elizabeth	

Item Suggestion 4

The students described the motion of the truck with and without the compass. Which is the better way of describing the motion of the truck? Why do you think this way is better?

Interpretation: Students who respond correctly to the above assessment items have demonstrated knowledge about "left" versus "right." They are able both to describe the motion of an object using his/her body as the reference point and to mentally place herself/himself in the position of another person and use that perspective as a reference point. Students can distinguish the contexts in which it is or is not appropriate to use one's body as a reference point for describing motion; that is, using a compass is more appropriate for describing motion in a scientific context. Students know how to interpret a compass to determine direction. Advanced students also know why, when using the compass, the description of an object's motion will be independent of the observer's position and why the cardinal points as references for motion are useful only for motion in a single plane.

Items to assess Identifying Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Table 19: Grade 12 Life Science Example of Generating and Interpreting Items

Grade 12: Structures and Functions of Living Systems—Organization and Development and Changes in Living Systems—Heredity and Reproduction

(The following example also appears in Appendix G.)

Content Statements

The content statements below describe two main functions of DNA: (1) providing instructions for assembling protein molecules, which then specify the characteristics of organisms and (2) carrying the hereditary information from generation to generation. Here also are the foundations for the genetic basis of evolution.

L12.2: Cellular processes are carried out by many different types of molecules, mostly proteins. Protein molecules are long, usually folded chains made from combinations of amino-acid molecules. Protein molecules assemble fats and carbohydrates and carry out other cellular functions. The function of each protein molecule depends on its specific sequence of amino acids and the shape of the molecule.

L12.8: Hereditary information is contained in genes, located in the chromosomes of each cell. A human cell contains many thousands of different genes. One or many genes can determine an inherited trait of an individual, and a single gene can influence more than one trait.

L12.9: The genetic information encoded in DNA molecules provides instructions for assembling protein molecules. Genes are segments of DNA molecules. Inserting, deleting, or substituting DNA segments can alter genes. An altered gene may be passed on to every cell that develops from it. The resulting features may help, harm, or have little or no effect on the offspring's success in its environment.

Commentary

Relationships among proteins, DNA, genes, and traits can be briefly summarized:

- Proteins carry out the work of cells.
- Proteins are long, usually folded, chains of amino acids.
- Protein function depends on its amino acid sequence.
- DNA provides instructions for assembling proteins.
- Genes are segments of DNA, which carry hereditary information and are located in the chromosomes of each cell.
- One or more genes can determine an inherited trait.
- Genes can be altered (e.g., inserting, deleting, and substituting DNA), and altered genes can be passed on to every cell that develops from them.
- Resulting features from gene alteration may help, harm, or have little effect on offspring's success in its environment.

For L12.2: Why does this content statement focus on proteins? From the perspective of biochemistry, proteins are the workhorses of cells. They are responsible for cell motility, movement of chromosomes during cell division, membrane transport, biosynthesis, and more. For example, nucleic acids assembled by proteins are biologically inert without proteins, which is why viruses aren't considered to be living things; phospholipids (themselves made by proteins) are passive barriers; and proteins transport ions and maintain the electrochemical gradients needed for such functions as nerve conduction. High school students are expected to understand specific examples of protein function as well as the generalization that proteins carry out the work of cells. (However, see exclusions in grade 12 content boundaries for "Organization and Development" and "Heredity and Reproduction.")

For L12.8: Chromosomal abnormalities can be used as contexts for assessment items (e.g., explaining Down's Syndrome).

The following are not recommended for assessment at grade 12:

- Names and structures of nucleotides
- Nature of bonding between DNA strands
- Details of DNA transcription and translation
- Mendelian genetic details

Examples of Performance Expectations

<u>Identifying Science Principles.</u> Students can:
- List three ways in which genes can be altered.

<u>Using Science Principles.</u> Students can:
- Use knowledge of DNA and proteins to explain the basis for differences between two organisms of the same species.
- Explain the basis for particular hereditary disorders, given information about the associated DNA alteration or about critical proteins.
- Predict possible consequences to offspring of an altered DNA sequence in a parent organism.
- Explain possible consequences to an organism's proteins and cells resulting from a change in its DNA.
- Describe the relationship between any two of the following: DNA, genes, proteins, and traits.

Examples of Performance Expectations (cont.)

Using Scientific Inquiry. Students can:

- Given data about particular genetic traits, critique possible explanations for them.
- Use pedigrees to establish when a mutation occurred in a family's history.
- Describe ways to determine possible causes of differences between organisms of the same species.

Using Technological Design. Students can:

- Given information about the relationships between particular DNA sequences, proteins, and traits of various plants within the same species, design the DNA of a plant to exhibit particular traits.
- Describe how to breed a model plant to exhibit a particular trait.

Items to Assess Identifying Science Principles

Item Suggestion 1

List two ways genes can be altered.

Key: Any two of the following: inserting, deleting, or substituting DNA

Item Suggestion 2

Which of the following is **not** a way genes can be altered?

A. Inserting DNA segments
B. Deleting DNA segments
C. Substituting DNA segments
D. Duplicating DNA segments

Key: D

Interpretation: Students answering these items correctly have recalled or recognized the "declarative knowledge" asked for in these items.

Item to Assess Using Science Principles

Item Suggestion

Suppose you observe 1000 flowers of the same species blooming. You observe that all flowers have four petals except for one flower, which has three petals. Assume the DNA of the parent flower had been altered. Describe the sequence of events that could have occurred to result in the three-petal flower.

Key: The DNA segment (gene) for number of petals was altered in some way in the parent flower. That altered DNA segment was passed from parent to offspring through reproduction. The altered DNA segment results in the offspring having three petals, rather than four petals.

Item to Assess Using Scientific Inquiry

Item Suggestion

> Suppose you observe 1000 flowers of the same species blooming. You observe that all flowers have four petals except for one flower, which has three petals. Describe two ways in which you could investigate whether the three-petal trait is due to a recessive gene in normal inheritance or a manifestation of an altered DNA sequence.

Key: Trace the lineage of the flower as far back as possible; compare the molecular sequence of the four-petal flower and the three-petal flower.

Interpretation: This item probes the upper end of the 12th grade achievement scale.

Item to Assess Using Technological Design

Item Suggestion

> Suppose you observe 1000 flowers of the same species blooming. You observe that all flowers have four petals except for one flower, which has three petals. You want to produce the three-petal flower because you believe it has aesthetic and economic value. Applying your understanding of genetics and the life cycle of organisms, how do you make certain you get a consistent set of three-petal flowers?

Key: Successively breed generations of three-petal flowers until they breed true.

Table 20: Grade 8 Earth and Space Science Example of Generating and Interpreting Items

Grade 8: Earth in Space and Time—Objects in the Universe
(The following example also appears in Appendix H.)

Content Statement

E8.2: Gravity is the force that keeps most objects in the solar system in regular and predictable motion. Those motions explain such phenomena as the day, the year, phases of the moon, and eclipses.

Commentary

This content statement encompasses two interrelated sets of concepts:

1. Gravity acts between and among all objects in the solar system, and it plays an essential role in the regular and predictable motions of planets around the sun and satellites around planets.

    - On Earth, gravity is experienced as a force that pulls everything "down" towards the center of the Earth. (A common naïve conception is that the atmosphere "pushing" things down causes gravity.)
    - Gravity is a force of attraction that is exerted by every object on every other object.
    - Gravity exists in space and on other planets. (A common naïve conception among students is that there is no gravity in space because space has no air.)
    - The almost circular motion of planets and satellites results from the force of gravity and the tendency of a body to continue moving through space in a straight line unless acted upon by a net force.

Commentary (cont.)

2. The regular and predictable motions of the Earth, the sun, and the moon cause the cyclic phenomena that can be observed in the sky.

   - The day-night cycle results from Earth's rotation on its axis once in 24 hours.
   - Annual changes in the visible constellations and the seasons result from Earth's revolution around the sun once every 365-1/4 days.
   - Moon phases result from the moon's orbit around the Earth about once a month, which changes what part of the moon is lighted by the sun and how much of the lighted part can be seen from Earth.

Note connection between this content statement and the Physical Science subtopic, "Forces Affecting Motion."

Students are not expected to use the inverse square relationship of gravitational force and distance to find the strength of the gravitational force between two objects.

Students need not know that the motion of planets and satellites is elliptical and not circular.

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Identify gravity as the force exerted by every object in the solar system on every other object.
- Identify gravity as the force that keeps the moon circling Earth, rather than flying off into space.
- Describe the regular motions of Earth through space, including its daily rotation on its axis, and its yearly motion around the sun.

Using Science Principles. Students can:

- Explain that the orbit of one object around another is due to the tendency of an object to move in a straight line through space and the force of gravity between the two objects.
- Explain how the monthly pattern of moon phases observed from a point on Earth results from the moon's orbit around Earth, which changes what part of the moon is lighted by the sun and what portion of the lighted part can be seen from Earth.
- Distinguish between explanations for lunar (moon) phases and lunar eclipses.
- Explain that astronauts and other objects in orbit seem to "float" because they are in free fall, under the influence of gravity.

Using Scientific Inquiry. Students can:

- Arrange a set of photographs of the moon taken over a month's time in chronological order and explain the order in terms of a model of the Earth-sun-moon system.
- Design a plan for observing the sun over a year's time to find out how the length of the day is related to the rising and setting point of the sun on the horizon.
- Design a series of observations or measurements to determine why some objects—such as certain asteroids or comets—visit the solar system just once, never to return.

Using Technological Design. Students can:

- Choose among several (qualitative) methods for aiming a rocket so that it reaches the planet Mars and give a rationale that shows understanding of orbital motion.
- Use scientific trade-offs in deciding whether or not to support a plan to observe and predict orbits of asteroids that enter the inner solar system.
- Given a scenario in which a person is shipwrecked on an island in the ocean, critique plans to create a calendar to keep track of the passage of time.

Items to Assess Identifying Science Principles

Illustrative Item 1

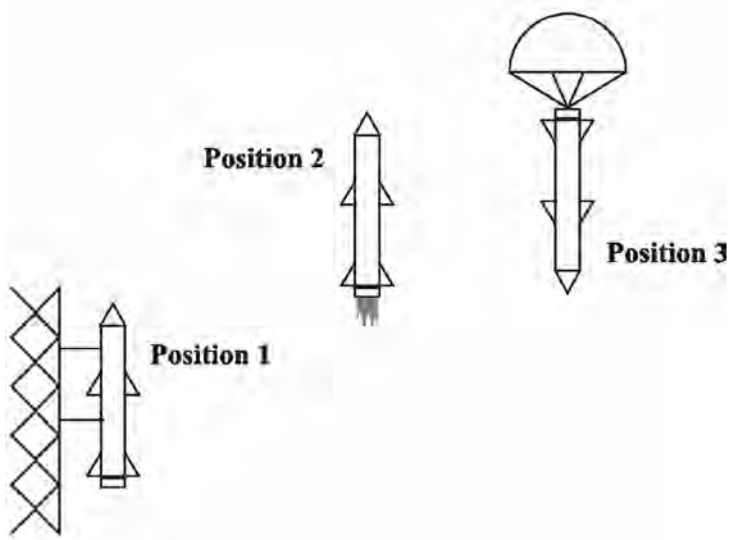
What force keeps the planets in our solar system in orbit around the Sun?

A. gravitational
B. magnetic
C. electrical
D. nuclear

Key: A
Source: Adapted from Massachusetts Department of Education, Massachusetts Comprehensive Assessment System (MCAS), 2000, Grade 8[27]

Illustrative Item 2

The drawings show a rocket being launched from Earth and returning.



In which of these positions does gravity act on the rocket?

A. 3 only
B. 1 and 2 only
C. 2 and 3 only
D. 1, 2, and 3

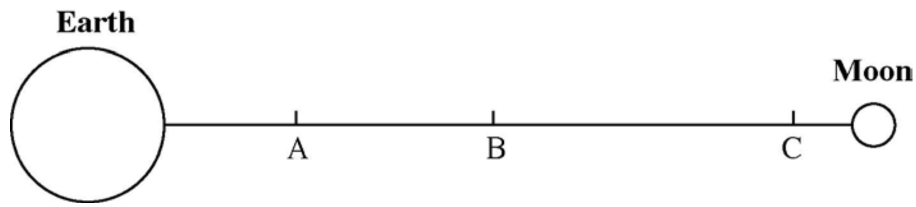
Key: D
Source: TIMSS 1999, Grade 8

[27] MCAS materials appearing in the *Specifications* have been released to the public by the Massachusetts Department of Education and is available at no cost on the Department's website at http://www.doe.mass.edu/mcas/testitems.html

Items to Assess Using Science Principles

Illustrative Item

A space station is to be located between the Earth and the Moon at the place where the Earth's gravitational pull is equal to the Moon's gravitational pull.
On the diagram below, circle the letter indicating the approximate location of the space station.



Explain your answer.

Source: NAEP 1996, Grade 8

Interpretation: The correct answer is C. Since the moon has 1/6 the amount of gravity as Earth, a body that experiences an equal gravitational force from Earth and the moon should be closer to the moon. Point C is the only point that is closer to the moon. Note: Point C is about 1/12 of the way between the moon and Earth; it should be 1/6 of the distance. (See Appendix D for the item scoring guide.)

Item Suggestion

Is there gravity in space? Which of the following gives the best response to this question?

A. No. You can see that astronauts float around weightless in their cabin.
B. No. There is no air in space, so gravity cannot exist there.
C. Yes. There must be gravity since planets keep circling the sun.
D. Some. The moon has one-sixth as much gravity as Earth, so we know there is some gravity in space.

Key: C

Interpretation: The correct answer is C. This question is drawn from a series of studies that show a common naïve conception—that there is no gravity in space because space has no air. The distractors are drawn from student interviews. It is likely that these naïve conceptions stem from images that students have seen of astronauts floating around in a "weightless" environment while in orbit. This item probes "schematic knowledge."

Item to Assess Using Scientific Inquiry

Item Suggestion

> A student is presented with a set of photographs of the moon taken over a month's time. The photos are not presented in chronological order. The student is asked to arrange them in the order in which they were taken and explain the reason for moon phases.

Interpretation: This suggestion reflects items used frequently in curricular materials (e.g., Schatz & Cooper, 1994). Students are asked to find patterns in the data. First, they should be sufficiently familiar with the lunar cycle to arrange the photographs in order, either in a line to represent a chronology or in a circle to represent a cycle (tapping "declarative knowledge" and "procedural knowledge" to a lesser extent). Then, students should be able to explain moon phases in terms of the moon circling Earth and the changing angle between the sun and moon as observed from Earth. This part of the item probes "schematic knowledge." This is a challenging question that many educated adults fail. However, studies show that middle school students can learn to do this by observing lunar phases and explaining them using a model of the Earth-sun-moon system (Barnett & Morran, 2002; Kavanagh, Agan, & Sneider, 2005).

Items to Assess Using Technological Design

Item Suggestion 1

> NASA wants to launch a spacecraft with rockets from Earth so that it will reach and orbit Mars. Which of the following statements about this flight is WRONG?
>
> A. In the first phase of its flight, the forces acting on the spacecraft are the thrust of the rocket engine, gravity, and friction from the Earth's atmosphere.
> B. When the rocket engine shuts off, the only force acting on the spacecraft is the force of gravity.
> C. Once the spacecraft is above the Earth's atmosphere and the rocket engine is off, it will travel at a constant speed since there is no gravity in space.
> D. If the spacecraft is aimed correctly and has the proper speed, the spacecraft will reach Mars and require only engine braking to attain orbit.

Key: C

Interpretation: The correct answer is C, since there is gravity in space and planning for such a rocket flight would need to take into account the gravity from Earth, Mars, and the sun ("declarative knowledge"). This question is drawn from a series of studies that show that the following naïve conceptions about gravity are common among many students at the middle school, high school, and even college levels: If a body is moving, there is a force acting on it in the direction of motion (Finegold & Gorsky, 1991; Gunstone & Watts, 1985; Sequeira & Leite, 1991); there is no gravity in space (Bar, Zinn, Goldmuntz, & Sneider, 1994; Chandler, 1991; Morrison, 1999); and gravity cannot act in space because there is no air in space (Bar & Zinn, 1998). One study showed that, with effective instruction, middle school students can overcome these naïve conceptions and learn that gravity does, in fact, act in space, where it keeps satellites and planets in their orbits (Bar, Sneider, & Martimbeau, 1997).

Items to Assess Using Technological Design (cont.)

Item Suggestion 2

Decisions about whether or not to develop new technologies always concern trade-offs. For example, most scientists today believe that the extinction of the dinosaurs and many other species was caused by the collision of a large asteroid with Earth 65 million years ago. As a result, there is a proposal to develop two new technologies: (1) the detection and tracking of all asteroids large enough to do considerable damage if they should strike Earth, and (2) the development of means of sending a spacecraft to meet the asteroid in space and change its path. Write a paragraph describing:

a. whether or not you think it would be possible to develop these technologies based on your knowledge of science; and
b. some of the scientific trade-offs that should be considered in deciding whether or not to develop these new technologies.

Interpretation: Look for evidence that the students understand what asteroids are, that scientists have observed asteroids, and that observations taken at several points in time allow for the prediction of an asteroid’s path. Also look for evidence that students understand that spacecraft can be built, launched, and navigated to intercept solar system bodies. (In fact, several spacecraft have intercepted asteroids and comets.) Regarding scientific trade-offs, look for evidence that students recognize the advantages of the proposed technologies (e.g., avoid a catastrophic collision in which billions of people and animals could die) as well as possible negative effects (e.g., break-up of the asteroid so there are many collisions rather than one, accidents on launch).

Learning Progressions

A learning progression is a sequence of successively more complex ways of reasoning about a set of ideas. For any important set of ideas in science, understanding increases over time as students learn more and more, moving from initially naïve knowledge of the natural world to increasingly more sophisticated knowledge and conceptual understanding; and this typically occurs in conjunction with educational experiences in and out of school (NRC, 2001). In other words, the progression from novice learner to competent learner to expert begins with the acquisition of relevant experiences, principles, concepts, facts, and skills and moves to the accumulation and organization of knowledge in a specific domain and finally to expertise *after extensive experience and practice* (e.g., Ericsson, 2002). The attention paid to growth of understanding may yield rich information about student progress.

Research has been conducted on students’ learning progressions in some areas of science and at some levels of students’ development. It is expected that this research will directly inform the development of assessment items at grades 4, 8, and 12. For example, the National Research Council (NRC) has commissioned papers on learning progressions in evolution (Catley et al., 2005) and in atomic molecular theory (Smith et al., 2004). Learning progressions provide opportunities for assessing specific content in greater depth.

Several caveats about learning progressions are in order. First, learning progressions are not developmentally inevitable but depend upon instruction interacting with the student's prior knowledge and construction of new knowledge. Thus, learning progressions will need to invoke assumptions about instruction. Second, there is no single "correct order." There may be multiple pathways by which certain understandings can be reached. Which pathway is taken may be influenced by prior instructional experiences, individual differences, and current instruction (NRC, 1999c, 2001). Thus, learning progressions will necessarily be complex, involving multiple specific paths at the micro level, and will need to be described in ways that encompass such diversity. Third, actual learning is more like ecological succession with changes taking place simultaneously in multiple interconnected ways. Thus, attempts to describe specific sequences of learning performances, including those in the Catley et al. (2005) and Smith et al. (2004) papers, must inevitably be artificially constrained and ordered. Finally, the learning progressions suggested in the *Framework* and *Specifications* are partly hypothetical or inferential, since long-term longitudinal accounts of learning by individual students do not exist.

Examples of learning progressions for "Floating, Sinking, and Density" "Earth in the Solar System," and "States of Matter," as well as examples of how learning progressions can inform item writing, are provided in Appendix I.

CHAPTER FIVE: TYPES OF ITEMS

Introduction

The performance expectations set forth in the last chapter reflect the variability and complexity of science knowledge, reasoning, and practices. In order to capture this variability and complexity in science, a variety of different assessment items is needed. This chapter specifies the nature of the assessment items to be found on the NAEP Science Assessment; these specifications apply to all grade levels assessed.

The assessment items envisioned for the NAEP Science Assessment differ in a number of ways so as to tap into the various performance expectations. First, the items vary in the response demands they set for students (e.g., selecting a response from a set of alternatives or constructing a response as a single word, short answer, essay, explanation, prediction, justification, summary of research findings). Second, the items vary in the stimuli needed to evoke the knowledge and thinking underlying the performance expectation (e.g., verbal, tabular, graphic, manipulative, computer simulated). Third, the items vary as to how they are scored (e.g., right or wrong, partial credit, scored by humans or computer software). Fourth, the items may either stand alone or be part of a cluster of items, which probe in some depth an important performance expectation. Fifth, the items vary as to the context into which they are set (e.g., history and nature of science, relationship between technology and science).

This chapter highlights some of the critical considerations in item development and concentrates on topics specific to the NAEP Science Assessment. Item writers should refer to directions for developing items provided by the assessment development contractor in addition to the information here.

Three types of textboxes are used throughout this chapter. A *Clarification* textbox provides details on a type of item. *Illustrative Item* textboxes provide assessment items that exemplify recommendations discussed in the text. Answers to selected-response illustrative items are indicated within the textbox; scoring guides for constructed-response illustrative items are provided in Appendix D. *Item Suggestion* textboxes also serve to illustrate points made in the text, but these textboxes present item ideas that require further development in that they are neither published nor field-tested items. Answers to selected-response item suggestions are indicated within the textbox; however, scoring guides have not been developed for constructed-response item suggestions. Although items in both Illustrative Item and Item Suggestion textboxes may assess more than one content statement or practice, only the primary content and practice designations are provided. This follows NAEP practice, which uses only primary designations for items in the analysis and reporting of student responses.

Principles of Good Item Writing

The principles of good item writing, first and foremost, seek to create assessment items that prompt *all* students, regardless of background, to respond in ways that display their science knowledge and reasoning. In other words, the intent is to create items that, as best as possible, tap students' knowledge and reasoning without special accommodations for certain groups of students. This goal can be seen in textbooks on item writing as well as in the principles espoused by proponents of "universal design" (Thompson, Johnstone, & Thurlow, 2002). The principles specified below build upon this goal. Nevertheless, accommodations for students with special needs (students with disabilities, English language learners) are appropriate when the principles are followed but additional considerations are necessary; the topic of accommodations is taken up in Chapter Six.

Before presenting these principles, one caveat is in order. The following principles of good item writing should be used as appropriate, depending upon the measurement intent of the item. For example, a principle dealing with item graphics states, "represent each important part of the item in the visual images." If the intent of the item is to have students recognize the importance of a piece of given information and use that information in their solution, then it might not be appropriate to include the information in the accompanying graphic. For example, a graphic display of distance traveled and time to travel should not include speed information if students are expected to calculate speed. Thus, these guidelines should be followed unless the targeted construct precludes doing so. For more information about grade-appropriate use of examples, instruments, units of measurement, representations, and technical vocabulary, see sections on content boundaries in Chapter Two.

Clear Measurement Intent

A critical step in good item writing is making sure that the measurement intent of the item is clear and that students understand what is being asked and what type of response is expected.

Clear Intent in Development

Item writers should provide a clear description of what each item is intended to measure. This will help classify items according to assessment specifications (content statement, practice, and cognitive demand), help develop clear scoring rubrics and scoring materials, reduce confusion in reviews, and provide evidence of the degree of alignment of the assessment to the *Framework* and *Specifications*.

The response to one item should not depend upon the response to another. For example, students should not be asked to determine the mass of an object in one assessment item and then use that measurement to determine the density of the object in another assessment item. Items can be related to one another, for example, through being based on the same graphic display or by having the same theme; but their response requirements must be independent.

Clear Intent for Assessment Takers

It should be clear to the student what is being asked for in each item. Writers should be careful not to make assumptions about how students will interpret an item's implicit requirements.

Constructed-response items should contain clear directions to students about how they can respond. For example, can the response incorporate graphics, or does it require the student to produce a verbal description? Is more than one type of response appropriate? Is more credit given for providing alternative responses to a problem? If, for example, a constructed-response item requires students to explain their answer or justify a prediction, the directions must clearly specify that students should include an explanation or justification and should note that it will be used in scoring the response.

Clarity of intent should not be taken to the extreme by over-scaffolding an item. Too much structure may result in measurement of a construct that is different from what was originally intended. For example, the rigid structure of some hands-on performance tasks may be result in measuring students' ability to follow step-by-step instructions rather than their ability to conduct relevant inquiry regarding a problem.

Plain Language

Plain language is a writing and editing tool designed to clearly convey meaning without altering what items are intended to measure. All items should use clear and concise language. Even when the intent of the item is for students to define, recognize, or use science vocabulary correctly, the surrounding text should be in plain language. Plain language guidelines often increase access and minimize confusion for students. The overarching principle is that assessment items should measure the performance expectation intended, not reading (or some other) ability (i.e., not "construct-irrelevant" performance). Some general guidelines are as follows:

- The reading level of the assessment should be below that of the science grade level assessed. Note that reading level must be determined by using data from pilot testing (e.g., from interviews in which students are asked what they think the item is about and what words or sentences they have problems understanding) and the judgment of experts (e.g., linguists and reading specialists, and teachers of the grade of the targeted reading level). Determining the reading level of assessment items should not rely on the use of readability formulas such as those included in some word processing software packages. To produce adequate measures of text complexity, readability formulas need to be developed for specific populations and specific types of text. Moreover, they need to be used with large amounts of texts, not assessment items, which, by definition, have small amounts of text (see Oakland & Lane, 2004).
- Write questions using brief, clear, and concise sentences or stems.

- Use the same structure for paragraphs throughout the assessment as much as possible (e.g., topic sentence, supporting sentences, concluding sentence).
- Use present tense and active voice.
- Do not combine interrogative and conditional modes in the same sentence (e.g., *How* many pens can a person buy *if* that person has $3.37*?*).
- Minimize paraphrasing.
- Avoid using pronouns; but if pronouns are necessary, be sure that pronoun references are clear.
- Avoid double negations (e.g., it is *not unusual*).
- Use high-frequency words as much as possible.
- Avoid colloquialisms, regionalisms, and contextual situations that are more familiar to certain socio-economic groups. Avoid stereotyping and racial, cultural, gender, and regional bias. See Appendix J, NAEP Item Development and Review Policy Statement, guiding principle 6.
- Use the science language (vocabulary) that appears in the content statements in Tables 6, 9, and 12. (Rare exceptions are noted in the "Content Boundaries" sections in Chapter Two.)
- Avoid using words with multiple meanings. If it is necessary to use words with multiple meanings, make sure the intended meaning is clear.
- Avoid using extraneous descriptive information unless it is related to the item's intent.
- Use format to clarify text (e.g., use bullets, allow space between pieces of text, use boxes and lines judiciously).
- Mathematics level should be approximately one to two grade levels below the science assessment grade level so as not to confound science achievement with mathematics achievement.

Graphics

Graphics such as pictures, charts, and diagrams are visual images reflecting information for an item or a set of items. Graphics can be very effective in supporting text, illustrating science concepts in the text, and increasing item access with minimal cost. If used improperly, however, graphics can add substantial confusion and distract assessment takers from what the items are asking students to do. When using graphics:

- Use visuals that mirror and parallel the wording and expectations of the text.
- Avoid illustrating extraneous information in the graphics unless it is related to the item's intent.
- Represent each important part of the item in the visual images.
- Pay attention to clarity and scale (see p. 207 for frequently used accommodations for visually impaired students).

Using Contextual Information Appropriately

Often, items will be designed so that they measure science in context (e.g., history and nature of science, relationship between science and technology). Contextual information includes problem scenarios, explanations, thorough directions, and background text. Using contextual information judiciously can place science concepts in full, often realistic, conditions. However, the contextual information should not interfere with the science being assessed or become a barrier to students' ability to demonstrate their science knowledge. Guidelines for using contextual information appropriately include the following:

- Use clear and concise language as much as possible.
- Use activities and graphics to increase item clarity.
- Spread the directions and explanations throughout the item or block instead of frontloading all of the necessary information.
- Use contexts that are meaningful to the science being assessed.
- Use contexts that are appropriate for the grade level assessed; if necessary, provide background information that assessment takers may not have.
- Use familiar contexts; avoid contexts that may confuse or be unfamiliar to some students taking the assessment.
- Avoid contextual information that could corrupt the measurement of the intended skill.

Writing Items with Multiple Access Points

Students vary in their abilities to access information and respond to tasks through visual, spatial, auditory, kinesthetic, and tactile pathways. When possible, items should be designed to allow students to approach and respond to the item via different pathways. Incorporating multiple pathways appropriately in both constructed-response and multiple-choice items can increase the ability of the assessment to elicit responses from students across the range of achievement without affecting the science content being assessed.

Guidelines for providing multiple access opportunities in how items are written or presented to students include the following:

- Use visuals to accompany and explain the text when appropriate.
- Use science laboratory equipment or materials.
- Use a brief activity when appropriate to address students with auditory, kinesthetic, or tactile strengths.

Guidelines for providing multiple response opportunities include the following:

- Write items that allow for multiple response formats. For example, students might be allowed to show their answers through illustrations, diagrams, or formulas.
- When possible, write items and scoring rubrics that measure the science knowledge and skills from students with a range of achievement.

Types of Items

The judicious selection of items lies at the heart of any effective assessment of science achievement. The Framework for the 1996-2005 NAEP Science Assessments called for three types of items: multiple-choice items (selected-response), open-ended paper-and-pencil items (constructed-response), and performance exercises. Multiple-choice items made up about 40% of the assessment, as measured by student response time, with open-response items comprising about 60% of assessment time. In addition, subsets of the students sampled were given an extra 20 minutes in grade 4 and 30 minutes in grades 8 and 12 to complete hands-on performance tasks. The *Framework* and *Specifications* generally follow the 1996-2005 recommendations in item structure but go beyond by specifying additional item types—some selected-response and others constructed-response. The 2009 recommendation for item distribution by item format, in terms of student response time, is 50% selected-response and 50% constructed-response.

Two further considerations are the need to develop items that probe students' ability to use communication skills and quantitative reasoning skills in science (see pp. 111-112). While there is no prescription in the *Specifications* about the amount of assessment time to be spent on items that require specific forms of communication or application of mathematics, it is expected that items requiring these skills will be represented at all three grade levels. As NAEP will not provide calculators on the Science Assessment, any required mathematical computations should be easily accomplished without the use of a calculator.

Justification for Variation in Types of Items

Issues of time and cost are paramount in any assessment. Accordingly, most of the item formats on the NAEP Science Assessment will be rather traditional selected-response and short constructed-response. However, some more complex items (e.g., hands-on performance tasks) should be part of any science assessment.

Responses to more complex items often correlate positively and cluster with responses to more efficient and simpler items. However, complex items are recommended in the *Framework* and *Specifications* for the following reasons:

- Items may correlate positively with one another, but they do not necessarily measure the same thing, that is, positive correlations can arise even when the cognitive demands of the assessment items vary. Research has shown that items vary in their cognitive demands for different kinds of knowledge and reasoning (e.g., Ayala et al., 2002; Leighton, 2004; Martinez, 1999).
- The NAEP Science Assessment signals the kinds of tasks, problems, and exercises, along with the kinds of knowledge and reasoning, that should be expected of students as a result of what is taught in the science curriculum, consistent with the *National Standards* and *Benchmarks*.

For these reasons, the *Framework* and *Specifications* call for a variety of item types for the 2009 NAEP Science Assessment.

Definitions of Types of Items

The *Framework* and *Specifications* distinguish selected-response from constructed-response item formats. For selected-response formats, students respond to a question by selecting the answer they believe to be most scientifically justifiable from a given set of alternatives. In contrast, with constructed-response formats, students respond to a question by "generating" or "constructing" a response. The constructed-response might be a single word, a short answer, an essay explanation, a summary of a laboratory investigation using concrete materials, or typed responses to a computer simulation.

In addition to these two main item formats, there are combination items that generally require more than one response. These include item clusters, Predict-Observe-Explain (POE) item sets, hands-on performance tasks, and interactive computer tasks. These combination items can use an all-selected-response format, an all-constructed-response format, or a mixture of these two main item formats.

Following are the main types of items to be used on the 2009 NAEP Science Assessment:

1. Selected-response
    - Individual multiple-choice items
2. Constructed-response
    - Short constructed-response items
    - Extended constructed-response items
    - Concept-mapping tasks
3. Combination items
    - Item clusters
    - Predict-Observe-Explain (POE) item sets
    - Hands-on performance tasks
    - Interactive computer tasks

Item clusters and POE item sets may use selected-response items, constructed-response items, or both. For example, a set of Predict-Observe-Explain items might include a multiple-choice item in which students select a prediction and a short constructed-response item in which students are asked to write a justification for their prediction.[28]

Hands-on performance tasks and interactive computer tasks also may use selected-response items, constructed-response items, or both. In recording their answers, students may be asked to respond to both selected-response and constructed-response items. For example, 12th graders might be asked to manipulate a computer simulation of a chemical reaction; a multiple-choice question could ask students to choose the correct mass of a

[28] See p. 177 for discussion of potential non-independence of items.

reaction product; and a short, constructed-response question could ask students to describe how mass is conserved in chemical reactions.

Selected-Response Items

Selected-response items include individual multiple-choice items. Following is a description of these types of items, as well as guidelines for their development.

Individual Multiple-Choice Items

Selected-response items most often take a multiple-choice format. Students read, reflect, and then select an answer from, say, four alternatives provided. The alternatives include the most scientifically justifiable response—the “answer”—as well as three “distractors.” The distractors should appear plausible to students but should not be scientifically justifiable; and, when feasible, the distractors should also draw from current understanding about students’ mental models and learning progressions. Whenever possible and especially when the focus is on Using Science Principles or “knowing why (schematic knowledge),” naïve conceptions, explanations, and predictions of the natural or human-made world should serve as distractors.

Developing Multiple-Choice Items

Multiple-choice items are an efficient way to assess knowledge and skills, and they can be developed to measure many but not all of the performance expectations. Well-constructed multiple-choice items can be developed to probe important facts, broad concepts and themes of science, as well as probe deductive reasoning skills.

Multiple-choice items have three components: (a) the stem, (b) an option that is the correct response, and (c) distractors or options that are incorrect. In a well-designed multiple-choice item, the stem presents the problem clearly to the student. The stem may be in the form of a question or a statement as long as it conveys what is expected of the student. The stem is followed by four possible alternatives, only one of which is correct; the remaining are “distractors.” In good multiple-choice items:

- The stem includes only the information needed to make the student’s task clear or to set the problem in an appropriate context; avoid using extraneous descriptive information unless it is related to the item’s intent.
- The stem presents a grammatically complete question or presents a statement, which includes a blank that can be replaced with one of the options.
- The stem should avoid negative phrasing (e.g., “all of the following *except*”; “which one is *not*”).
- The following is avoided: use of the same word, phrase, or concept in both the stem and the correct answer but not in any of the distractors.

Multiple-choice items have the following requirements for options or alternatives:

- There must be only one clearly identifiable correct option for each question. If scientifically justifiable arguments can be made for more than one option, then the item is unacceptable.
- The correct option must be a concise answer that will satisfy any qualified judge as being an adequate short answer to the question. The response must not answer more than the stem question asks.
- Distractors (foils) should be incorrect due to the examinee’s predisposition, unsound reasoning, misinformation, naïve science conception, or casual reading of the responses.
- In items requiring the use of mathematics, distractors should probe students’ science knowledge (e.g., correct use of a formula), not common errors in calculation.
- Distractors should be plausible to students. Options that are obviously wrong or silly effectively reduce the number of possible correct answers and, thus, reduce the potential value of the item. Sources of good distractors include common misinterpretations, naïve conceptions and errors in reasoning, statements that are true but that are not correct answers to the questions posed in the stems, and carefully worded incorrect statements that may sound plausible to the uncritical thinker.
- Distractors must be written with as much care and precision as the correct option so that all alternatives are equally attractive to an examinee who guesses. Each option should be a separate and distinctly different response to the stem. Responses should not overlap or include other responses.
- All options should be parallel in point of view and grammatical structure.
- All options should be similar in length.
- Options such as “none of the above” or “all of the above” should not be used.

Examples of multiple-choice items that fulfill these characteristics are shown below. Both require Identifying Science Principles, and both tap the cognitive demand, “knowing that (declarative knowledge).” The first item taps simple factual content. The second item taps more conceptually sophisticated content. For some students, if they cannot easily recall the science content, this second item may require Using Science Principles (i.e., tapping more of “knowing why” than “knowing that”).

Illustrative Item

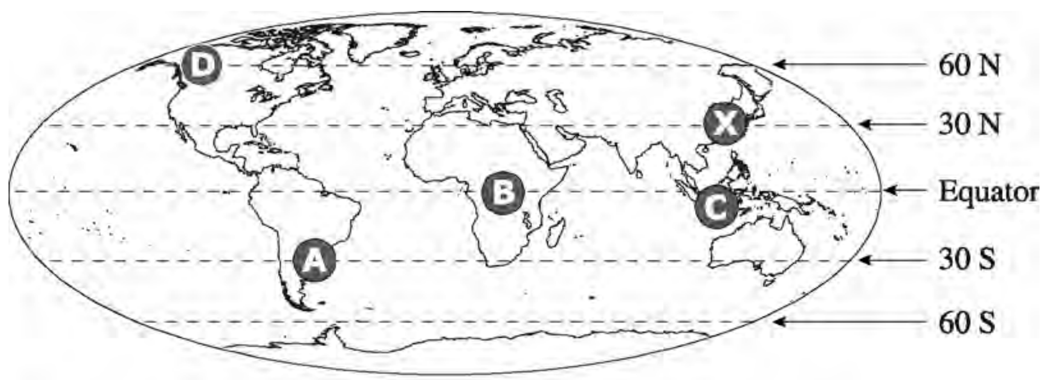
Air is made up of many gases. Which gas is found in the greatest amount?

A. Nitrogen
B. Oxygen
C. Carbon dioxide
D. Hydrogen

Key: A

E8.7, Identifying Science Principles
Source: TIMSS 1995, Grade 8

Illustrative Item



The diagram above shows a map of the world with the lines of latitude marked. Which of the following places marked on the map is most likely to have an average yearly temperature similar to location X ?

A. A
B. B *
C. C
D. D

Key: A

E8.12, Identifying Science Principles
Source: TIMSS 2003, Grade 8

Constructed-Response Items

Constructed-response items include short constructed-response items, extended constructed-response items, and concept-mapping tasks.

Constructed-response items can be used to provide insights into students' levels of conceptual understanding and can serve to assess their abilities to communicate in science. They can also be used to probe students' abilities to generate information related to science content statements and their interconnections (e.g., how two or more cyclic events are related). Constructed-response items can also be particularly useful for probing the practices of Using Scientific Inquiry or Using Technological Design (e.g., interpret given data, provide a solution to a real-world problem).

The type of constructed-response item used depends on the science performance expectation being assessed. The general components of a constructed-response item are (a) a question to which the student must respond, (b) stimulus material (e.g., selected text, data table) that the student must use, (c) instructions for how the student is to use the material and that suggest limitations or structure of the desired response, and (d) a scoring rubric.

Developing Constructed-Response Items and Scoring Rubrics

Item writers should draft each constructed-response item to tap a particular performance expectation, that is, the type of constructed-response item being developed—the nature of the task and the response format—depends on the performance expectation to be assessed. Stimulus material should be generated to allow for a variety of student responses, including numbers, diagrams, drawings, graphs, or written narrative responses. For development of stimulus materials, constructed-response items can utilize text taken from materials students commonly encounter (e.g., magazines, newspapers, textbooks, monographs). Whenever possible, stimulus material should draw upon subject matter that is of interest to students. Students can be asked to read a text (written at a reading level below the grade assessed), generate data from textual, graphical, or tabular material, and/or draw conclusions from such material and the data they generate. However, in choosing stimulus material, care must be taken to minimize the potential for inequity, since some students may have greater exposure to certain materials and thus possess an unfair advantage (see Lee, 1999).

The initial scoring rubric should be developed at the same time as the item to ensure that both the item and the rubric reflect the construct being measured. In developing the scoring rubric for an item, writers should think about the performance expectation measured and the kinds of student responses that would show increasing degrees of knowledge and understanding. Indeed, where research literature exists on the performance expectation measured, writers should be familiar with and consult this literature. Writers should sketch condensed sample responses for each level of knowledge and understanding and create corresponding score categories (e.g., correct, partially correct, incorrect). Item writers also should include justification or explanation for each rubric category description. Doing so will assist the writer in developing a clear scoring rubric as well as provide guidance for scoring the item. Item writers should refer back to the measurement intent of the item when they are developing its scoring rubric. The number of categories used in the rubric should be based upon the science demand of the item, that is, the scoring system should match the task demands and the anticipated levels of response quality. As the item goes through the review process and student responses to the item become available, the scoring rubric should be updated to reflect these empirical data.

Aligning Items and Rubrics: Defining the Score Categories

Each score category (e.g., correct, partially correct, incorrect) must be distinct from the others; descriptions of the score categories should clearly reflect increasing understanding and skill in the science constructs of interest. Distinctions among the categories should suggest the differences in student responses that would fall into each category; the definitions must be clear enough to use in training scorers. Each score level should be supported by the intent of the item. Factors unrelated to the measurement intent of the item should not be evaluated in the rubric. For example, if an item is not meant to measure writing skills, then the scoring rubric should be clear that the demonstration of science in the response does not need to be tied to how well the response is written. However, if an extended explanation is part of the item requirement, the rubric should reflect that such explanations should be clear and understandable.

Aligning Items and Rubrics: Measuring More than One Concept

If an item is measuring more than one skill or concept, the description of the score categories in the rubric should clearly reflect increasing understanding and achievement in each area. For instance, if a performance assessment task is measuring both students' understanding of a science topic and their skill in developing an appropriate problem solving procedure, then the description of each category in the rubric should explain how students' understanding and skill are evaluated. If an item requires both an acceptable solution and procedure, then the rubric should show how these two requirements are addressed in each score category.

Aligning Items and Rubrics: Specifying Response Formats

Unless the item is measuring whether or not a student can use a specified approach to a problem, each score category should allow for various approaches to the item. It should be clear in the rubric that different approaches to the item are allowed. Varied approaches may include:

- Using different problem-solving procedures pathways (e.g., controlling and isolating variables in an experiment may lead to the same solution, no matter the order in which variables were tested).

- Using different formats (e.g., it may be appropriate to use a graph, diagram, or mathematical formula to solve or explain a solution to a question on an object's speed).

As previously discussed in the section "Principles of Good Item Writing," care must be taken on constructed-response items to identify the response elements that should be present to constitute a satisfactory answer; that is, the stimulus must be defined to indicate to the student how the item will be scored. Whenever an objective interpretation is desired, information must be presented (scaffolded) in the stimulus that informs students which essential elements to include in their answers (without giving the answer away). Care must be taken, in order to ensure exercise reliability, to scaffold all constructed-response items that elicit the same type of answers the same way.

Scorer training materials should include examples of student work that illustrates various appropriate and inappropriate approaches to the item. Moreover, training should contain borderline ("fence-sitters") papers that explain important distinctions between scoring categories; the absence of such explanations could give rise to large-scale disagreement among scorers. Scorer training and scoring protocols must emphasize avoiding biases toward longer or shorter responses.

Short Constructed-Response Items

This item type generally requires students to supply the correct word, phrase, or quantitative relationship in response to the question given in the item, illustrate with a brief example, or write a concise explanation for a given situation or result. Thus, students must generate the relevant information rather than simply recognize the correct answer from a set of given alternatives, as in selected-response items.

Some short constructed-response items are written to be scored dichotomously. Short constructed-response items with two scoring categories should measure knowledge and skills in a way that multiple-choice items cannot or provide greater evidence of the depth of students' understanding. Such short constructed-response items might be appropriate for measuring some conceptual knowledge, for example, to avoid guessing or when recall rather than recognition—a factor in a multiple-choice item—is important. Short constructed-response items are also useful when there is more than one possible correct answer, when there are different ways to display an answer, or when a brief explanation is required. Item writers should take care that short constructed-response items would not be better or more efficiently structured as multiple-choice items; they should not be simply multiple-choice items without the options. Item writers must draft a scoring rubric for each short constructed-response item. For dichotomous items, the rubrics should define two categories—0 and 1.[29]

[29] These categories could be labeled as "1=correct" and "0=incorrect." For three scoring categories, appropriate labels might be "2=correct," "1=partial," and "0=incorrect." However, the exact labels for scoring categories are as yet undetermined and will be consistent with NAEP practice across all subject areas.

The following illustrates an item with two score categories.

Illustrative Item

On a hot, humid day the air contains a lot of water vapor. What happens to the water vapor in the air when the air becomes very cold?

(See Appendix D for item scoring guides.)

P4.6, Identifying Science Principles
Source: TIMSS 2003, Grade 4

Some short constructed-response items are written to be scored on a three-category scale—0, 1, and 2. Short constructed-response items with three scoring categories should measure knowledge and skills that require students to go beyond giving an acceptable answer. These items allow for degrees of accuracy in a response so that a student can receive some credit for demonstrating partial understanding of a concept or skill. For items with three score categories, the rubrics should clearly define each category.

The following two textboxes illustrate items with three score categories.

Illustrative Item

When a population of mice is infected with parasites, many of the mice die from the parasitic infection, but some mice appear as healthy as they were before being infected. Some people are considering using these parasites to control the mouse population in people's homes.

Give one advantage and one disadvantage of using these parasites instead of mouse traps or poisons to limit the population of mice.

(See Appendix D for item scoring guides.)

L8.6, Using Technological Design
Source: NAEP 1996, Grade 8

Illustrative Item

One day when the temperature was just below 0°C, Peter and Ann made snowballs. They put a thermometer into one of the snowballs and it showed 0°C. They tried to make the snowball warmer by holding it in their hands. What do you think the thermometer showed after two minutes? Explain your answer.

(See Appendix D for item scoring guides.)

P8.6, Using Science Principles
Source: TIMSS 1995, Grade 8

Extended Constructed-Response Items

This item type is generally multi-dimensional; that is, it taps into multiple content statements, practices, and/or cognitive demands. These types of items can provide particularly useful insight into students' level of conceptual understanding and reasoning. They can also be used to probe students' ability to communicate in the sciences. Such items generally present a situation within or across content areas and require students to analyze the situation, choose and carry out an alternative plan for addressing it, and interpret their response in light of the original situation. Students may also be given an opportunity to explain their responses, their reasoning processes, or their approach to a problem situation. However, care must be taken, particularly with 4th graders and English language learners, that language ability is not confounded with science ability.

In general, extended constructed-response items ask students to solve a problem by applying and integrating science concepts or require that students analyze a science situation and explain a concept, or both. Extended constructed-response items should typically have no more than five scoring categories—0, 1, 2, 3, and 4.

The final decision about the number of categories should depend on the nature of the task and the performance expectation measured.

Item writers should develop a draft scoring rubric specific to each extended constructed-response item. The rubric should clearly reflect the measurement intent of the item.

The following item involves reasoning with "mental models" (on carbon cycling) and thus attempts to probe the practice of Using Science Principles and taps into the cognitive demand of "knowing why."

Illustrative Item

> The biosphere (living organisms), the lithosphere (rocks and soils of the Earth's crust), and the atmosphere are all involved in the cycling of carbon atoms. Describe the role that each plays in the carbon cycle.

(See Appendix D for item scoring guides.)

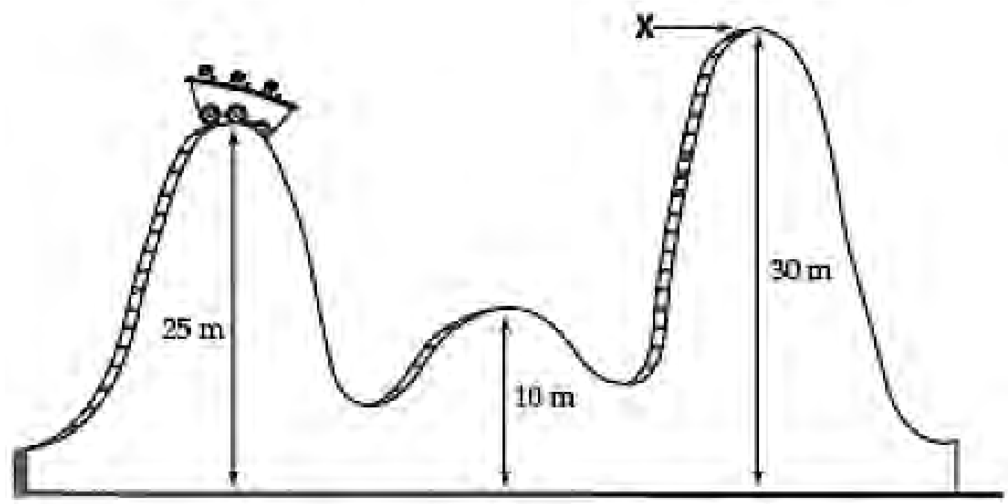
E12.12, Using Science Principles
Source: New Standards Spring Field Test 1999 for High School[30]

[30] The New Standards Spring Field Test 1999 for High School is the property of the University of Pittsburgh and the National Council on Education and the Economy (NCEE) and may not be used, reproduced, or distributed without the express written permission of the University of Pittsburgh and NCEE.

Parts A and B of the next item assess Identifying Science Principles and tap into the cognitive demand of "knowing that." Part C of the item involves students' reasoning from a mental model of energy transfer, forces, and motion, thus requiring the Using Science Principles practice and tapping the cognitive demand of "knowing why."

Illustrative Item

Look at the picture of a roller coaster below.



The car on the roller coaster is released from the position shown and allowed to roll freely.

a. Name **two** of the forces that affect the motion of the car while it moves on the roller coaster.
b. Describe how the potential and kinetic energy of the car change as the car rolls downhill.
c. Explain why the car **cannot** reach point **X** on the third hill, as shown in the picture.

(See Appendix D for item scoring guides.)

P8.16 (a), P8.12 (b and c), Identifying Science Principles (a and b), Using Science Principles (c)
Source: Colorado State Assessment Program, 2001, Grade 8 Science

Concept-Mapping Tasks

It is highly recommended that concept-mapping tasks be included on every NAEP Science Assessment at grades 8 and 12 in the form of an interactive computer task (ICT) (see later in this chapter). A concept map (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996a) is a graph in which the nodes represent concept terms, the directed lines that link pairs of nodes represent relationships, and the labels on those lines represent the nature of the relation between the concept pairs (Shavelson & Ruiz-Primo, 1999). Concept maps have been used in science curriculum materials for a number of years (e.g., see Ruiz-Primo & Shavelson, 1996a; Shavelson & Ruiz-Primo, 1999) and can also be used as a reliable and valid assessment of students' ability to make connections among science principles (Ruiz-Primo & Shavelson, 1996a). That is, concept-mapping tasks address the practice of Identifying Science Principles and the cognitive demand of "declarative knowledge," in particular the organization of this knowledge. Ruiz-Primo, Shavelson, et al. (2001) provided statistical and "think aloud" evidence that interpretations of concept maps as reflecting students' knowledge structures were warranted.

In a concept-mapping task, students should be given a set of six to eight concept terms (e.g., "RAIN," "CLOUDS") and be asked to construct a map linking pairs of terms with directed arrows.[31] Students should label each arrow with a word or phrase that explains the relationship between a pair of concept terms (e.g., "comes from"). An arrow-linked pair of concept terms is called a proposition (e.g., "RAIN comes from CLOUDS"). Thus, propositions on a concept map (see below) can be read by selecting a node and reading along the labeled arrow to the next node (e.g., "CLOUDS contain WATER"; "RIVERS contain WATER"; "RIVERS flow to OCEANS").[32]

Item Suggestion

Student Response

An eleven-year old student constructed the following concept map using terms associated with the water cycle:

E8.11, Identifying Science Principles
Source: White and Gunstone (1992, p. 16)

[31] In order to easily distinguish concept terms from linking phrases, concept terms in this section are written using CAPITAL LETTERS.
[32] Since concept maps are representations of structural relations, the logic with which they are constructed may be heavily influenced by relational patterns inherent in an individual's first (or dominant) language. This is of particular importance in the use of concept mapping with English language learners; research is needed on the linguistic demands of this form of assessment on such students.

Although certain types of limited concept-mapping tasks may save time in administration and ease scoring, they do not measure the intended performance expectation (e.g., Ruiz-Primo, Shavelson, et al., 2001). For example, some forms of concept-mapping tasks provide the map structure with a few nodes, arrows, and/or propositions filled in; and the student is asked to complete the balance of the concept map. These forms do not tap the organization of students' "declarative knowledge" (Ruiz-Primo, Shavelson, et al., 2001). NAEP Science concept-mapping tasks should require students to provide the structure of the concept map themselves, and care should be taken to ensure that modifications to concept-mapping tasks do not change the performance expectation being probed.

A concern may arise due to the limited amount of time that students have to construct a concept map. While the length of time required to complete a concept-mapping task may be adjusted by varying the number of concept terms to be used by students, it is highly recommended that concept-mapping tasks utilize a computer-based (ICT) format. Computer programs can ensure that students follow the rules for construction of a concept map; for example, such programs do not allow illegal moves (e.g., unlabeled lines, lines without direction); they prompt for linking terms; and they automatically turn lines into directed arrows. Thus, utilizing the ICT format for concept-mapping tasks would minimize the amount of time required to familiarize students with the task.

The NAEP Science Assessment should show students how to construct a concept map. The directions should be clear; step the student through map construction with exemplars (unrelated to the content domain being assessed); and be sufficient to enable students to construct syntactically correct maps without supervision. The use of ICTs for concept mapping can help guide students by offering reminders about map construction and by not allowing mapping errors. Research (e.g., Ruiz-Primo, Schultz, Li, & Shavelson, 2001) suggests that students as young as 4th graders come to understand the procedure rather rapidly. Concept maps are prime candidates for "cognitive labs" in the item development process (see later in this chapter).[33] In particular, cognitive labs can be used to refine concept-mapping tasks so that these tasks measure students' "declarative knowledge" structures and not their mapping skills.

[33] See concept map research by Ruiz-Primo and colleagues in the Bibliography.

The following textbox provides an illustrative set of instructions for a concept-mapping task.

Clarification: Concept-Mapping Task Instructions

Once they are familiar with how to construct concept maps, students might encounter a set of task instructions that resemble the following:

Examine the concept terms listed below. The terms selected focus on the topic, [insert topic].

Construct a concept map using the terms provided below. Organize the terms in relation to one another in any way you want. Draw an arrow between the terms you think are related. Label the arrow using phrases or only one or two linking words.

You can construct your map on the blank pages attached. When you finish your map, check that (1) all the lines have an arrow; (2) all the arrows have labels; (3) your concept map uses all the terms provided; and (4) your map shows what you know about [insert topic].

List of terms: [insert list of terms]

Source: Adapted from Ruiz-Primo, Shavelson, et al. (2001, p. 107)

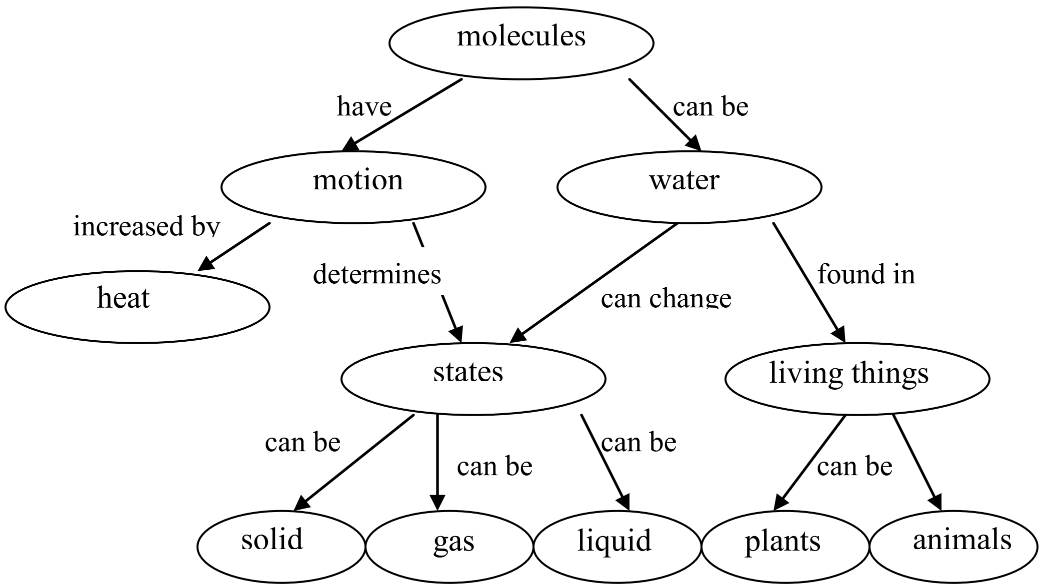
Students' concept maps are typically scored as to the accuracy of propositions in their maps (Ruiz-Primo & Shavelson, 1996). Yin, Vanides, Ruiz-Primo, Ayala, and Shavelson (2005) review the various ways concept maps have been scored, and this resource should be consulted for scoring concept maps in the NAEP Science Assessment. In brief, they defined a total proposition accuracy score as the sum of the individual proposition scores in a student's map. Each proposition was scored on a four point scale: 0—wrong or scientifically irrelevant; 1—partially incorrect; 2—correct but scientifically "thin"; 3—scientifically correct and scientifically stated. They implemented the scoring system in an Excel database.

In the following concept map, the proposition, "WATER found in LIVING THINGS" would receive a score of 2—it is correct but scientifically thin. If the proposition stated, "WATER is a component of all LIVING THINGS," it would receive a score of 3—scientifically correct and scientifically stated. On the other hand, if the proposition stated, "WATER sometimes found in LIVING THINGS," it would receive a score of 1—partially incorrect—because water is always (not sometimes) found in living things. Finally, if the proposition stated, "WATER not found in LIVING THINGS," it would receive a score of 0—wrong. Absent propositions are not scored (e.g., the absence of a linking line between "WATER" and "LIVING THINGS" or between "HEAT" and "STATES" would receive 0 score points by default). Once each proposition in the concept map is scored, the sum of all individual proposition scores would comprise the aggregate score for the entire concept map. This aggregate score could then be standardized to NAEP scale scores.

Item Suggestion

Student Response

A student constructed the following hierarchical concept map:



P8.6, Identifying Science Principles
Source: Adapted from Novak and Gowin (1984, p. 18)[34]

If concept-mapping tasks are administered as ICTs, the computer could create a database of students' propositions for each pair of concept terms. This database would consist of all of the labels used by students on the linking arrows between each valid pair of concept terms. (Some concept terms do not pair together and so would not be expected to have a labeled arrow connecting them.) Human raters would judge the quality of each proposition for a linked pair of concept terms and enter this score into the database. Once a score has been assigned to each proposition, the computer can score subsequent maps automatically. If a new proposition appears that the computer does not recognize, the proposition would be submitted to human raters before the computer continues with the scoring process for that particular student's concept map. Checks on computer scoring and reliability analyses should be performed throughout the scoring process. Even if concept-mapping tasks are administered using a paper-and-pencil format, a computer database is still recommended, with the only difference being that each proposition created by a student would need to be inputted manually.

A final concern that may arise about concept-mapping assessment tasks is that teachers may teach to them directly, having students memorize certain widely used concept maps

[34] Reprinted with the permission of Cambridge University Press.

in instruction. However, it becomes clear that, unless the exact same set of concept terms appears on the NAEP Science Assessment as those on which students have been coached, quite different concept maps might be expected. Anticipating this concern, Ruiz-Primo, Schultz, et al. (2001) examined the equivalence of concept maps in which the terms were randomly sampled from a universe of terms. They found that concept maps constructed from two randomly generated sets of concept terms provided equivalent information about students' knowledge structures. Consequently, concept-mapping tasks on NAEP could be generated in such a way that the chances of "teaching to the test" would be very slim.

Combination Items

Combination items include item clusters, Predict-Observe-Explain item sets, hands-on performance tasks, and interactive computer tasks. Combination items may consist of an all selected-response format, an all constructed-response format, or a mixture.

Item Clusters

Item clusters should be included on every NAEP Science Assessment at grades 4, 8, and 12. Item clusters are groups of related items that provide for more in-depth analysis of student performance than would a collection of discrete, unrelated items. Considerable research has been done on the use of clusters to measure the depth of student knowledge within a particular content area, that is, learning progressions. Most of the following discussion is centered on the use of item clusters for this purpose. However, clusters may also be designed to measure the breadth of student understanding of specific concepts, principles, or procedures across content areas, that is, crosscutting content.

Item clusters may cross content areas, vary in the knowledge tapped, and vary across grade levels. In particular, item clusters that probe students' "mental models," naïve conceptions, predictions, or explanations of the natural world are good candidates for the NAEP Science Assessment. Their development should be guided by current research on different forms of these items. In this type of item set, two or more items focus on an important idea or "mental model." Hence, these items tap the practice of Using Science Principles and the cognitive demand of "knowing why." Where there is a rigorous body of research available on students' conceptions (as there is about the solar system), item clusters provide opportunities to assess students' understanding of a particular key science principle at some depth. This type of item set can probe the conceptions and "mental models" that underlie students' explanations of and reasoning about the natural world. Since there is more than one item probing the same conception, guidelines in addition to those for individual multiple-choice items are set forth here:

- All options should reflect alternative conceptions that are plausible to students based, wherever possible, on the research literature.[35] With respect to floating and sinking of objects in water, alternatives could include the relative density of water and the object (correct), the mass of the object, the size of the object, and other similar options.
- The correct option should reflect the most scientifically justifiable alternative conception.
- Because there is a series of related items, care should be taken to ensure that the answer to one of the items is not contained in another item in the cluster and that the answer to one item does not depend on an answer to a prior item.
- As some items will be dropped based on review and field test, almost twice the number of items should be written for a cluster as would be used operationally in the cluster.
- Each item should be tagged as to the cluster to which it belongs and, for multiple-choice items, each distractor should be tagged as to the particular conception (e.g., "mental model") it represents.
- Where feasible, an item cluster should include an "orienting" item, typically one that asks students to *identify* the science principle ("declarative knowledge" and reasoning) that is the focus of the item cluster with subsequent items focused on, perhaps, *using* science principles. In this way, students are cued to the conceptual domain of focus in the cluster.

Following are examples of two different approaches to item clusters.

Sadler (1998) built a 47-item assessment of people's "astronomical ideas." Small clusters of items on the assessment tap important ideas in astronomy such as why there is night and day on Earth and why there is a change of seasons in the continental U.S. These item clusters are built considering the non-linear development of students' conceptions about the natural world that evolve as students gain understanding. Each multiple-choice item on Sadler's assessment contains an alternative that constitutes the scientifically accepted response; the remaining distractors contain naïve conceptions that people are known to hold. The following were part of a set of items probing high school students' mental models (percentages of student responses to each option are given in parentheses):

[35] As a start, assessment developers may wish to consult the following resources, which document research on students' science conceptions: AAAS (1993, chapter 15); Driver, Guesne, and Tiberghien (1985); Driver, Squires, Rushworth, and Wood-Robinson (1994); Duit (2004); Gentner and Stevens (1983); Hewson, Beeth, and Thorley (1998); NRC (1999c, 2001); Osborne and Freyberg (1985). The FACET Innovations website lists additional resources on students' physics conceptions; see http://www.facetinnovations.com/main/diag-refs.htm (retrieved February 20, 2006).

Illustrative Items

What causes day and night?

A. The earth spins on its axis. (66%)
B. The earth moves around the sun. (26%)
C. Clouds block out the sun's light. (0%)
D. The earth moves into and out of the sun's shadow. (3%)
E. The sun goes around the earth. (4%)

Key: A

E8.2, Using Science Principles

The main reason for its being hotter in summer than in winter is:

A. The earth's distance from the sun changes. (45%)
B. The sun is higher in the sky. (12%)
C. The distance between the northern hemisphere and the sun changes. (36%)
D. Ocean currents carry warm water north. (3%)
E. An increase occurs in "greenhouse" gases. (3%)

Key: B

E8.12, Using Science Principles
Source: Sadler (1998, pp. 274-276)

It is possible to link a particular alternative to students' general knowledge in astronomy. Sadler showed that as general science ability (as measured by an astronomy test's total score) increases, so does the probability that students will select "earth spins on its axis" for the cause of day and night. (Ironically, in some cases, as students learn more, they may develop new naïve conceptions as they progress to fuller understanding.) Moreover, with several items clustered around an astronomical idea such as change of seasons, it is possible to characterize the percentage of students holding justified or alternative conceptions.

Another approach to probing students' conceptions of the natural world is to develop a cluster of ordered multiple-choice items (Briggs, Alonzo, Schwab, & Wilson, 2006). These items track students' performance along a learning progression from naïve understandings through more reasoned naïve conceptions to full and scientifically justified understandings. In this approach, the progression is first described and then divided into levels so that multiple-choice items can be designed specifically to assess the performance level that a student (or group of students) has reached. (See Appendix I for additional information.)

As noted earlier, constructed-response formats are not precluded for use in item clusters (see Appendix I for an example). Although items in a cluster are related to one another by virtue of their tapping similar performance expectations, they should be scored individually. In other words, individual items in a cluster are scored separately, but these scores could be added together or response patterns could be analyzed for research purposes. See p. 219 for more on reporting NAEP results.

Predict-Observe-Explain (POE) Item Sets

POE item sets (White & Gunstone, 1992) should be included on every NAEP Science Assessment at grades 4, 8, and 12. These types of items ask the student to predict, observe, and/or explain as follows: A situation is described, and the student's task is to provide a prediction for what will happen (sometimes with justification), and/or to provide an explanation for what appears to be an anomaly. POE items tend to tap the practice of Using Science Principles and the cognitive demand of "knowing why (schematic knowledge)."

As noted earlier, a POE item can take either a selected-response or a constructed-response format. In the selected-response format, students choose from a set of possible alternatives (based on known alternative mental models). In the constructed-response format, the student's task is to write out (with justification) a prediction or an explanation.

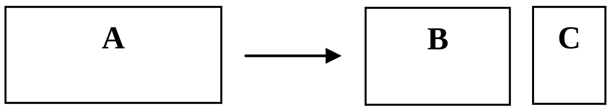
A POE selected-response item set that includes a justification takes advantage of the multiple-choice format and yet preserves information inherent in students' justifications. An example of such an item set focusing on prediction based on a mental model of buoyancy is provided below. In the first item, used at the middle school level, students are asked to *choose among a set of predictions* for whether the 1/3 and 2/3 parts of the block will sink or float. They can choose an answer from several alternatives that include naïve conceptions among them. In the second item, students justify their answers by selecting from a number of possible explanations. Again, these alternatives include the correct explanation along with distractor explanations based on naïve conceptions.[36]

---

[36] Valuable information on the prevalence of certain naïve conceptions can be provided by secondary analyses of which distractors students choose.

Illustrative Item[37]

Rich cut block A into two unequal parts. Part B is 2/3 of the original block A, and part C is 1/3 of the original block A. Block A sinks in water. What will happen to B and C when placed in water?



A. Both B and C will float.
B. B will sink; C will float.
C. B will subsurface float; C will float.
D. Both B and C will sink.

Key: D

P8.4, Using Science Principles
Source: Adapted from Shavelson (2006) and Shavelson et al. (2005)

Item Suggestion

Which of the following best describes the reason for your answer to the preceding question?

A. The mass of B is less than the mass of A, but more than C.
B. The density of B and C is the same as the density of A.
C. The volume of B is less than the volume of A, but more than C.
D. The density of C is less than the density of A and B.

Key: B

P8.4, Using Science Principles

Instead of using the multiple-choice format, students could be told that the full block sinks in water and be asked to *write their prediction* (with justification) as to what will happen when the two parts are placed in water. Or, they could *observe* a simulation or video of what happens to the full block and the two parts and then be asked to *explain* what they have just observed.

Although interdependence of items is to be generally avoided (see p. 154), non-independence is expected for items asking students to explain or justify their response to a preceding item. The sequential responses may be integrated into the scoring system. In some cases, a score point may be awarded for a correct evidence-based justification in the second item, even if the justification is for an incorrect response to the first item. For example, if each of two items is worth 1 score point, it is possible for a student to receive 0, 1, or 2 total points for the two items combined—0 points for incorrect responses to

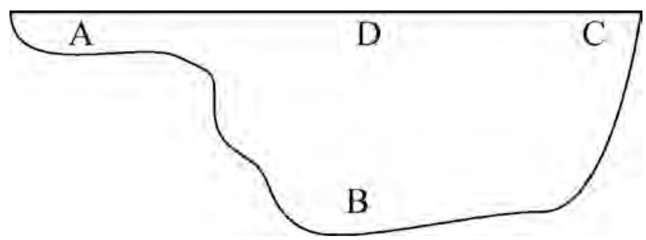
[37] This item assumes that the entire block consists of a completely homogeneous material. The possibility that the block is made of heterogeneous material is unlikely to occur to middle school students.

both items, 1 point for a single correct response, 2 points for correct responses to both items.

Expectations regarding predictions with justifications may be different at different grade levels. For example, the following grade 12 item requires students to make a correct prediction *and* supply at least a partially correct explanation (justification) to receive any score points. Students receive a score of zero if the prediction is incorrect or if the explanation is incorrect or missing.

Illustrative Item

Here is a cross-section of a lake in the mountains. The air temperature gets below freezing in the winter and stays below freezing for 3 months.



Not all of the water in the lake freezes. Which part of the lake will remain the warmest? Explain.

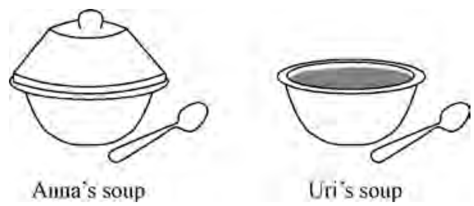
(See Appendix D for item scoring guides.)

P12.1, Using Science Principles
Source: TIMSS 1995, Grade 12

By contrast, the following grade 4 item requires only that students make the correct prediction to receive one score point. An additional score point may be earned if the explanation (justification) is also correct.

Illustrative Item

Anna and Uri had identical bowls of soup, both at the same temperature. Anna put a cover on her bowl.



Whose soup do you think would stay hot longer? __________________

Give a reason for your answer.

(See Appendix D for item scoring guides.)

P4.2, Using Science Principles
Source: TIMSS 1995, Grade 4

Hands-on Performance Tasks

Hands-on performance tasks will be administered to a subset of the students sampled for every NAEP Science Assessment at grades 4, 8, and 12. In hands-on performance tasks, students manipulate selected physical objects and try to solve a scientific problem involving the objects. These exercises, if carefully designed, can probe students' abilities to combine their science knowledge with the investigative skills reflective of the nature of science and inquiry. In large-scale assessments such as NAEP, uniform administration must be ensured. In the past, this has been accomplished through the use of standardized performance assessment kits, with each exercise proctored and scored by trained personnel. Special accommodations may be necessary for some students.

A particularly cogent criticism of most hands-on performance tasks administered in large-scale assessments is that, rather than tap into students' ability to inquire into a problem, typical performance assessments instead measure students' ability to follow step-by-step instructions to arrive at the expected answer. Assessment developers are likely to create these recipe-types of exercises since they need to take into account the vast differences in students' science courses and experiences. Given these differences, the absence of structure might produce unanticipated responses that might be problematic for the assessment either at the time the data are collected or when students' performances are scored by raters. Although both the concern for structure to make large-scale assessment manageable and the criticism of highly-structured performance tasks are well taken, there is evidence that valid performance exercises can be designed, developed, administered, and scored without encountering major problems (e.g., Ruiz-Primo & Shavelson, 1996b; Shavelson, Baxter, & Pine, 1991; Solano-Flores & Shavelson, 1997).

In designing hands-on performance tasks, the following should be kept in mind. The degree to which students engage in some aspect of scientific inquiry depends upon who selects the problem to be studied, who selects the procedures to be carried out in tackling the problem, and who selects the answer. In NAEP, the assessment should provide students with a challenging problem. However, students must be given the opportunity to determine scientifically justifiable procedures for addressing the problem and arriving at a solution. Indeed, the problem to be solved is in setting forth procedures that manipulate the variable of interest, control extraneous variables, and provide solid data to be used in arguing for and justifying a problem solution. In addition to allowing students to determine the procedures for carrying out the experiment, NAEP hands-on performance tasks should be “content rich,” in that they require knowledge of science principles to carry them out.

The basic elements of a performance exercise are the challenge, the response, and the scoring system:

- <u>The challenge:</u> The task should pose a problem requiring concrete materials that react to the actions (one or more operations) taken by students in carrying out the task. At grade 4, that operation may be as simple as taking the temperature of a container of water or weighing an object and recording the measurement. It can also be a problem to solve (e.g., “Given the objects before you, make the bulb light”; “How many ways can you make the bulb light?”). At grade 8, the challenge can be more complicated. Students can be asked to link several operations together, make a prediction, and/or create a data table, which they then can interpret. For example, students might be asked to put a substance that they carefully measure in water, causing an exothermic reaction; they could then be told to add additional measured amounts of the material to the water and record the temperature after each addition; finally, they could be asked to interpret or make some extrapolations based on their data. At grade 12, students may be asked to solve an open-ended problem, given certain materials. For example, students could be told to plan a procedure to test how several blue liquids provided react with one another; follow the procedure and record their answers; and, from the results obtained, provide as much information as possible relative to the nature and identity of each liquid.
- <u>The response:</u> Usually, more than one type of student response is required in such tasks, and the required responses may vary in format. For example, there could be a series of multiple-choice items that measure content knowledge or other aspects of the problem situation; there could be several short and one longer written response to portions of the items in the hands-on task (e.g., summary of conclusions, descriptions or lists of steps in a procedure, diagrams, calculations, graphs, tables). The required response format should be based on the challenge. Previous caveats apply with respect to interdependency.

- The scoring system: The scoring system should tap the performance expectation (construct) for which the task was selected. It should also capture the "right" answer(s), the justifiability of the procedures and evidence used to arrive at the answer(s), and some aspects of the students' problem solving procedures and errors (Ruiz-Primo & Shavelson, 1996b). The scoring system should capture the wide variety of performances expected of students. Typically, performance tasks are scored on a 4- or 6-point scale, which combines an evaluation of student performance on several attributes (e.g., manipulation of an independent variable, control of extraneous variables, appropriate measurements, accuracy of conclusions).

The challenge should require students to be engaged in solving a problem rather than simply performing a procedure for no other reason (e.g., making a measurement, weighing an object). While activities such as making measurements can be exciting for 4th graders, they are boring for older students. Also, procedural types of activities tend to assess whether students have been exposed to them or can carry them out, rather than what they know and can do in using and applying science principles.

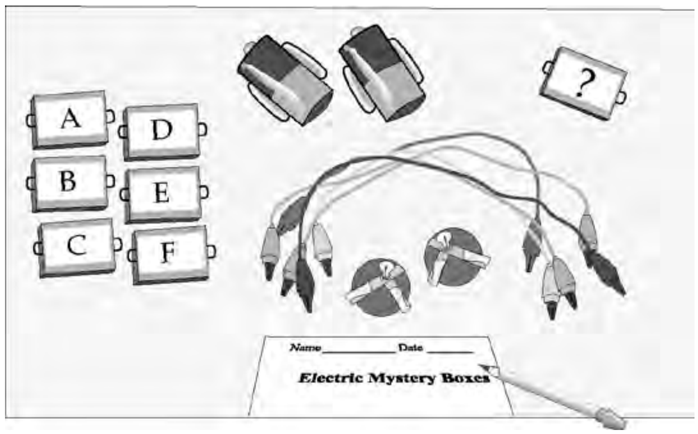
The hands-on performance task protocol will require the use of standardized kits. Each task should be proctored and scored by highly trained personnel. The assessment development contractor will develop appropriate protocols for administering hands-on performance tasks, including the physical layout, sequencing, and timing (see Solano-Flores & Shavelson, 1997).

Hands-on performance tasks require the use of "laboratory" equipment and materials, such as balance scales, rulers, beakers, and material objects. These materials should be supplied to students when they receive the block containing these tasks. Developers should design the assessments so that tasks requiring the use of specific equipment or materials are contained in the same block. Equipment and materials must be chosen with student safety as a critical criterion (e.g., Bunsen burners would not be appropriate for wide distribution and use).

In brief, any hands-on performance task included in the NAEP assessment should present students with a concrete, well-contextualized task (problem, challenge) along with "laboratory" equipment and materials, and a response format that leaves the exercise process open. Students' scores should be based on both the procedures created for carrying out the investigation and the solution (Shavelson et al., 1991). The assessment, then, should provide the problem that draws on science principles and practices (performance expectation of interest) and leave students free to design and carry out the exercise to arrive at an answer or solution. The following items, designed for 5th graders, are examples of such tasks.

Illustrative Item

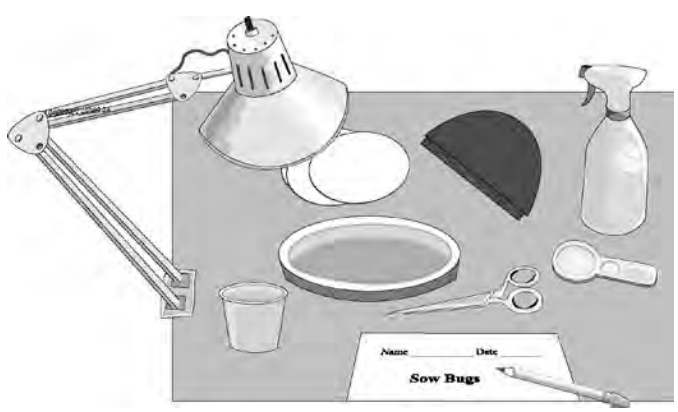
Students are asked to identify the contents of each of the six boxes (A-F) by using the batteries, bulbs, and wires they are given to complete a circuit. This task requires knowledge of series circuits but leaves problem-solving procedures up to the student. (See Appendix D for further description of this task.)



P4.11, Using Scientific Inquiry
Source: Shavelson et al. (1991)

Illustrative Item

Students are provided with materials including a plastic dish, blotter paper, black filter paper, scissors, a lamp, a spray bottle, a magnifying glass, and 5-6 non-rolling sow bugs (see below). Students are asked to conduct a series of experiments to determine sow bug behavior when given choices of the following environments: (1) light or dark, (2) damp or dry, and (3) combination of the two.



(See Appendix D for a sample scoring form.)

L4.2, Using Scientific Inquiry
Source: Shavelson et al. (1991); Shavelson, Gao, and Baxter (1995)

At each of grades 4, 8, and 12, students should be permitted 30 minutes to complete their hands-on performance tasks.

There is a risk that some students will not complete process-open hands-on performance tasks. Because NAEP aims to provide opportunities for all students to perform, it may be necessary to create a "branch" for students who require further prompts to carry out these tasks. For example, a branch might include "recipe-like" instructions for students who remain at a standstill after approximately 30 percent of the allotted task time has passed. Specific directions on how to determine when to use branching and record this for scoring must be provided to assessment administrators. Responses from students who use such branches should be scored on the lower end of the item scoring scale; it may be possible to differentiate among student responses within this lower end.

In summary, the following list (taken from a more complete discussion offered by Solano-Flores and Shavelson, 1997, p. 21) notes some of the important issues in the form of questions that assessment developers need to consider in designing hands-on performance tasks:

- Is the task amenable to many solutions varying in correctness?
- What actions should students be explicitly asked to report?
- Does the response format elicit the intended type of response?
- Do students actually understand the problem?
- Is student diversity (e.g., gender, ethnicity, SES) being considered?
- Does the problem promote the students' active participation?
- Does the equipment react consistently to the students' manipulations?
- Can the students use the equipment with ease? Are they familiar with it?
- What are the equipment's packaging, storage, transportation, and handling requirements?
- What are the classroom physical conditions (e.g., illumination, water availability) needed to properly administer the assessment?
- What activities (e.g., hand-out material, watch students) must be carried out to properly administer the assessment?
- What procedure should scorers use to review the students' responses?

Interactive Computer Tasks

Interactive computer tasks (ICTs) will be administered to a subset of the students sampled for every NAEP Science Assessment at grades 4, 8, and 12. The 2009 NAEP Science Assessment should include some but not necessarily all of the following four types of ICTs: (1) information search and analysis, (2) empirical investigation, (3) simulation, and (4) concept mapping. The broad purpose of ICTs in the NAEP Science Assessment is to tap performance expectations that are more advantageously assessed in this format. ICTs are intended as a complement to the hands-on performance tasks, not as a replacement.

At each of grades 4, 8, and 12, students should be permitted 30 minutes to complete their ICTs.

Computers and other media provide potential solutions to a variety of practical challenges posed by complex assessment exercises. The messiness and logistical challenges of hands-on performance tasks can be circumvented with computer simulation (e.g., Pine, Baxter, & Shavelson, 1993). Extensive databases can be presented to assess students' ability to select and evaluate information relevant to the situation or problem they are asked to address (Persky, Bennett, Weiss, & Jenkins, 2005; Quellmalz et al., 2004). Moreover, the difficulty of providing materials and training for complex tasks such as concept maps can, as well, be circumvented with computers (Herl, O'Neil, Chung, & Schacter, 1999; Yin et al., 2005). To avoid cheating and teaching to the concept map, concept terms can be randomly sampled for a particular map (Ruiz-Primo, Schultz, et al., 2001). Finally, student response databases and natural language processors can be used to score complex performance.

ICTs should be used where the format offers advantages over other assessment modes. To summarize, these include, but are not necessarily limited to, assessing student knowledge, skills, and abilities related to the following situations:

- 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 speeded-up (e.g., erosion caused by a river). It is also useful when it is necessary to freeze action or replay it.
- For modeling scientific phenomena that are invisible to the naked eye (e.g., the movement of molecules in a gas).
- For working safely in lab-like simulations that would otherwise be hazardous (e.g., using dangerous chemicals) or messy in an assessment situation.
- For situations that require several repetitions of an experiment in limited assessment time, while varying 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).
- For searching the Internet and resource documents that provide high-fidelity situations related to the actual world in which such performances are likely to be observed.
- For manipulating objects in a facile manner such as moving concept terms in a concept map.

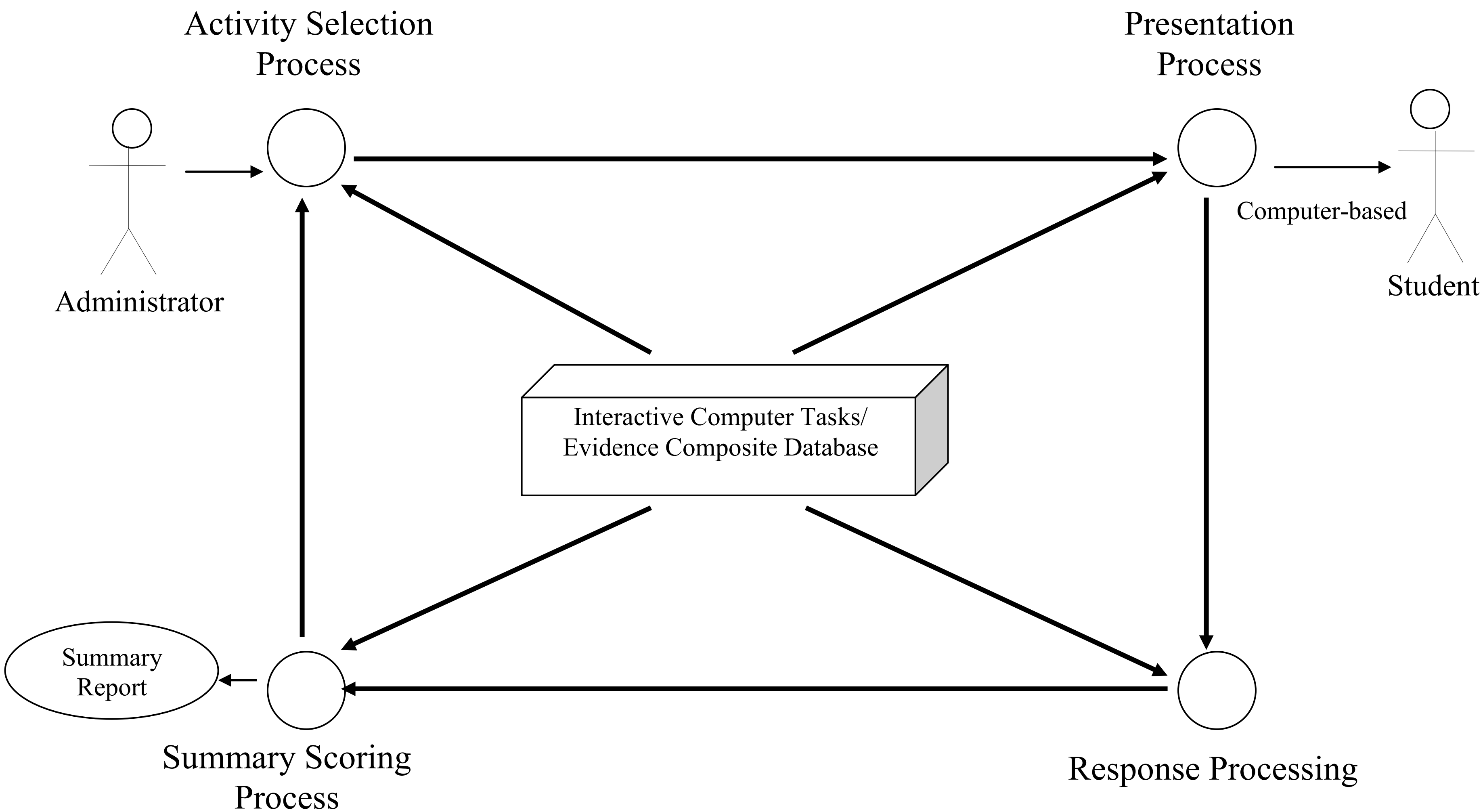
Extended constructed-response items, concept-mapping tasks, and simulated performance tasks are especially strong candidates for interactive computer tasks. In this way, the complex science understandings and practices that need to be probed in the NAEP Science Assessment might very well be captured with less time, cost, and logistical challenges and with greater opportunity for divergent problem-solving tasks than has been the case in the past.

Specifications of the ICT Assessment System

The following set of specifications follows the *Four-Process Architecture* proposed by Almond, Steinberg, and Mislevy (2003), which provides a very useful framework for describing, at a general level of specificity, the components of complex, computer-based assessments.

Figure 3 shows a diagrammatic representation of the assessment delivery system that could be used to administer the ICTs.

Figure 3. The Four Processes in the ICT Assessment Cycle



Adapted from Almond et al. (2003)

In Figure 3, the two key human players are the administrator and the student. The administrator in the case of the NAEP Science Assessment is the contractor charged by NCES with the task of setting up and maintaining the assessment. The administrator is the starting point for the assessment process and determines the choices made in the assessment system. The other players in the system are the students who take the NAEP Science Assessment at grades 4, 8, and 12. They interact with the various ICTs that are presented to them via the Presentation Process, which is later discussed in detail.

The following sections give specifications relevant to each of the four processes of the assessment cycle shown in Figure 3.

<u>Activity Selection Process</u>

The Activity Selection Process is the process responsible for the selection and sequencing of tasks from the ICT database represented in the center of Figure 3. The ICT database may comprise four types of ICTs, all using different kinds of simulations:

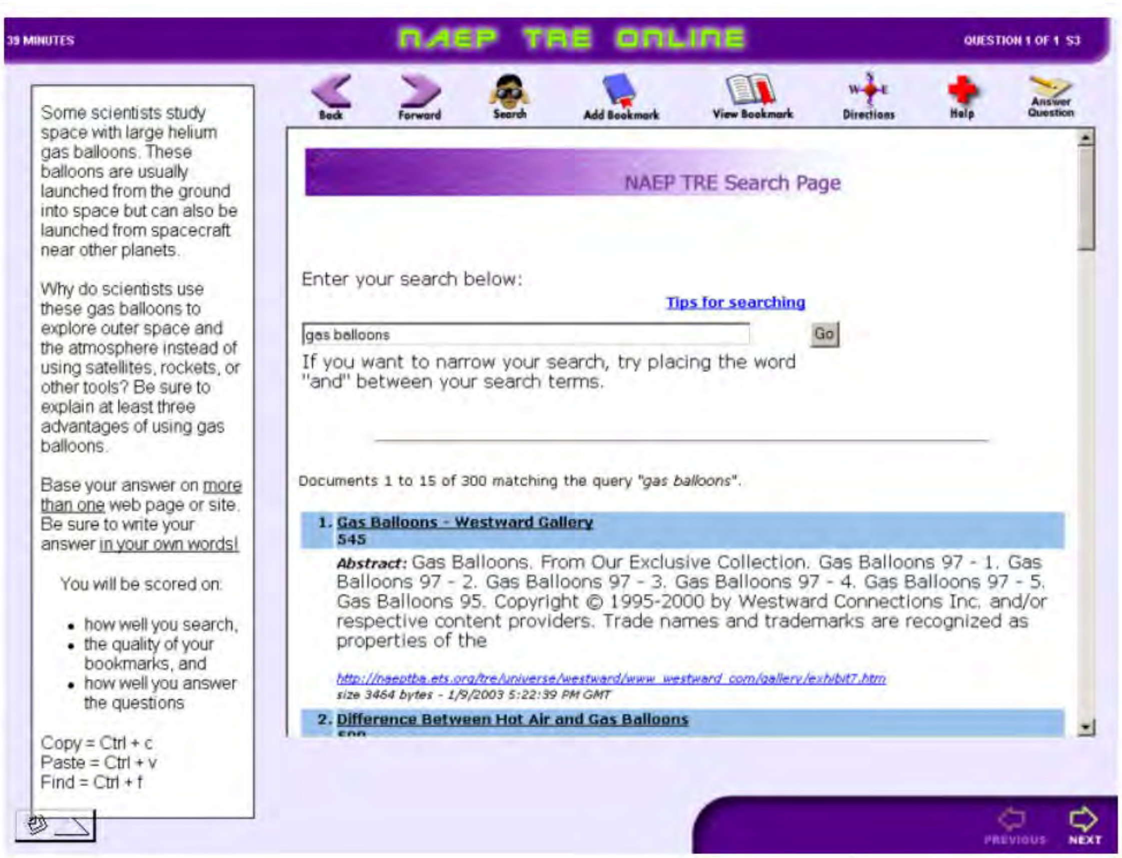
1. <u>Information search and analysis.</u> A student is given a problem and then uses a simulated Internet environment—one that has been constructed specifically for this task— to search out information (scientific data, theory, social issue, etc.) and bring it to bear to solve the problem.
2. <u>Empirical investigation.</u> In a virtual environment, students design and conduct an experiment in which they manipulate equipment, record results, and report on their analysis and conclusions in a notebook. Responses in this format require students to record and analyze data, which necessitates access to a spreadsheet software program or similar custom-designed software tool. Students also answer selected-response questions and type in their answers to constructed response questions.
3. <u>Simulation of phenomena and models.</u> Students interact with a simulation of a natural phenomenon, such as the behavior of gases, or with models of real-world systems like a food cycle. Then, questions are posed to students, and they manipulate the simulation to gather data and solve given problems. Responses require the same sort of tools described for simulated hands-on performance tasks.
4. <u>Concept mapping</u>. To build a concept map, students spatially manipulate concept terms provided to them and explain the relationship between pairs of concept terms by typing in labels for the directed lines.

Following are examples of these four types of ICTs.[38] Static screen shots are used throughout this document to illustrate examples of ICTs. Note that these screen shots represent only a small subset of the many screens students see when engaged in actual ICTs.

[38] In addition to the following examples, see www.colorado.edu/physics/phet/ for online simulations created by the University of Colorado's Physics Education Technology project (PhET); Perkins and Wieman (2005) provide an overview of these simulations. Although the PhET simulations are not designed as assessment tasks, they clearly illustrate the interactive possibilities of ICTs.

Information search and analysis items pose a scientific problem and ask students to query an information database to bring conceptual and empirical information to bear, through analysis, on the problem. The following is a screen shot from such an ICT developed for 8th graders.

Illustrative Item

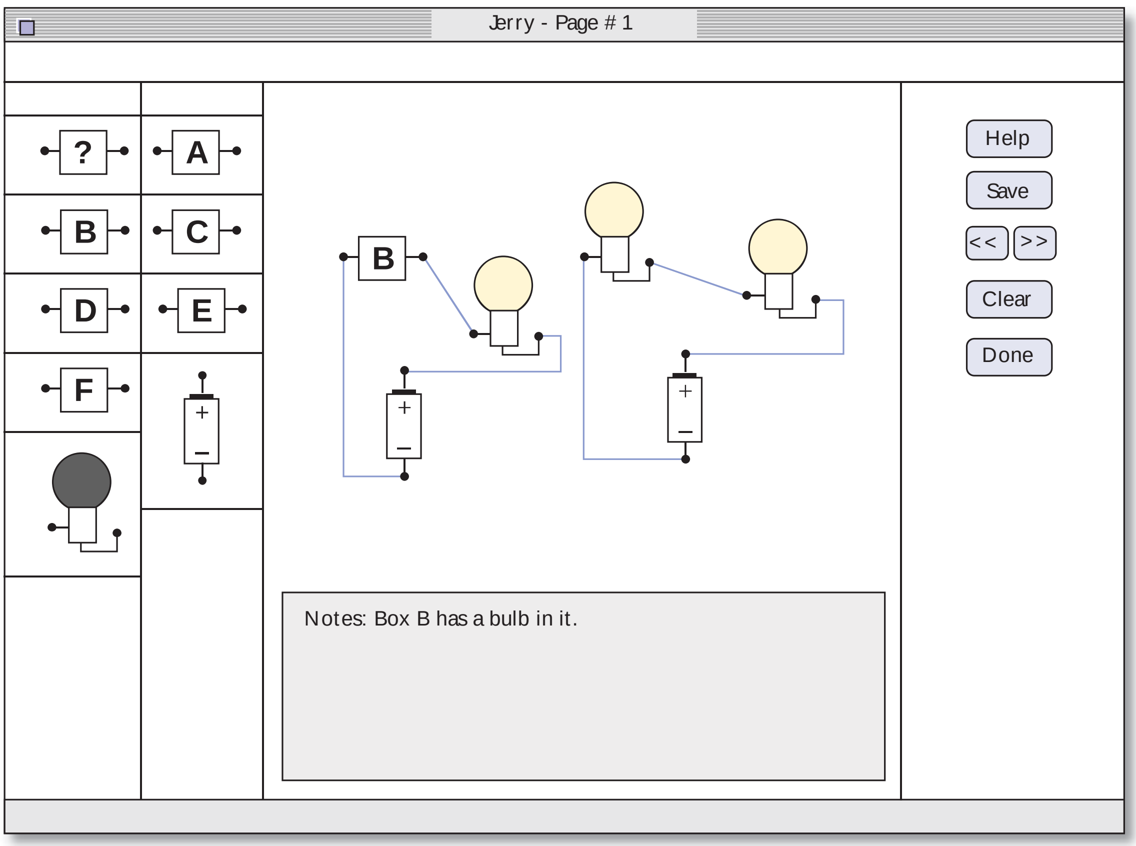


(See Appendix D for further description of this task.)

P8.16, Using Scientific Inquiry
Source: Persky et al. (2005)

Empirical investigation items put hands-on performance tasks on the computer and invite students to design and conduct a study to draw inferences and conclusions about a problem. Whether the computer simulated experiment assesses the same skills, knowledge, and understandings as hands-on performance tasks has not been established (but see Rosenquist, Shavelson, & Ruiz-Primo, 2000), and a special study is proposed to address this question (see p. 222). The following is a screen shot from a computer version of the Electric Mysteries task (see p. 182 and Appendix D for further description of this as a hands-on performance task).

Illustrative Item



P4.11, Using Scientific Inquiry
Source: Shavelson et al. (1991, p. 357)

Simulation items model systems (e.g., food webs), pose problems of prediction and explanation about changes in the system, and permit students to collect data and solve problems in the system. The following screen shot comes from such a simulation ICT.

Illustrative Item

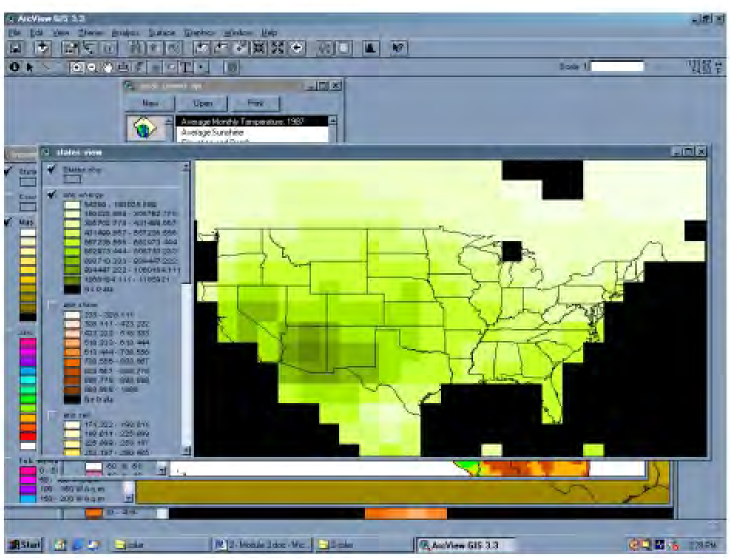
High school (grade 12) students are asked to identify locations appropriate for solar power generation. (See Appendix D for description of the full task.) In order to complete the task, they must:

- Evaluate GIS map visualizations.
- Compare and contrast visualizations of different types of data.
- Use analytical extension to perform computations with visualization data.

Provided below is an example of one of several tasks that a student completes:

Your task is to identify two states that will have a high annual solar energy and will be able to generate the maximum amount of electricity from their solar panels. Name two states that you predict will have a good annual electrical yield. In the rest of this performance assessment, you will generate visualizations and calculate which states will generate the best annual electricity yield from solar panels.

The following screen shot provides an example of a student response:



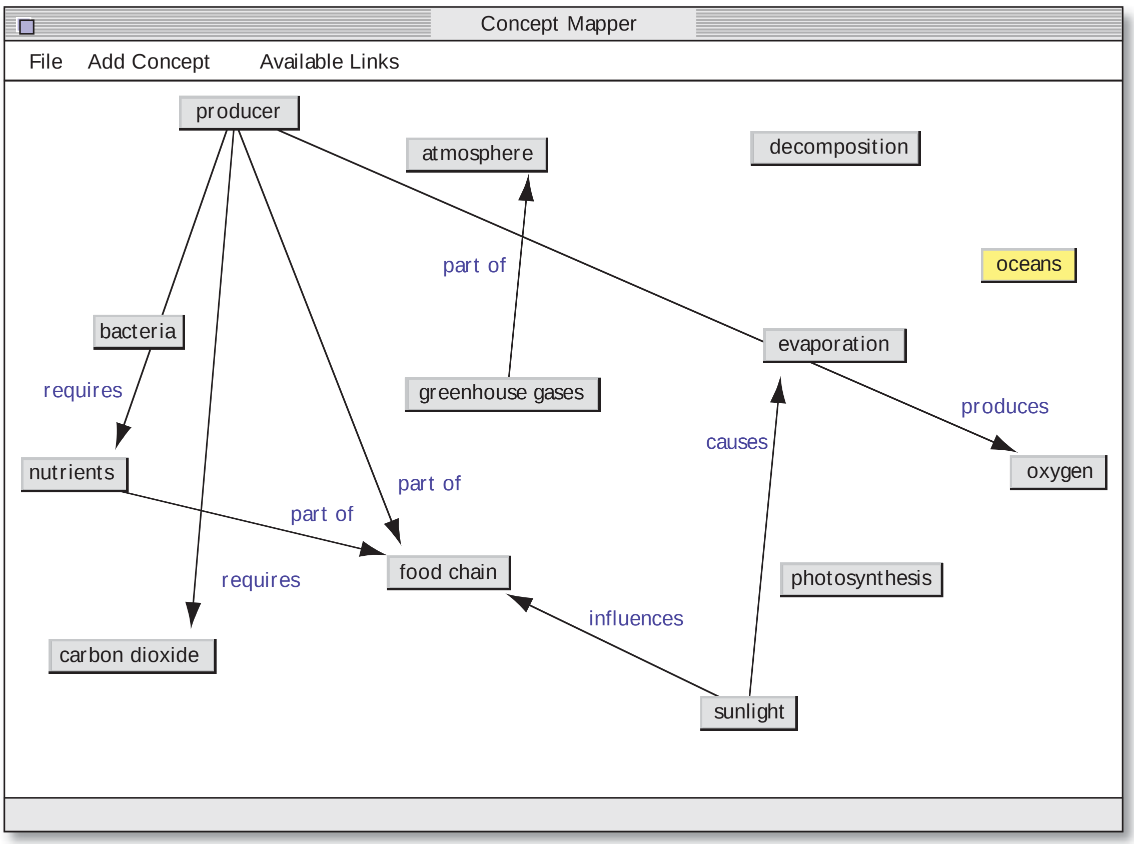
This visualization shows that the states with the most annual solar energy are Arizona and Nevada. This data alone, however, does not completely justify the location of the panels, due to temperature issues.

E12.10, Using Technological Design
Source: Quellmalz et al. (2004)

Finally, concept mapping can be done by providing concept terms and asking students to build propositions on the computer by linking pairs of terms with arrows and words or phrases. The following is a screen shot of a completed concept map.

Illustrative Item

In this task, middle and high school students used a customized software program to create concept maps. Students were given eighteen environmental science terms, as well as seven link labels.[39] Students could drag and drop these concepts onto the grid space of the mapping program and add, erase, and link the items in their newly constructed maps. (See Appendix D for further description of this task.)



L12.5, Identifying Science Principles
Source: Adapted from Herl et al. (1999)

[39] Link labels should not be provided to students on the 2009 NAEP Science Assessment. See the section on "Concept-Mapping Tasks" earlier in this chapter.

Presentation Process

The presentation process is responsible for presenting the ICTs to the student via computer connected to the Internet. The presentation must be accessible from different hardware platforms (PC and Mac) using a range of common web-browsing software (e.g., Microsoft Internet Explorer, Safari, Mozilla Firefox). Any software plug-ins needed to run the presentation of the assessment must be freely available to schools. The presentation process should run smoothly at download speeds typical of common broadband Internet access in schools. Access to the presentation process will be secured by usernames and passwords. All data transactions should be secure.

To maximize the use of ICTs in the database, it is envisioned that components of the ICT items will be stored in separate data collections. For instance, separate components might include—but are not limited to—the following:

- The simulation scenario
- Variable objects in the simulations
- Values of objects in the simulations
- Assessment tasks
- Response tools

The selection process (controlled by the administrator) will determine the appropriate components for a given ICT, and in the presentation process, these components will be presented to the student with which to interact.

For example, an ICT might involve a hands-on laboratory simulation of a vehicle (whose mass can be varied by adding weights) that is allowed to run down a slope under the force of gravity where it hits another stationary vehicle. The task for the student is to run several trials using the vehicle with different weights attached, record the results in a data table, plot the results on a graph, and answer questions about the energy transferred from the moving vehicles to the stationary one. The assessment selection process would select a simulation scenario (e.g., inclined plane, slope). The angle of slope would be obtained from the values table. Also, the other variable objects in the simulation would be presented in the scenario, one moving vehicle and the other stationary one, together with the weights that can be attached to the moving vehicle. The masses of the two vehicles would be determined from the values tables, as would the masses of the weights that are to be added to the moving vehicle. The assessment tasks would be selected and presented, including instructions for the task and questions. Response tools appropriate to the task, a balance scale for measuring mass of the vehicles and weights, a table for data entry, a graphing tool, and a text entry box would be chosen in the selection process and presented to the student as part of the ICT. The selection processes and presentation processes combine to make a seamless experience for the student. In this way, such a task can be varied over time. For example, the slope can change as can the mass of the cars and weights, and even gravity could be designated as a variable that changes in the simulation. What remain constant are the procedures in the investigation, collection of data, and interpretation of results.

Total interaction time for each ICT will be 30 minutes. As students interact with the ICTs, their responses will be captured and sent as “work products” to the response processing part of the assessment cycle. Response processing is described in the following section.

<u>Response Processing</u>

The actions and inputs that students make in response to the ICTs in the presentation process are gathered as work products particular to the different elements of each ICT. These work products are, at this stage, just raw data that must be processed before it can be scored in the next stage of the assessment cycle. The response processing identifies the essential features of the response that provide evidence about the student’s current knowledge, skills, and abilities.

In the ICT assessment system, there will be several different work products generated. First, the student’s interaction with the simulation will produce student actions that relate to the skills to be measured. For example, in the Internet search simulation, the specificity of the search terms used by a student, resources accessed when a website is visited, and the number of sources tapped all provide measurable responses. Similarly, in the simulations of laboratory experiments, actions of interest would include the equipment selected to conduct the experiment, how the equipment is connected and used, or the number of trials run to gather data. In the simulations of scientific phenomena, measurable parts of the interaction might be the manipulation of variables, whether multiple trials are run to investigate a problem, and sampling across the range of potential values of the variables.

Other work products will come from the questions that students answer during or after their interaction with the simulation. These questions may be selected-response questions or constructed-response questions.

Processing for each work product response type will vary. For example, if measuring, then the number of times a student conducts a trial in an experiment will have to be interpreted. Certainly, conducting only one trial would produce the lowest score, but conducting 20 trials may be too many and show that the student’s skill in conducting the experiment is less than a student who chooses a moderate number of trials. Thus, there needs to be some processing of the response data in order to convert them to a form that is interpretable as part of the summary scoring (discussed in the next subsection).

A potential—but not necessarily the only—way of scoring actions and combining them to an interpretable single score is “bundling.” For example, Table 21 shows how, on a series of four measurable actions in a simulation, different students might get the same raw score. The first column shows the actions upon which the students were judged. The four right-hand columns labeled Student A through Student D show the dichotomous scores allocated for completing each action successfully (1) or not (0). Note that the total score for each student is the same (3) but that each obtained it in a different way. Suppose it is of interest to see whether students can select and assemble the correct equipment and that to score highly on this task, students should do both correctly. This means that it would be preferable to give students C and D a higher score than students A and B. A “bundled” score for these scoring patterns is shown in the bottom row of the table. Students C and D are given four points, while students A and B get only three points.

Table 21. Example of Bundling Student Scores

Action	Student A	Student B	Student C	Student D
Selected correct equipment	0	1	1	1
Assembled equipment correctly	1	0	1	1
Ran sufficient trials	1	1	0	1
Accurately recorded data	1	1	1	0
Total Raw score	3	3	3	3
Bundled score	**3**	**3**	**4**	**4**

Processing of selected-responses is also necessary. Some may simply generate dichotomous scores where there are correct (1) and incorrect (0) options; however, there may be situations where one selection is optimal, and other selections generate partial scores.

It is recommended that the computer-based system be used to score constructed-responses as often as possible. Constructed-responses will be short answers that a student types into the computer. Response processing of the text of a student’s short written response will be handled by such technology as natural language processing to categorize responses according to level. Such technologies generally require that the system be “trained” by scoring a few hundred student responses with previously known scores (obtained by human scoring). Occasional checks by human scorers of subsequent computer scoring of assessment task responses also should be instituted.

Information on the various response types, response processing information, and weightings for responses will be stored in the ICT tasks/evidence database shown at the center of Figure 3. This information is called upon in the Activity Selection Process as an ICT task is specified for presentation and response processing. The evidence from all of a student’s interactions with a particular ICT are processed and sent to the next step in the cycle, the Summary Scoring Process, described in the next section.

<u>Summary Scoring Process</u>

The summary scoring process forms the second stage of the scoring. It accumulates the evidence that is generated from the response processing stage to update the Scoring Record, which represents the interpretations about the student’s knowledge, ability, and skills measured in the NAEP Science Assessment.

The NAEP Science Assessment adopts a student model for overall mastery (Almond et al., 2003) in which there is a single (continuous) student model variable that indicates the student's achievement level (Basic, Proficient, or Advanced) in science. This lends itself to a unidimensional IRT-based statistical process for producing the summary score, but other measurement models may be applicable. This statistical model will accumulate the information from across the elements of the task and produce a summary score on the NAEP achievement scale.

Development of Accessible Interactive Computer Tasks

The readability of text is critical on a computer screen because of the variations in the number of pixels that can represent letters in different font styles, sizes, and on monitors of differing resolutions (Nielsen, 2000). In general, the same fonts and sizes that maintain acceptable legibility on paper apply equally to text on computer screens. However, additional concerns include the contrast between background colors and fonts. Black text on a white background (known as "positive text") provides optimal legibility and should be used, unless the design of the ICT calls for some other background or text color. If other colors are required, a combination with a strong contrast ratio should be selected. Ease of reading and comprehension may also be affected by the layout of text on the screen. It is desirable, wherever possible, to make certain that text passages fit onto one screen and that scrolling is not required.

Students' ability to type on a keyboard will vary across grade levels assessed in the NAEP Science Assessment because of physical differences in the size of their hands (Waner, Behymer, & McCrary, 1992) and variations in their level of keyboarding skills due to opportunity to learn those skills (Fleming, 2002). Students at grade 4 may lack some of the keyboard skills to efficiently type a written response to an ICT. So, the design of items should minimize the amount of typed input at the 4th grade level and that input should be mainly via the mouse or other simple keyboard commands, such as the directional arrow keys and the "enter" button. Although it is more reasonable for 8th and 12th grade students to type written responses, to maintain high levels of accessibility to the assessment, the simplest and least cognitively demanding input method should be used. Even accuracy in the use of the mouse is dependent on age and the size of the "target" that children are expected to click on (Hourcade, Bederson, Druin, & Guimbretière, 2004). Paying attention to such differences in the design stages will help to ensure that the ICTs measure the relevant science performance expectation and that students are not hindered in their responses because they lack computer skills that are unrelated to what is being assessed.

A good way to ensure that ICTs are easy to use for the intended student group and that they produce measurable results is to conduct usability testing at key stages during development. The development of ICTs should follow good usability engineering practices for software that include the application of usability heuristics, testing and assessments methods (Nielsen, 1993; Nielsen & Mack, 1994). Usability testing should include trying the ICT on a range of different computers that are typical of the variety found in schools, since the platforms, operating systems, processor speeds, and software vary tremendously in educational settings. The usability testing should involve a typical cross-section of students who take the NAEP Science Assessment, including students with disabilities and English language learners. Also, given that the format of an ICT developed for the NAEP Science Assessment is likely to be unfamiliar to most students, students should be given an opportunity to view a demonstration of how to respond to such an item and given the opportunity to practice with a sample item prior to engaging with an item that will be scored.

Distribution of Items by Item Type

As measured by student response time, 50% of the assessment items at each grade level should be selected-response items and 50% should be constructed-response items (short constructed-response, extended constructed-response, concept-mapping tasks). NCES and the assessment development contractor should determine the relative proportion of kinds of items within the constructed-response category. Individual selected-response items and individual constructed-response items used within each of the item clusters, POE item sets, hands-on performance tasks, and ICTs should be included to count toward this 50%-50% distribution. If variation from the 50%-50% distribution becomes necessary as items are developed, preference should be given to constructed-response items.

With respect to the combination item types, the NAEP Science Assessment should contain *at least one of each* of the following at each grade level: item clusters, POE item sets, hands-on performance tasks, and ICTs. Also, it is highly recommended that, in grades 8 and 12, each assessment include at least one concept-mapping task.

Hands-on performance tasks and interactive computer tasks are administered to a subset of the students sampled. The number of hands-on performance and interactive computer tasks is specified in Table 22.

Table 22. Distribution of Items by Type of Item and Grade

	Grade 4 (# of tasks)	Grade 8 (# of tasks)	Grade 12 (# of tasks)
Hands-On Performance Task (HT)	≥ 1	≥ 1	≥ 1
Interactive Computer Task (ICT)	≥ 1	≥ 1	≥ 1
Total HT + ICT	≥ 4	≥ 4	≥ 4

In any grade, the number of interactive computer tasks (ICTs) should not exceed the number of hands-on performance tasks (HTs). No student will be administered both an ICT and a HT.

Assessment Item Contexts

There are certain components of science content, such as the history and nature of science, and the relationship between science and technology, with which science-literate citizens should be familiar. In Chapter Three, the nature of science is largely addressed through a discussion of the science practices (particularly Using Science Principles and Using Scientific Inquiry). The relationship between science and technology is partially addressed in a discussion of the Using Technological Design practice. The history and nature of science not only clarify facets of science practices but also the human aspect of science and the role science has played in various cultures. Students can see that science changes, and new conclusions can be made on the basis of new empirical data (e.g., development of the theory of plate tectonics, use of gaps in early versions of the Periodic Table to "discover" new elements). The reciprocal relationship between science and technology can be seen, for example, in that scientists use technological tools to empirically test proposed explanations for questions about the natural world; and engineers develop adaptations to the natural world to address human problems, needs, and aspirations based in part on science. From stimulating the development of germ theory by designing new health practices (19^{th} century) to crafting new alloys that stimulate scientists to search for explanations of the useful and emergent properties of the new metals (20^{th} century), crafts and engineering have spurred scientific thought and advancement.

When items are written to particular content statements, they may be framed in these contextual components of science content. Aspects of the history and nature of science and the relationship between science and technology should thus be incorporated into the contexts of assessment items as illustrated in the item below.

Item Suggestion

Ernest Rutherford found that when he fired alpha particles at a thin gold foil, some were scattered at large angles. What caused this scattering?

A. The gold's positive atomic nuclei attracted the negatively charged alpha particles.
B. The gold's negative atomic nuclei repelled the negatively charged alpha particles.
C. The gold's negative atomic nuclei attracted the positively charged alpha particles.
D. The gold's positive atomic nuclei repelled the positively charged alpha particles.

Key: D

P12.2, Identifying Science Principles

Item Development and Review

The item development and review process shall be consistent with the Guiding Principles set forth by NAGB for NAEP assessments. (See the "NAEP Item Development and Review Policy Statement" in Appendix J.)

In order to ensure the development of items that adequately represent the performance expectations and exhibit proper psychometric characteristics, as well as to construct an item pool that will facilitate the establishment of *a priori* achievement levels, the following guidelines should be followed. NCES and the assessment development contractor should conduct the initial item development and preliminary item review phase according to Guiding Principle 1, engaging "a minimum of 20% of the membership of the framework project committees in each subject area [science] to serve on the item writing and review groups as the NAEP test questions are being developed" (see Appendix J, p. 349). In addition, the review group should include scientists, teachers, science educators, measurement experts, reviewers trained in sensitivity review procedures, and representatives of state education agencies. The pool of items approved for further development should then go to field test in classrooms and cognitive labs (see below). Empirical information along with evidence of the link between the item and the *Framework* and *Specifications* documents should be subject to further expert review. The intent is for external reviewers to have at their disposal not only logical evidence that the item links to the *Framework* and *Specifications* and that it appears to evoke the intended performance expectation, but also empirical data that confirm, raise questions about, or disconfirm the logical evidence. Indeed, as item development iterates, it is important that the expert review group continue their reviews at several points during the assessment development process. This group, updated as needed because of changes in curriculum or assessment methodology, should function throughout the life of the *Framework* and *Specifications*.

The assessment development contractor should build careful review and quality control procedures into the assessment development process. Although large-scale field testing provides critical item-level information for assessment development, other useful information about the items should be collected before and after field testing. Before field testing, items and scoring rubrics should be reviewed by the expert review group described above. After field testing, the items and the assessment as a whole should be reviewed to make certain that they are as free as possible from irrelevant variables that could interfere with students' demonstrating their science knowledge and skills.

Sensitivity reviews are a particularly important part of the assessment development process. Reviewers, including educators and community members who are experts in the schooling or cultural backgrounds of the primary minority and special needs students who will be taking the assessment, evaluate assessment items and associated materials. The reviewers focus on identifying offensive or stereotypical subject matter and other construct irrelevant factors in the assessment. They provide valuable guidance about the context, wording, and structure of items; and they identify flaws in the items that confound the validity of the inferences for the groups of students they represent.

Two particularly useful procedures for collecting information about how items are working—classroom tryouts and cognitive labs—are described below. The information collected is valuable for determining whether items are measuring the construct as intended and for refining the items and scoring procedures before field testing. The information that the assessment development contractor garners from classroom tryouts and cognitive labs should be provided to item writers to help them develop new items and revise existing items before field testing, and it can be used to enhance item writing training and reference materials.

Classroom Tryouts

Classroom tryouts are an efficient and cost-effective way to collect information from students and teachers about how items and directions are working. Tryouts allow the assessment developer to troubleshoot the items and scoring rubrics. Classroom tryouts usually involve a non-random but carefully selected sample; the students should reflect the range of student achievement in the target population as well as represent the diversity of examinees. For example, the tryout sample should include urban and rural schools; schools in low, middle, and high economic communities; schools from different regions; and schools with students in all the major racial/ethnic categories in the population. The more the sample represents various groups in the testing population, the more likely the tryout will identify areas that can be improved in the items.

In addition to providing student response data, tryouts can provide various kinds of information about the items, including what students and their teachers think the items are measuring, the appropriateness of the associated assessment materials, and the clarity of the instructions. Students can be asked to edit the items, for example, by circling words, phrases, or sentences they find confusing and making suggestions for improvement. Teachers can ask students what they thought each item was asking them to do and why they answered as they did, and provide the information to the assessment developer. Teachers can also be asked to edit items and associated assessment materials. Item tryouts also are an efficient way to test how accommodations work and to try out new equipment and materials for hands-on performance tasks and interactive computer tasks.

Student responses to the items should be reviewed by content and measurement experts to detect any problems in the items and should be used along with the other information gathered to refine the items and scoring rubrics. Using a sample that includes important groups in the population will allow reviewers to look for issues that might be specific to these groups. Responses also are useful in developing training materials and notes for scorers.

Cognitive Labs

In cognitive labs (e.g., Leighton, 2004; Ruiz-Primo, Shavelson, et al., 2001; Taylor & Dionne, 2000), students are interviewed individually while they are taking or shortly after they have completed a set of items. These labs, then, provide *empirical* evidence on the extent to which the intended task demand is interpreted by the student in the same manner as intended by the item developer, and that the science practices and cognitive processes (cognitive demands) evoked by the items are the ones intended by the performance expectation. That is, as cognitive labs highlight measurement considerations in a more in-depth fashion than other administrations can, their use can provide important information for item development and revision. For example, cognitive labs can identify if and why an item is not providing meaningful information about student achievement, provide information about how new formats are working, or identify why an item was flagged by differential item functioning (DIF) analyses. Cognitive labs are particularly important for the *all new* NAEP Science Assessment to be administered in 2009.

The student samples used in cognitive labs are much smaller than those used in classroom tryouts. Students should be selected purposefully to represent the full range of diversity found in the populations assessed by NAEP and to allow an in-depth understanding of how an item is working and to provide information that will help in revising items or in developing a particular type of item.

Cognitive labs are time consuming and expensive. Consequently, not all items developed for the new science assessment will be tried out in such labs. (Nevertheless, all items should be field tested and item statistics collected and reported.) The following priorities should be followed in conducting cognitive labs on candidate assessment items: (1=Top) concept-mapping tasks, POE items, hands-on performance tasks, and interactive computer tasks; (2) short and extended constructed-response items; and (3) multiple-choice items that are a part of item clusters or tap the practices of Using Scientific Inquiry or Using Technological Design.

CHAPTER SIX: ADMINISTRATION OF THE ASSESSMENT

Introduction

Almost all students are eligible to be assessed in NAEP. This includes most students with disabilities (SWD) and those who are English language learners (ELL). Because the NAEP Science Assessment is designed to measure the academic achievement of all assessment takers at a given grade, students with a range of backgrounds and experiences, including those with disabilities and those who are mastering English as a second language, should be included in the assessment. To make certain that NAEP results provide a meaningful representation of all students' knowledge and skills, the science assessment should be designed to provide access for *all* students, that is, to provide all students with the opportunity to demonstrate what they know and are able to do.

There are two ways that NAEP addresses the issue of accessibility. One is to follow item and assessment development procedures that build accessibility into the standard assessment (see Chapter Five). The other is to provide accommodations for students with disabilities and for English language learners.[40] To the extent possible, in the 2009 science assessment and beyond, accommodations will be used to measure the full range of performance expectations represented in the *Framework* and *Specifications*.

In what follows, guidelines are set forth for administering the NAEP Science Assessment, for selecting students, and for reporting requirements and achievement levels. Particular attention is paid to assessment accessibility for all students (good item writing) and accommodations for SWDs and ELLs. This chapter concludes with recommendations for special studies.

NAEP Administration and Student Samples

As currently planned, the 2009 NAEP Science Assessment will be administered between January and early April, with the administration conducted by trained field staff. Each science assessment booklet will contain two separately timed, 25-minute sections of science items. There will also be extra 30-minute blocks for hands-on performance tasks and interactive computer tasks that are given to a subset of the students sampled. The assessment is designed with multiple forms of the test booklets. Items will be distributed across the booklets using a matrix sampling design so that students taking part in the assessment do not all receive the same items. In addition to the science items, the assessment booklets will include background questionnaires, administered in separately timed sessions.

---

[40] For more information about research that has guided the increased inclusion of students in NAEP, see Olson and Goldstein (1997) and Mazzeo, Carlson, Voelkl, and Lutkus (1999).

The assessment is designed to measure science achievement of students in the nation's schools in grades 4, 8, and 12 and to report the results at the national, regional, state, and urban district levels. To implement this goal, schools throughout the country are randomly selected to participate in the assessment. The sampling process is carefully planned to select schools that accurately represent the broad population of U.S. students and the populations of students in each state participating in State NAEP.

The selection process is designed to include schools of various types and sizes from a variety of community and geographical regions, with student populations that represent different levels of economic status; racial, ethnic, and cultural backgrounds; and instructional experiences. Students with disabilities and English language learners are included *to the extent possible*, with accommodations as necessary (see later in this chapter for more information about inclusion criteria and accommodations). The sophisticated sampling strategy helps to ensure that the NAEP program can generalize the assessment findings to the diverse student populations in the nation and participating jurisdictions. This allows the program, in aggregate, to present information on the strengths and weaknesses in students' understanding of science and their ability to apply that understanding; provide comparative student data according to race/ethnicity, eligibility for free/reduced-price school lunch, gender, state, and jurisdiction; describe trends in student performance over time; and report relationships between student achievement and certain background variables.

Student Access

When assessments are first conceptualized, they need to be thought of in the context of the entire population that will be assessed (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; NRC, 1999a; Thompson et al., 2002). Accessibility refers to the degree to which the assessment provides all students in the targeted population with the opportunity to demonstrate their achievement in relation to the performance expectations defined by the *Framework* and *Specifications*. The design for the NAEP Science Assessment must address issues of student access—considerations that can either facilitate or obstruct the goal of obtaining valid measurements of the targeted test takers' achievement in science.

The NAEP Science Assessment is designed to measure the achievement of students across the nation. Therefore, it should allow students to demonstrate what they know and are able to do in science including students who have learned science in a variety of ways, following different curricula and using different instructional materials; students who have mastered the content to varying degrees; students with disabilities; and students who are English language learners. The question to ask in developing the assessment is, "What is a reasonable way to measure the same intended performance expectations for students who come to the assessment with different experiences, strengths, and challenges; who approach the constructs from different perspectives; and who have different ways of displaying their knowledge and skill so that the widest possible range of students have appropriate opportunities to demonstrate what they know and can do?"

The central requirement of such an assessment is that the same science performance expectations are measured across diverse groups of students. To this end, the assessment should maintain the rigor of the science expectations in the *Framework* and *Specifications* while providing the means for all tested students to demonstrate their levels of knowledge and skills.

NAEP uses two methods to design an accessible assessment program: (1) developing the standard assessment so that it is accessible and, where necessary, (2) providing accommodations for students with special needs. The general item specifications in Chapter Five incorporate methods for making the assessment accessible to most students, (see "Principles of Good Item Writing"); these methods will be assumed, not reiterated, in this chapter. This chapter addresses accommodations and additional ways of making the assessment accessible to students with disabilities and English language learners.

Ways to strengthen access include the following:

1. Paying careful attention to how items are presented to students in the assessment (e.g., plain language editing procedures, use of graphics, item format considerations, use of laboratory equipment and materials)
2. Designing constructed-response items that allow for multiple ways of responding, as appropriate to the knowledge and skill assessed
3. Developing scoring rubrics so that the targeted knowledge and skills are evaluated at all score levels
4. Formatting assessment booklets to allow enough space between items, using boxes and lines judiciously
5. Providing proper training procedures and materials for scorers
6. Providing adapted forms for students with disabilities and students literate in Spanish
7. Providing administration accommodations for students with disabilities and English language learners

Before turning to these topics, a caveat is in order. There is "only limited systematic information about the use of assessment accommodations for elementary and secondary students with disabilities and even less about the effects of accommodations on the validity of scores" (Koretz & Barton, 2003-2004, p. 31; NRC, 2005c). Moreover, "[n]either educators nor researchers have reached consensus about the appropriateness of various accommodations …" (Koretz & Barton, 2003-2004, p. 35). The same holds true for English language learners (see NRC, 1997b). In addition, depending on the particular disability or the nature of the language difference, it may not be possible to assess all students in a manner that addresses their disability or language difficulty on the one hand and tests the same construct across such individual differences on the other hand (see Koretz & Barton, 2003-2004; NRC, 1997a). Given the trade-off of access for *all* students and adequate coverage of the science performance expectations in the *Framework* and *Specifications*, preference should be given to coverage of the science performance expectations, whether assessed through paper-and-pencil items, hands-on performance tasks, or interactive computer tasks.

NAEP Inclusion and Accommodations

NAEP should strive to develop science assessments that allow for the participation of the widest possible range of students, so that interpretation of scores of all who participate leads to valid inferences about the levels of their performance, as well as valid comparisons across states and with state assessments. All students should have the opportunity to demonstrate their knowledge of the concepts and ideas that the NAEP Science Assessment is intended to measure.

According to the National Research Council:

> Fairness, like validity, cannot be properly addressed as an afterthought once the test has been developed, administered, and used. It must be confronted throughout the interconnected phases of the testing process, from test design and development to administration, scoring, interpretation, and use (1999b, pp. 80-81).

As NAEP looks to measure the educational progress of *all* students in the nation's classrooms, assessment developers will encounter challenges that require giving deeper thought and consideration to the development of items, providing as fair a context as possible for all students. SWDs and ELLs present special challenges in how their science competencies can be assessed validly.

Attention to the identification and classification of students is needed not only to reduce current wide variation among states in identification criteria but also to determine appropriate accommodations, if necessary. NAEP is seeking to further refine the inclusion criteria and process in order to work towards greater standardization. Current NAEP inclusion criteria for students with disabilities and English Language Learners are as follows (quoted from NCES, 2005a, Current Policy section):[41]

[41] Identification of special needs students is problematic in and of itself. While NAEP defines inclusion criteria, student identification and classification is under the purview of states and localities. Consequently, there is wide variation from state to state on the percentage of students identified and classified as either SWD or ELL status. For example, the percentage of students identified as having a qualifying disability for the 2005 NAEP mathematics assessment varied across states from 10% to 24% at grade 4 and from 9% to 18% at grade 8 (Perie, Grigg, & Dion, 2005, p. 36). Similarly, the percentage of students identified as ELL for the 2005 NAEP mathematics assessment varied across states from 1% to 33% at grade 4 and from 1% to 21% at grade 8 (Perie, Grigg, et al., 2005, p. 36). Of course, much of the ELL variation depends on the overall demographics of particular states, but some of the variation is influenced by identification criteria and procedures. Indeed, Koretz and Barton (2003-2004) and Abedi (2004) have raised questions about the reliability and validity of student classification into SWD and ELL, respectively.

> A student identified on the Administration Schedule as having a disability [SWD], that is, a student with an Individualized Education Plan (IEP) or equivalent classification, *should be included* in the NAEP assessment *unless:*
> - The IEP team or equivalent group has determined that the student cannot participate in assessments such as NAEP, *or*
> - The student's cognitive functioning is so severely impaired that he or she cannot participate, *or*
> - The student's IEP requires that the student be tested with an accommodation that NAEP does not permit, and the student cannot demonstrate his or her knowledge of [name of subject area] without that accommodation.
>
> A student who is identified on the Administration Schedule as limited English proficient (LEP) or as an English language learner (ELL is the term used by NAEP for the 2005 assessment and beyond) and who is a native speaker of a language other than English *should be included* in the NAEP assessment *unless*:
> - The student has received reading or mathematics instruction primarily in English for *less than 3* school years including the current year, *and*
> - The student cannot demonstrate his or her knowledge of [name of subject area] in English even with an accommodation permitted by NAEP.[42]

Despite the differences between SWDs and ELLs, there are some commonalities, not the least of which is considerable heterogeneity within each of these groups as to assessment needs. In addition:

- Conceptual frameworks based on appropriate theories of language development and proficiency and of various forms of disabilities will be needed to build inclusive assessments.
- Financial and human resources will be needed over what is usually allocated in order to develop, administer, and interpret performance on relevant tasks.

Accessibility for Students with Disabilities and English Language Learners

Two general recommendations address both groups in the context of good assessment design for all students: readability of written text and alignment to content statements. Indeed, though commendable, efforts to conceptualize SWD and ELL testing have not gone much beyond putting forth criteria that apply to all good test development (e.g., Koretz & Barton, 2003-2004).

Students' ability to read and respond to written text often determines successful performance on assessments. Assessment items may pose an unfair disadvantage for some students if there is a heavy burden on reading skills when reading is not the target of the assessment. Language that is both straightforward and concise and that uses everyday words to convey meaning is needed. The goal of ensuring that language has these characteristics is to improve the comprehensibility of written text while preserving

[42] However, there will be a Spanish-language version of the 2009 NAEP Science Assessment to ensure inclusion of Spanish-speaking students.

the essence of its meaning. The use of language that reduces the linguistic demands placed on students reduces the effect of reading skills and language proficiency on students' science performance and assessment scores. (See p. 155 for more information on reading level.)

Items on the NAEP Science Assessment must be aligned to the content statements and science practices with the same depth and breadth of coverage and the same cognitive demands as specified in the *Framework* and *Specifications*. The emphasis in assessment design should be on accessibility using different formats, technologies, designs, and accommodations to include as many students as possible. It must be clear from the beginning that, to be equitable, assessments need to measure the achievement of all students on the same content and achievement standards.

To these ends, field tests should sample every type of student expected to participate in the final assessment administration, including students with a wide range of disabilities, English language learners, and students across racial, ethnic, and socio-economic lines. Field-testing NAEP items with a broad range of students will not only help determine whether items are unclear, misleading, or inaccessible for certain groups of students, but will also help ensure that assessment procedures are accessible to students when the NAEP Science Assessment is fully implemented.

Accommodations for Students with Disabilities and English Language Learners

NAEP strives to assess all students selected by its sampling process. Rigorous criteria are applied to minimize the number of English language learners and students with disabilities excluded from NAEP assessments. Participating students with special needs are permitted to use accommodations. An accommodation is a departure from standard testing (Koretz & Barton, 2003-2004; NRC, 1997a). The National Center for Educational Outcomes offers the following definition (see http://education.umn.edu/nceo/TopicAreas/Accommodations/Accomtopic.htm):

> Accommodations are changes in testing materials or procedures that enable students to participate in assessments in a way that allows abilities to be assessed rather than disability. They are provided to "level the playing field." Without accommodations, the assessment may not accurately measure the student's knowledge and skills.

Examples of accommodations range from extended time to use of a magnification device. As stated in current NAEP policy:

> All special-needs students may use the same accommodations in NAEP assessments that they use in their usual classroom testing unless the accommodation would make it impossible to measure the ability, skill, or proficiency being assessed, or the accommodation is not possible for the NAEP program to administer (NCES, 2005a, Current Policy section, ¶ 4).

For many students with disabilities and students whose native language is not English, the standard administration of the NAEP Science Assessment will be most appropriate. For some SWDs and ELLs, the use of one or more administrative accommodations will be more suitable. How to select and provide appropriate accommodations is an active area of research, and new insights are emerging (e.g., Koretz & Barton, 2003-2004). The NAEP accommodations policy allows for a variety of accommodations depending upon the needs of each student. Most accommodations that schools routinely provide in their own testing programs are allowed in the science assessment, as long as they do not affect the construct tested. Up to the 2005 NAEP assessment administration, the most frequently used accommodations included (quoted from the *NAEP 2005 Assessment Administrator Manual*, p. 2.37):

- Bilingual glossary (science bridge only)
- Bilingual booklet (mathematics operational and science only)
- Bilingual dictionary (mathematics operational and science only)
- Large-print booklet
- Extended time in regular session
- Read aloud in regular session (mathematics and science only)
- Small group
- One-on-one
- Scribe or use of a computer—used to record answers
- Other—includes format or equipment accommodations such as a sign language translator, amplification devices… (if provided by the school)
- Breaks during test
- Magnification device
- School staff administers

Other possible accommodations that may be available in states are not permitted by NAEP because they would alter the construct or skill being assessed, or they are not possible to administer due to logistics (e.g., audio tape administration of the assessment, extending sessions over multiple days) (NCES, 2005a, Current Policy section).

Accommodations are offered in combination as needed (e.g., students who receive one-on-one testing generally also use extended time).[43]

Students with Disabilities: Assessment Issues and Recommendations

A framework for students with disabilities should consider issues of classification, special access issues, and accommodations. Before turning to these considerations, note that most SWDs will take the standard assessment without accommodations, and those who take the assessment with accommodations will use the standard version of the test also. Item writers and the assessment developers should minimize item characteristics that could hinder accurately measuring the science achievement of students with disabilities by following these recommendations:

[43] The desirability of greater standardization in the use of accommodations is recognized, and NAEP is seeking to further refine its policies and administration guidelines on accommodations.

- Avoid layout and design features that could interfere with the ability of the student to understand the requirements and expectations of the item.
- Use plain language.
- Develop items so that they can be used with allowed accommodations.
- Address alternatives for students who are not able to use the equipment and materials necessary for responding to an item.

As noted, issues of identifying and classifying SWDs in disability categories are substantial (e.g., Koretz & Barton, 2003-2004). Classification is inconsistent at the teacher level, school level, and importantly, at the state level. The reported prevalence of specific disability categories shows great variability across states (Koretz & Barton, 2003-2004). These inconsistencies pose serious difficulties for assessment such as differential access to special accommodations across states; and if the accommodations are successful, test-performance interpretations (validity) become hazardous. Moreover, inconsistencies raise questions about assessment design and where the greatest emphasis on access of SWDs should be given. Finally, identifying accommodations on the basis of classification becomes problematic.

While states and localities retain the right to classify individuals into disability categories, there are several things that NAEP can do to improve access and provide opportunities for SWDs to demonstrate what they know and can do. First and foremost, assessment developers should take into account the relative prevalence of disabilities within the U.S. and design tests with these groups in mind, including students with specific learning disabilities and attention deficit disorders. The above advice on item characteristics for SWDs is relevant.

Second, in identifying appropriate accommodations for SWDs, consideration should be given to three things: (1) classification, (2) Individualized Education Plan (IEP), and (3) instructional environment. It is likely that student classification is a somewhat unreliable indicator for the necessary types of accommodations because of inconsistencies in classification and the paucity of research linking classifications to effective accommodations, that is, accommodations that do not change the performance expectation measured. Nevertheless, student classification is a starting point. The student's IEP may provide some insight as to what might be appropriate accommodations. Indeed, if permitted by NAEP, accommodations that are used during classroom instruction should also be used in assessment. On the other hand, if students do not need the accommodation for instruction, they may not need it in the testing situation. Finally, an understanding of the classroom instructional environment that has been adapted to assist student learning provides further ideas about appropriate accommodations.

Certain physical disabilities such as limited vision, hearing, or physical dexterity present challenges in fully assessing the science performance expectations for SWDs. This may be especially the case when special item formats are used to assess students' ability to conduct a scientific investigation (hands-on performance tasks) or to carry out a computer simulation (ICTs). To the extent possible, the assessment items should be designed in such a way as to allow permissible accommodations.

In some cases, ICTs may serve in lieu of hands-on performance tasks. This is especially the case with performance assessments where there appears to be an equivalence of hands-on and computer-generated science investigations although not all hands-on tasks can be done on the computer (e.g., Shavelson et al., 1991).

This said, given the importance of the performance expectations, SWDs in some cases may not have full access to the assessment even when their performance is accommodated. Priority should be given to assessing the science performance expectations as fully and validly as possible in the larger population.

English Language Learners: Assessment Issues and Recommendations

English language learners are bilingual. There are mental processes that are specific to being bilingual (Bialystok, 1991) that should be taken into account in making testing decisions about bilingual individuals. This section provides a brief framework with implications for identification of ELL students and their testing and accommodations.

In order to properly assess ELLs' science achievement, they should be classified as to their linguistic proficiencies in both their native language (L1) and their second language (L2) (Valdés & Figueroa, 1994), recognizing that bilingual individuals vary considerably as to their patterns of language dominance (Genesee, 1994). Not recognizing this fact makes ELL misclassifications likely. For example, a student who has been classified as limited English proficient based on oral communication skills in English is tested in her native language. Yet, this student's reading and writing proficiency in English might differ significantly from her English listening and speaking skills. Moreover, the reading and writing skills in the native language might be significantly lower than the reading and writing skills in English.

Simple criteria for classifying students' linguistic proficiencies may render flawed information about ELLs. For example, classifying ELL students based on the types of programs in which they are enrolled (e.g., dual language programs or full immersion programs) fails to consider the wide variety of patterns of language dominance among bilinguals. It also dismisses the notion that bilingual programs vary tremendously as to the fidelity with which they are implemented and the multitude of factors that shape their success (Cummins, 1999).

Ideally, proper testing decisions for ELLs—both in test development and in decisions about appropriate accommodations—should be based on accurate information on their skills across the four language modes (listening, speaking, reading, and writing) in both L1 and L2. At the very least, classifications of students according to their linguistic proficiencies should be based on information on the students' reading and writing skills in both L1 and L2. This is a sensible compromise, since large-scale assessment mainly depends on these two language modes.

To properly address language proficiency in testing, the relatively independent nature of language development across language modes (see Mackey, 1962) should be taken into consideration when identifying accommodations. For example, someone can be very good at reading but speak very poorly in a second language. In selecting accommodations for ELL students, then, consideration should be given to these modes such that the student has maximal access to performing on the test while not compromising the measured construct (performance expectation). By the same token, testing accommodations' policies for ELLs should be based on knowledge of their reading and writing skills in both L1 and L2. For example, simply providing the same accommodation to ELLs and SWDs (e.g., large font size and enhanced lighting conditions when ELLs are tested in English) or providing glossaries based on the assumption that all ELLs being assessed at a given grade level have comparable English reading skills may produce the illusion that fair and valid measures of academic achievement for ELLs are being obtained. Also, many students who are learning English as their second language will take the standard, English-only version of the assessment. They may have trouble understanding what items are asking for on assessment forms administered in English.[44]

ELL students are diverse both across and within their language groups. For example, this is the case with Spanish language speakers who come from various countries in Latin America. Among communities of native Spanish speakers in the U.S., there are considerable cultural differences, differences in socio-economic background, and differences in language usage (e.g., terms, idiomatic expressions) in both English and Spanish.

There are five actions that can be taken throughout the process of assessment development to address this tremendous diversity. Rather than thinking of them as exchangeable, these actions should be used complementarily.

First, even if they are to be tested only in English, ELLs should be included, along with native English speakers, in the pilot stages of assessment-item development—when item wording gets revised. By including ELLs in the samples of pilot students, test developers increase their opportunities to identify characteristics of the items that may pose unnecessary linguistic demands. They also increase their opportunities to properly address English variation among broad linguistic groups of ELLs (e.g., native Spanish speakers and native Tagalog speakers) and among groups of ELLs within the same linguistic group (e.g., native Spanish speakers from different regions of the U.S.). This is not a trivial matter, as there is evidence that the dialect of the language in which ELLs are tested (e.g., standard versus local) can be an important source of measurement error. This holds regardless of whether the language used to test them is English or their native language (Solano-Flores & Li, 2006).

Second, the samples of pilot ELL students should be large enough to be statistically representative. Their sizes should reflect the proportion of ELLs in the population of

[44] For more information about designing assessments that are accessible to English language learners, see Kopriva (2000).

students in U.S. schools. Also, to the extent possible, an effort should be made to ensure that the makeup of the ELL pilot samples reflects the diversity that exists within this segment of the population.

Third, appropriate specialists (e.g., linguists, cultural anthropologists, specialists in second language acquisition, specialists in general language development) should be involved in the entire process of test development, along with psychometricians, teachers, and content specialists. The contribution of these specialists should not be underestimated. For example, many issues regarding vocabulary (meaning of individual words), semantics (meaning of words in a sentence), syntax (grammatical structure of sentences), and pragmatics (interpretation of words and sentences in context) cannot be properly addressed without applying the methods and reasoning from anthropology and linguistics (see Solano-Flores & Trumbull, 2003). These specialists should participate as test developers for a significant number of hours, not just be hired as consultants near the end of the process of test development to perform reviews of test materials.

Fourth, in order to minimize misinterpretation of items by ELL students, language in item writing has to be addressed at several levels of complexity. The following recommendations address some of the most important issues:

- Item reading level should be below grade level so as not to confound the performance expectation being measured and reading ability. However, reading level should be determined based on piloting items and using the expertise of the professionals mentioned above. As discussed in Chapter Five, the use of readability formulas as tools for determining item reading level should be discouraged, as they may render spurious information on the reading difficulty of items. The use of readability formulas can be appropriate for analyzing large amounts of text, not test items (see Oakland & Lane, 2004).
- Item writers should be aware that words that pose linguistic difficulty in science may not only be technical terms but also logical connectives (e,g., *simultaneously*, *essentially, in addition to*), common terms that have more than one meaning (e.g., *bond*), and terms that have subtle different meanings in science and in everyday life (e.g., *mass*) (see Wellington & Osborne, 2001). This fact, which is relevant to testing in science for all students, regardless of their proficiency in English, becomes even more important in testing ELLs. Additional issues that are particularly relevant to ELL testing include the use of cognates and false cognates. Cognates are words in two languages that have a common ancestor and are similar in form and meaning (e.g., *plant* in English and *planta* in Spanish); test writers can capitalize on them to minimize the linguistic demands of items. False cognates are words with similar appearance but different meaning in two languages (e.g., *billion*, which in English means one thousand millions and *billón*, which in Spanish means two million).

- Item writers also should be aware that language is relevant to both the ways in which disciplinary knowledge is transmitted and the ways in which acquisition of that knowledge is assessed. This dual role of language poses a challenge to item writers. They need to be able to determine when language is part of the construct that an item is intended to assess and when language is a source of bias and measurement error. Only where the use of *scientific language* is at the core of the performance expectation tested should technical words, scientific discourse, scientific notation, and other aspects of scientific language be included in an item. For instance, if the intent is to measure students' ability to use the term, "mass" in an application setting, the term should not be substituted by "weight," as the two terms have subtle but important differences in meaning. By contrast, if the intent of the item is not to evaluate students' understanding of scientific terminology and "mass" may not be understood by the student, then it is appropriate to use "weight" instead of "mass." See textbox on p. 26 for more detail on the distinction between weight and mass.

Fifth, several important issues related to the proper interpretation of ELL students' responses must be properly addressed to ensure sound and fair scoring of constructed-response assessment items. Literacy issues and varied background experiences have an impact on how well scorers can properly read, understand, and evaluate the responses of English language learners to constructed-response items.[45]

ELL students' written responses to constructed-response items administered in English may be difficult to interpret properly. One obvious reason is that, because ELLs are still developing their second language, the characteristics of their writing may deviate from the norms and conventions of written Standard English (e.g., spelling, syntactical structure). Another, more subtle reason, is that, as a part of the natural process of bilingual development, ELLs use features (e.g., terms, meanings, expressions, grammatical forms, discursive styles) that belong to their native languages in their writing in English. Scorers who are not aware of this fact may wrongly dismiss ELLs' written responses as incorrect, based on their superficial features, without carefully considering the essence of the ideas these students intend to express.

It is incorrect to assume that certain types of assessment items (e.g., concept-mapping tasks, ICTs) are language-bias free simply because the reading or writing demands are different from other types of items (e.g., extended constructed-response items). Each type of item is sensitive to different types of reasoning and thus poses different (rather than lesser or fewer) sets of linguistic demands to students; these demands are especially challenging for ELLs. The evidence that exists about the linguistic demands posed by a given assessment item must be carefully examined before any assumptions are made about the item's ability to reduce the impact of language proficiency on the validity of assessment scores.

---

[45] For more information about interpreting and scoring responses from English language learners, see Kopriva and Saez (1997) and Solano-Flores, Lara, Sexton, and Navarrete (2001).

It is important for scorers to evaluate responses based on the measurement intent of the item and recognize when an unusual response is actually addressing that intent. The following actions should be taken to fairly and properly score ELL student responses:

- Bilingual teachers should be involved in the different aspects of scoring, from the selection of benchmark and training responses to the development of scoring rubrics to the scoring of student responses.
- Experts in reading responses of ELLs should be available to scorers throughout the scoring process.
- Scoring leaders should have additional training in recognizing and properly interpreting responses from English language learners.
- To ensure accurate interpretations of ELL students' written responses, scorer training materials and benchmark student responses should illustrate features typically observed in the writing of ELLs in English. Examples of these features are as follows:
    1. Code-switching—use of both the student's native language and English, ranging from the insertion of some words from the native language to the use of the two languages in the same sentence without violating the syntactical rules of either one of the languages (see Romaine, 1995)
    2. Use of native language phonetics in attempting to write English or beginning-stage English phonetic spelling (e.g., *de ticher sed*—"the teacher said")
    3. Use of writing conventions from the native language (e.g., *today is monday*—the names of the weekdays are not capitalized in standard Spanish)
    4. Use of technical notation conventions from the student's culture (e.g., $25,00 to express twenty-five dollars)
    5. Imprecise use of words (e.g., *I called with him*—"I spoke with him")
    6. Unusual sentence structures (e.g., *I don't know what is the response*)
    7. Unusual discourse structure (e.g., lack of a topic sentence at the beginning of a paragraph, which may reflect the discursive style that is acceptable in the student's formal written language)
    8. Over-reliance on non-verbal forms of communication (such as charts or pictures) with limited writing
    9. Others, including transposition of words (e.g., *the cat black*), and omission of tense markers (e.g., *Yesterday he learn at lot*), articles (*I didn't see it in notebook)*, plurals (*the horse are gone*), prepositions (e.g., *explain me what you said*), or other words

Novel interpretations and responses are common for English language learners and often reflect background experiences quite different from that of most native English speakers. Circular, indirect, deductive, and abbreviated reasoning writing styles are encouraged by some cultures, and scorers should be trained to appropriately score these types of responses. Also, it is not unusual for scorers to give high scores to long, elaborate responses. However, some cultures discourage long responses to questions among children, especially when adults or authority figures ask the questions. Such a pattern of communication can be reflected in the written responses of ELLs to open-ended questions as short sentences. Despite being short, some responses to open-ended items may be correct. When a specific writing style is not the measurement intent of the item, scorers need to both understand the nature, conventions, and approaches of these kinds of styles and know how to separate the structure and sophistication of the written response from the substantive content being evaluated.

Accessibility and Accommodations for Interactive Computer Tasks (ICTs)

The accessibility and accommodation issues discussed in this chapter apply to all parts of the NAEP Science Assessment, including the interactive computer tasks (ICTs). The computer-based delivery of ICTs raises some additional concerns for accessibility, but it also offers the opportunity to embed some accommodations that are not available in a paper-and-pencil format. As with items in other parts of the assessment, ICTs should be constructed with accessibility as one of the general principles of item design. Accessibility concerns for ICTs go beyond those for paper-and-pencil formats in the ways discussed in the following paragraphs.

While computer screens present some extra challenges for legibility of text, computers also provide opportunities for building into the design of the ICT features that will improve accessibility to text for SWD and ELL students. Allowing students to enlarge the font size or change the background/text colors can improve the readability for students with visual impairments. Students with more severe visual impairment or ELL students whose understanding of spoken English is above their comprehension of written English may benefit from an option to hear the text in the ICT read by the computer, either by a text reading program or via pre-recorded readings of the text (e.g., Ketterlin-Geller, 2005). Similarly, ICTs should be designed whenever possible to offer options that will improve accessibility in other ways. For example, when graphics are used, students could be given the option of increasing the size of the graphic.

Some general advice on developing computer-based assessments that are inclusive of students with disabilities is provided by the National Center on Educational Outcomes (Thompson, Thurlow, & Moore, 2003). The following proposed development steps are adapted from their work:

- Step 1. Assemble a group of experts to guide the development. Include experts on assessment design, accessible test and Web design, and assistive technology, along with state and local assessment and special education personnel and parents.
- Step 2. Decide how each accommodation will be incorporated into the computer-based test. Examine each possible accommodation in light of computer-based administration.
- Step 3. Consider each accommodation or assessment feature in light of the constructs being tested.

- Step 4. Consider the feasibility of incorporating the accommodation into computer-based tests. Construct a specific plan for building in features that are not immediately available and conduct extensive pilot tests with a variety of equipment scenarios and accessibility features.
- Step 5. Consider the implications for scoring when dealing with responses to ICTs from SWD and ELL students.
- Step 6. Consider training implications for staff and students. The best technology will be useless if students or staff do not know how to use it. Special consideration needs to be given to the computer literacy of students and their experience using features such as screen readers. Information about the features available on computer-based tests needs to be available to IEP teams to use in planning a student's instruction and assessments. Practice tests that include relevant features need to be available.

Reporting Requirements and Achievement Levels

The assessment should be designed so that reliable and valid inferences from the results can be drawn for the population and for subgroups by scale scores and achievement levels. Because the results are intended to describe the achievement of all students in the nation, results should provide an accurate picture of achievement across the entire score scale. The test developer should design the assessment so that the performance measured is aligned with the performance expectations in the *Framework* and *Specifications* and the knowledge and skills described in the achievement levels.

Any assessment, no matter how carefully crafted, including NAEP, taps only partially the students' achievement in a subject-matter domain. This caveat is particularly important in science, given the breadth of the field and the complexity and the depth of knowledge available. Likewise, the methods of science are challenging. An understanding of these methods, even at a rudimentary level, is important as is the understanding of how these methods warrant knowledge claims about the natural and human-made worlds. More on the uses and challenges of NAEP data is provided below.

Uses of NAEP Data

For more than four decades, NAEP has provided information integral to reporting on the condition and progress of education at grades 4, 8, and 12 for the nation, and more recently, for the states and for a set of large, urban school districts. Legislation concerning NAEP states that its purpose is to provide, in a timely manner, a fair and accurate measurement of student academic achievement and reporting of trends in such achievement in reading, mathematics, and other subject matter (Public Law 107-279).

Because of its rigorous design and methodology, NAEP reports are increasingly used for monitoring the state of education in the subjects that are assessed, as models for designing other large-scale assessments, and for secondary research purposes.

Monitoring

As the nation's only ongoing survey of students' educational progress, NAEP has become an increasingly important resource for obtaining information on what students know and can do. Because the information it generates is available to policymakers, educators, and the public, NAEP can be used as a tool for monitoring student achievement in reading, mathematics, science, and other subjects at the national, state, and selected district levels. For example, NAEP reports, known as "The Nation's Report Card," compare student performance in a given subject across states, within the subject over time, or among groups of students within the same grade. NAEP also reports long-term achievement trends for 9-, 13-, and 17-year-olds in reading and mathematics (e.g., Perie, Moran, & Lutkus, 2005). To the extent that individual state standards reflect the common core of knowledge and skills specified in the *Framework* and *Specifications*, state comparisons can legitimately be made. If a state has unique standards, any comparison is limited by the degree of mismatch between NAEP content and state content. Even with this caveat, NAEP still stands as a key indicator of what students know and can do in science at grades 4, 8, and 12.

Model of Assessment Development and Methods

NAEP assessment frameworks and specifications documents are themselves used as resources for international, state, and local curriculum and assessment. The broad-based process used in the development of the frameworks and specifications means that current thinking and research is reflected in these descriptions of what students should know and be able to do in a given subject. In addition, NAEP uses a rigorous and carefully designed process in developing the assessment instruments themselves. Pilot tests and internal and external reviews ensure that NAEP assessments are reliable and valid with respect to what they attempt to accomplish. This sophisticated methodology serves as a model for other assessment developers. Given the requirements contained in *No Child Left Behind* to assess students in science, states may wish to use NAEP as one model to guide their own assessment development.

Research and Policy

The data NAEP provides include subject-matter achievement results (reported as both scale scores and achievement levels) for various subgroups; background information about schools, teachers, and students at the subgroup level (e.g., course-taking patterns of Hispanic male 12th graders); state-level results; reports for a set of large urban districts; history of state and district participation; and publicly released assessment questions, student responses, and scoring guides. The NAEP Web site, http://nces.ed.gov/nationsreportcard, contains user-friendly data analysis software to enable policymakers, researchers, and others to examine all aspects of NAEP data, perform significance tests, and create customized graphic displays of NAEP results. These data and software tools can be used to inform policymaking and for secondary analyses and other research purposes.

Challenges of Developing a NAEP Assessment

There are three major challenges in developing NAEP assessment frameworks and specifications documents, and, in particular, this *Framework* and *Specifications*. One such challenge arises from measurement constraints and the nature of the items included on the assessment; the next relates to time and resource constraints and how much can be assessed in NAEP; and the last comes from the time horizon for the *Framework* and *Specifications* and the difficulty of developing ten-year recommendations with the rapid explosion of knowledge in the Information Age. Each of these is discussed in detail below.

Measurement Constraints

NAEP, like any large-scale assessment in education, the workplace, or clinical practice, is constrained in what it can measure. This has implications for the proper interpretation of NAEP Science Assessment results. The *Framework* is an assessment framework, not a curriculum framework. Although the two are clearly interrelated, each has a different purpose and a different set of underlying assumptions. 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 pick and choose what and how they teach. An assessment framework is a subset of the achievement universe from which assessment developers must choose to develop sets of items that can be assessed within time and resource constraints. Hence, the science content to be assessed by NAEP has been identified as that considered central to the Physical, Life, and Earth and Space Sciences. As a result, some important outcomes of science education that are difficult and time-consuming to measure—such as habits of mind, sustained inquiry, collaborative research—but valued by scientists, science educators, and the business community, will be only partially represented in the *Framework* and *Specifications* and on the NAEP Science Assessment. Moreover, the wide range of science standards in the guiding national documents that could be incorporated into the *Framework* and *Specifications* had to be reduced in number so as to allow some in-depth probing of fundamental science content. As a result, the *Framework* and *Specifications* represent a careful distillation that is not a complete representation of the original universe of achievement outcomes desirable for science education.

Time and Resource Constraints

What NAEP can assess is limited by time and resources. Like most standardized assessments, NAEP is an “on demand” assessment. It ascertains what students know and can do in a limited amount of time (50 minutes for paper-and-pencil questions and, for a subset of students sampled, an additional 30 minutes for hands-on performance or interactive computer tasks), with limited access to resources (e.g., reference materials, feedback from peers and teachers, opportunities for reflection and revision). The national and state standards, however, contain goals that require extended time (days, weeks, or months). To assess the achievement of students in the kinds of extended activities that are a central feature of the national and state standards and many science curricula, then, it would be necessary to know, for example, the quality of students'

- reasoning while framing their research questions;
- planning for data collection and the execution of that plan;
- abilities to meet unpredictable challenges that arise during an actual, ongoing scientific investigation;
- lines of argument in deciding how to alter their experimental approach in the light of new evidence;
- engagement with fellow students and/or the teacher in interpreting an observation or result and deciding what to do about it; and
- deliberations and reasoning when settling on the defensible conclusions that might be drawn from their work.

NAEP, like other "on demand" assessments, then, cannot be used to draw conclusions about student achievement with respect to the full range of goals of science education. States, districts, schools, and teachers can supplement NAEP and other standardized assessments to assess the full range of science education standards. In addition to describing the content and format of an examination, assessment frameworks and specifications, like these, signal to the public and to teachers what elements of a subject are important. The absence of extended inquiry in NAEP, however, is not intended to signal its relative importance in the curriculum. Indeed, because of the significance of inquiry in science education, the *Framework* and *Specifications* promote as much consideration of inquiry as can be accomplished within the time and resources available for assessment.

<u>Balancing Current and Future Curricula</u>

The *Framework* and *Specifications* attempt to strike a balance between what can reasonably be predicted about future school science and what students are likely to encounter in their curriculum and instruction now and in the near future. It is a significant challenge to write recommendations for the future. Cutting-edge science research creates new knowledge at the intersection of disciplinary boundaries. For example, research on human and natural systems has generated new understandings about environmental science that are closely linked to knowledge generated in the physical, life, and Earth and space sciences. Although the *Framework* and *Specifications* are organized into the more traditional Physical, Life, and Earth and Space Sciences, features of current science research are woven throughout.

Another example of burgeoning knowledge relates to technology and technological design and the role of both in the NAEP Science Assessment. Technology and technological design are closely interrelated with science, yet the focus of this assessment is science. Hence, technology and technological design are included in the *Framework* and *Specifications* but are limited to that which has a direct bearing on the assessment of students' science achievement. (See Chapter Three.)

The *Framework* and *Specifications* are intended to be both forward-looking (in terms of what science content will be of central importance in the future) and reflective (in terms of current school science). Because it is impossible to predict with certainty the shape of school science in 2009 and beyond, the choices made for 2009 should be revisited in response to future developments in school science.

Reporting Requirements

The NAEP Science Assessment reports results for the nation's students in grades 4, 8, and 12, and for participating states' students in grades 4 and 8, as well as for subgroups of the population defined by specific demographic characteristics such as race/ethnicity, eligibility for free and reduced-priced school lunch, gender, and type of school attended (see www.nationsreportcard.gov for more information about subgroup scores). Scores are not reported for individual students or schools.

Results are typically reported in two ways: scale scores and achievement levels. Scale scores can range from 0 to 500 (0-300 on the 1996-2005 NAEP Science Assessment), and the average scaled score and standard deviation are reported. The scale score range for the new 2009 NAEP Science Assessment has not yet been determined. Achievement levels identify students at three levels—Basic, Proficient, and Advanced—and the percentage of students who reach these achievement levels are reported. These achievement levels describe what students should know and be able to do in science.

The NAEP Science Report Card should report a composite scale score and content area subscale scores (for Physical Science, Life Science, and Earth and Space Science). As the assessment resulting from the *Framework* and *Specifications* will start a new NAEP "science trend," there will be no trend data reported in the 2009 NAEP Science Report Card; trend data should be reported for subsequent NAEP Science Assessments. Because trends over time need to be maintained in both the composite scale score and the subscale scores, each content area needs to be sampled adequately to ensure parallel item pools across years.

To the extent possible, scaled scores for the four science practices should be reported on the NAEP website. Alternatively, the practices of Identifying Science Principles and Using Science Principles might be combined and a scaled score reported; likewise, scaled scores might be reported for Using Scientific Inquiry and Using Technological Design combined.

With the new types of items and assessment tasks introduced by the *Framework* and *Specifications*, additional reporting requirements are needed. First, the NAEP Science Report Card should include exemplars of each of the new types of items and assessment tasks. Additional exemplars should be released for inclusion in the NAEP data tool. Second, in addition to these items' scores being interpreted with NAEP scaled scores, it is recommended that the NAEP Science Report Card and the NAEP data tool provide additional information on released items. For released items and item clusters focusing on learning progressions (naïve conceptions), the percentage of students at each of the various performance levels should be reported. For released hands-on performance tasks and ICTs, the percentage of students taking one or another solution path (correct or incorrect) should be reported. Indeed, the NAEP data tool should permit the public to explore students' alternative conceptions of the natural world or their solution paths for hands-on performance tasks or ICTs.

Achievement Levels

Public Law 107-279 specifies NAGB's responsibilities regarding NAEP, including the identification of appropriate achievement goals for each age and grade in the subject areas assessed by NAEP. The achievement level descriptions are statements of what students should know and be able to do on NAEP at grades 4, 8, and 12. To fulfill its statutory responsibility, the Governing Board developed a policy in 1989 (modified in 1993) to guide the development of achievement levels for all NAEP subjects. Three levels of achievement were identified to provide the public, educators, and policymakers with information on student performance on NAEP. These levels—Basic, Proficient, and Advanced—are used as a primary means of reporting NAEP results to describe "how good is good enough" at grades 4, 8, and 12. See Appendix C for the NAEP Science Preliminary Achievement Level Descriptions.

Table 23 displays the Board's generic policy definitions for Basic, Proficient, and Advanced achievement that pertain to all NAEP subjects and grades.

Table 23. Generic Achievement Level Policy Definitions for NAEP

Achievement Level	Policy Definition
Advanced	This level signifies superior performance.
Proficient	This level represents solid academic performance for each grade assessed. 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.
Basic	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.

During the framework development process, the project committees are asked to develop preliminary achievement level descriptions, based on the generic policy definitions, to guide item development. Essentially, the purpose of these statements is to provide examples of what students performing at the Basic, Proficient, and Advanced achievement levels should know and be able to do in terms of the science content and practices identified in the *Framework* and *Specifications*. The intended audiences for these preliminary descriptions are the NAEP assessment development contractor and item writers. The descriptions are used to ensure a broad range of items (covering Basic, Proficient, and Advanced levels) is developed at each grade level.

Tables 24 to 26 (Appendix C) present the preliminary achievement level descriptions for grades 4, 8, and 12 to clearly illustrate the science content and practices expected at each grade level. Members of the Planning Committee used their expert judgment to draft these preliminary descriptions and were partially guided by the following considerations:

- Advanced descriptions could be composed of highly challenging science content and highly challenging components of a given science practice.
- Basic descriptions could be composed of both moderately challenging content and practice components.
- Proficient descriptions could be composed of a mixture of highly and moderately challenging content and practice components.

The preliminary descriptions include only *illustrative* statements drawn from the science content and practices defined in the *Framework* and *Specifications*. The statements are not intended to represent the entire set of objectives from the content and practice dimensions, nor do the preliminary achievement level descriptions denote a sense of priority or importance based on the statements selected. In addition, these descriptions should be interpreted neither as assessment items nor as suggestions for assessment items. Although adjectives are not explicitly included in every statement, each was written with the expectation that student performances would be "good" and "reasonable." For example, the phrase "design an investigation" implies "design an organized and logical investigation" and "propose a framework" implies "propose a rational framework."

After the assessment is administered, broadly representative panels engage in a standard-setting process to determine the achievement level cut scores on the NAEP scale. The cut scores represent the minimum score required for performance at each NAEP achievement level. A second outcome of this standard-setting process is a set of paragraphs, derived from the preliminary achievement level descriptions, to be used in reporting the NAEP science results to the general public and other audiences. At each grade level, there will be paragraphs describing what students should know and be able to do at the Basic, Proficient, and Advanced levels in terms of the science content and practices identified in the *Framework* and *Specifications*.

Further information on NAEP achievement levels can be found at www.nagb.org.

Special Studies

Special studies bearing on aspects of the 2009 NAEP Science Assessment are presented in the *Framework* and *Specifications*. Each would contribute to a further understanding of science assessment.

Group 1 Special Studies are recommended as having highest priority:

- "Exchangeability" of Hands-on Performance and Interactive Computer Investigations
- Impact of Variation in Item Format and Language Demand on the Performance of English Language Learners and Students with Disabilities
- Computer Adaptive Testing to Assess the Development of Student Understanding of Earth Systems

Group 2 Special Studies are recommended as having lower priority:

- Knowing What Students Know about Technological Design
- Extended Investigations by Students

Note that the order in which studies are listed does *not* imply priority within Group 1 or Group 2. Group 1 Special Studies are presented below, and Group 2 Special Studies are located in Appendix E.

Group 1 Special Studies

<u>"Exchangeability" of Hands-on Performance and Interactive Computer Investigations</u>

Inquiry is at the heart of knowing and doing science. A fundamental aspect of inquiry is the design, conduct, and interpretation of empirical investigations to answer a question or test a hypothesis. While a full assessment of inquiry is not possible on any test that is given on demand, hands-on performance investigations attempt to approximate this aspect of inquiry under time, space, cost, and logistic constraints. For this reason, hands-on performance investigations have been a part of the NAEP Science Assessment since 1996.

These hands-on performance investigations (HI), however, have been criticized as costly, logistically difficult, and too highly structured. On the other hand, interactive computer tasks or, in this case, interactive computer investigations (ICI), are logistically simpler and lower in cost, as well as more open-ended. Consequently, the purpose of this study is to explore whether ICI and HI are exchangeable. The question is *not* whether ICI could replace HI either on NAEP or in the classroom—it should not. Even if these two approaches produce quite similar performances and scores, each affords somewhat different opportunities; simulations are just that and are not exchangeable with actual practice. This is an assessment question: can the cost and logistical challenges of HI be reduced with ICI and *still measure the same competencies as reliably and validly?* Some research suggests that the two methods of assessing student inquiry are, to a fair degree, exchangeable (e.g., Pine et al., 1993; Rosenquist et al., 2000). Yet, further research is needed on several different investigations to provide a satisfactory answer for large-scale assessment.

Specifically, this study would address the following research questions:

- Does choice of ICI or HI limit the questions that may be asked? Specifically, is there something of value in HI that cannot be asked if the ICI is administered?
- Are scores on HI and ICI equally reliable?
- Are scores on HI and ICI of equal magnitude?
- To what extent does performance on HI predict performance on ICI of the same investigation?
- Do scores on HI and ICI correlate fairly equally with scores on another measure of science inquiry or achievement?
- Are similar thinking processes evoked by HI and ICI?
- Do the answers to these questions depend on individual differences among students, such as gender, English proficiency, race/ethnicity, socio-economic status, and geographic location? (Also of interest is variation in student access to computers, which may be confounded with the other variables listed here.)

<u>Impact of Variation in Item Format and Language Demand on the Performance of English Language Learners and Students with Disabilities</u>

English language learners and students with disabilities do not perform as well on standardized achievement assessments, even accounting for background. Recent studies, for example, have pointed to a systematic relationship between the linguistic complexity of the assessment and the test scores of English language learners (e.g., Abedi, 2003) and students with certain disabilities such as those related to reading and information processing. Science assessments, with their heavy verbal load, may exacerbate performance disparities. In cases where this relationship is demonstrable and test items are high in language complexity, the differences become sources of measurement error and construct irrelevant variance, so that the nature of the assessment item must be addressed. Until this dimension of the assessment item is more clearly understood, any interpretation of the performance of English language learners or students with disabilities on a content assessment is problematic; language proficiency, for example, and science understanding cannot be disentangled.

Preliminary results from several studies of scaffolded science assessments that are designed to minimize language complexity and provide alternative response modalities—including graphic organizers or drawn representations of the concepts—indicate that English language learners and students with disabilities may be able to demonstrate content knowledge at a higher level if a variety of response options are available to them (Dalton, Morocco, Tivnan, & Rawson, 1994; Delgado, 2005). Further research is needed to clarify the relationship between language complexity, scaffolded assessment items, and the performance of English language learners and students with disabilities.

Specifically, this study would address the following research questions:

- Can the language complexity of a content-based assessment be systematically measured?
- Can content-based assessment items be designed to minimize the language demand while conserving the content information obtained?
- If the content-based assessment contains a graphic response modality, do English language learners and students with disabilities demonstrate higher understanding of the content concept being assessed relative to more linguistically demanding response modalities?
- When the content-based assessment with a graphic response option is also computer-based, is there a further benefit in terms of content concept conservation and these students' performance?

<u>Computer Adaptive Testing to Assess the Development of Student Understanding of Earth Systems</u>[46]

A common critique of large-scale assessment is that its necessary reliance on easily-scored, decontextualized, and decomposed items has led to an impoverished range of potential learning activities from which valid and reliable measures might be derived (Resnick & Resnick, 1992). Among attempts to find alternatives have been (a) the Facets approach (Minstrell, 1998), which posits a strong model of "facets" of student knowledge for certain science topic areas and uses coordinated sets of multiple-choice items to hone in on students' particular conceptions and misconceptions and (b) the progress variable approach (Masters, Adams, & Wilson, 1990), which posits a learning progression and uses Item Response Theory to scale students' responses to (typically open-ended) items to estimate in which part of the learning progression students are most likely located.

This special study combines the strengths of each of these approaches to develop a new type of "branching" item that can be used to investigate (a) the more complex types of knowledge structures and (b) complex procedural steps involving contingencies such as those common in inquiry-related contexts, and yet maintain the efficiency of traditional multiple-choice testing. Specifically, the Facets approach will contribute its strong knowledge structure and convenient scoring, and the progress variables approach will contribute the interpretational framework of the learning progression and the flexible statistical modeling available through recent advances in item response modeling (De Boeck & Wilson, 2004). Together, these make possible the utilization of item bundles such as that shown in Figure 4 to provide both the usual result in terms of student ability estimation, as well as potentially more educationally informative results such as the prevalence of particular classes of misconceptions among the student body.

---

[46] Currently, a number of states are using computerized testing, consisting largely of translating traditional paper-and-pencil items into computer-based delivery systems. The study suggested here, as well as the first study on p. 222, makes more extensive use of the capabilities inherent in computer-based assessments.

Figure 4. Storyboard Showing Item Design for a "Branching" Item Bundle on Ions and Atoms (adapted from Scalise, 2004)

Matter Composition: Ions and Atoms Item Bundle

1. Lead-based paint contains Pb^{2+} ions and lead pipes are made up of Pb atoms
The main difference between Pb^{2+} ion and Pb atom is:
A. They are basically the same. (go to question 2)
B. They have a different number of electrons. (go to question 3)
C. They have a different number of protons (3)
D. They are different but not in the ways described. (go to question 2)
E. I don't know (0)

2. Choose the answer with which you most agree.
Pb^{2+} ion and Pb atom are the same except:
A. Pb^{2+} has ionic bond and Pb has atomic bonds. (3)
B. Pb^{2+} ion and Pb are similar but used differently. (1)
C. Pb^{2+} is a liquid, Pb is a sold. (2)
D. Pb^{2+} requires two Pb atoms. (2)

3. Pick the best answer below:
A. Pb^{2+} has 2 fewer electrons than Pb. (go to question 4)
B. Pb^{2+} has a larger e- density cloud. (3)
C. Pb^{2+} is positively charged so has 2 extra valence electrons. (4)

4. Electrons can be thought of as arranged in shells or orbitals.
Select the ground state electron configuration for carbon:
A. $1s^2 2s^1 2p^3$ (7)
B. $1s^2 2s^2 2p^2$ (go to question 5)
C. $1s^2 2s^2 2p^1$ (6)
D. None of these (5)

5. Certain main group elements exhibit multiple oxidation states, including Group 14 of which carbon is a member. It would seem that a possible oxidation state for carbon is +2, with two electrons available for bonding in half-filled $2p_x$ and $2p_y$ orbitals. Based on valence bond theory, show an electron configuration justifying the formation of four bonds for carbon rather than two and describe the type of orbitals formed. EXPLAIN your answer as fully as possible.
(All respondents who reach this screen will be scored as an 8.
Answers submitted here will be used to build distractors for this question.)

Possible scores and parameters for this item bundle under iota model (assume constraint on cases):

score	0	1	2	3	4	5	6	7	8
# paths	1	1	2	3	1	1	1	1	1
parameters	δ_{71}	δ_{72}	δ_{73}, ι_{731}	$\delta_{74}, \iota_{741}, \iota_{742}$	δ_{75}	δ_{76}	δ_{77}	δ_{78}	δ_{79}

Specifically, this study would address the following research questions:

- Can the "branching" item type be developed and delivered in a logistically efficient way for use in NAEP?
- Can the information from sets of "branching" item bundles be used to provide reliable, valid, and useful information on both student overall ability in science and the classification of students into educationally-useful categories?

The study would focus on a specific Earth system that is of practical and environmental significance such as the biogeochemical carbon cycle. Understanding this system and related environmental issues (e.g., global climate change) requires connected understandings in the Physical, Life, and Earth and Space Sciences, many of which are characterized in the *Framework* and *Specifications*. For example, students who understand global warming understand how photosynthesis, cellular respiration, and fossil fuel combustion affect carbon dioxide concentrations in the atmosphere. This connected understanding can be tracked as a learning progression.

APPENDIX A

NAEP SCIENCE STEERING COMMITTEE, PLANNING COMMITTEE, PROJECT STAFF, AND CONTRIBUTING GROUPS

Steering Committee

Gerald Wheeler, Chair
Executive Director
National Science Teachers Association
Arlington, VA

Mary Yakimowski, Co-Chair*
Director of Assessment
Council of Chief State School Officers
Washington, DC

Vance R. Ablott
Executive Director
Triangle Coalition for Science and Technology Education
Arlington, VA

Eileen M. Ahearn
Project Director
National Association of State Directors of Special Education
Alexandria, VA

Margaret A. Bartz
Council of the Great City Schools (representative)
Assessment Specialist, Administrator
Chicago Public Schools
Chicago, IL

Tim Beagley
National Association of State Boards of Education (representative)
Member, District 6
Utah State Board of Education
West Valley, UT

Ann Elizabeth Benbow
Director of Outreach, Education & Development
American Geological Institute
Alexandria, VA

E. Jane Gallucci
Secretary/Treasurer
National School Boards Association
Clearwater, FL

Yolanda Scott George
American Institute of Biological Sciences (representative)
Deputy Director and Program Director
American Association for the Advancement of Science
Washington, DC

Lisa Brady Gill
Partnership for 21st Century Skills (representative)
Director, Office of Education Policy
Educational and Productivity Solutions Texas Instruments, Inc.
Dallas, TX

Heidi Glidden
Senior Associate
American Federation of Teachers
Washington, DC

Gary L. Heath
Assistant State Superintendent Accountability and Assessment
Maryland State Department of Education
Baltimore, MD

Jack G. Hehn
Director, Education
American Institute of Physics
College Park, MD

Paul Hickman
American Physical Society (representative)
Science Consultant
Andover, MA

* CCSSO's Rolf Blank co-chaired the last meeting of the Steering Committee.

Linda Hodge
President
National Parents and Teachers Association
Chicago, IL

Marilynn Messick Kelley
National Catholic Educational Association (representative)
Science Chair/Physics Teacher
Gonzaga College High School
Washington, DC

Ioannis N. Miaoulis
President and Director
Museum of Science, Boston
Boston, MA

Jacqueline S. Miller
Senior Scientist, Center for Science Education
Education Development Center, Inc.
Newton, MA

Jean Moon
Director, Board on Science Education
The National Academies, Center for Education
Washington, DC

Brett D. Moulding
Past President, Council of State Science Supervisors (representative)
Science Coordinator
Utah Office of Education
Salt Lake City, UT

Carlo Parravano
Executive Director
Merck Institute for Science Education
Rahway, NJ

Leslie Ann Pierce
National Association of Biology Teachers (representative)
Biology and Chemistry Teacher
Thomas A. Edison High School
Alexandria, VA

Andrea I. Prejean
Senior Professional Associate, Student Achievement
National Education Association
Washington, DC

Judith Richardson
Resident Practitioner for Assessment & Accountability
National Association of Secondary School Principals
Reston, VA

Cheryl Riggins-Newby
Associate Executive Director
National Association of Elementary School Principals
Alexandria, VA

Mary Lucy "Doc" Sennett
Association for Supervision and Curriculum Development (representative)
Secondary Science Teacher
Northwest Rankin High School (MS)
Belhaven College (MS)
Brandon, MS

Sally Goetz Shuler
Executive Director
National Science Resources Center
Washington, DC

Finbarr C. Sloane[**]
Program Director, Research, Evaluation
& Communication
National Science Foundation
Arlington, VA

Patrick Sean Smith
Senior Research Associate
Horizon Research, Inc.
Chapel Hill, NC

Conrad Stanitski
American Chemical Society
(representative)
Visiting Professor, Chemistry
Department
Franklin & Marshall College
Lancaster, PA

Angela Diane Wenger
Association of Science-Technology
Centers (representative)
Executive Vice President and Chief
Operating Officer
New Jersey Academy for Aquatic
Sciences
Camden, NJ

Anne S. Yates
Consultant, Center for Collaborative
Research & Education
DuPont Company
Wilmington, DE

[**] Larry Suter substituted for Finbarr Sloane at the third Steering Committee meeting.

Planning Committee

Richard J. Shavelson, Chair
Margaret Jacks Professor of Education and Senior Fellow, Stanford Institute for the Environment
Stanford University
Stanford, CA

Senta A. Raizen, Co-Chair
Director, National Center for Improving Science Education
WestEd
Arlington, VA

Charles W. (Andy) Anderson
Professor, Department of Teacher Education
Michigan State University
East Lansing, MI

Myron Atkin
Professor of Education (Emeritus), Center for Educational Research at Stanford (CERAS)
Stanford University
Stanford, CA

Rodger W. Bybee
Executive Director
Biological Sciences Curriculum Study
Colorado Springs, CO

Audrey Champagne
Professor, Department of Educational Theory and Practice and Department of Chemistry
University at Albany, State University of New York (SUNY)
Albany, NY

Nancy Chesley
Teacher
Mabel I. Wilson School
Cumberland Center, ME

Linda M. Crocker
Professor (Emeritus)
University of Florida
Hawthorne, FL

Richard Duschl
Professor, Graduate School of Education
Rutgers, the State University of New Jersey
New Brunswick, NJ

Janice Earle
Senior Program Director
National Science Foundation
Arlington, VA

Arthur Eisenkraft
Distinguished Professor of Science Education and Senior Research Fellow
University of Massachusetts, Boston
Quincy, MA

Jim Ellingson
Assistant Professor of Education
Concordia College
Moorhead, MN

Barbara A. Hopkins
Outreach Director, UNH Leitzel Center
University of New Hampshire
Durham, NH

Jane Butler Kahle
Condit Professor of Science Education, School of Education and Allied Professions
Miami University
Oxford, OH

Michele Lombard
Science Specialist
Arlington Public Schools
Arlington, VA

María Alicia López Freeman
Executive Director, California Science Project
University of California, Los Angeles
Los Angeles, CA

Carolee S. Matsumoto
Director of Research and Dissemination, Center for University, School and Community Partnerships
University of Massachusetts, Dartmouth
New Bedford, MA

Jim Minstrell
Research Scientist
FACET Innovations
Seattle, WA

Michael J. Padilla
President, National Science Teachers Association
Professor and Director for Educator Partnerships, College of Education
University of Georgia
Athens, GA

Jo Ellen Roseman
Director, Project 2061
American Association for the Advancement of Science
Washington, DC

Cary Sneider
Vice President for Educator Programs
Museum of Science, Boston
Boston, MA

Elizabeth K. Stage
Director, Lawrence Hall of Science
University of California, Berkeley
Berkeley, CA

Anne Tweed
Past President, National Science Teachers Association
Senior Science Consultant
Mid-continent Research for Education and Learning (McREL)
Aurora, CO

David J. White
Science Assessment Coordinator
Vermont Department of Education
Berlin, VT

Mark Wilson
Professor, Graduate School of Education
University of California, Berkeley
Berkeley, CA

NAGB Staff

Mary Crovo
Assistant Director for Test Development
National Assessment Governing Board
Washington, DC

Sharif Shakrani
Deputy Executive Director and Contract Officer
National Assessment Governing Board
Washington, DC

Charles Smith
Executive Director
National Assessment Governing Board
Washington, DC

WestEd Staff

Alice Fu
Science Specialist, National Center for Improving Science Education
WestEd
Arlington, VA

Joyce Kaser
Senior Program Associate
WestEd
Albuquerque, NM

Susan Mundry
Associate Director, Mathematics, Science, and Technology Programs
WestEd
Newburyport, MA

Senta Raizen
Director, National Center for Improving Science Education
WestEd
Arlington, VA

Steve Schneider
Director, Mathematics, Science, and Technology Programs
WestEd
Redwood City, CA

Mike Timms
Senior Research Associate, Mathematics, Science, and Technology Programs
WestEd
Oakland, CA

CCSSO Staff

Rolf Blank
Director of Education Indicators Programs
Council of Chief State School Officers
Washington, DC

Arthur Halbrook
Senior Associate
Council of Chief State School Officers
Washington, DC

Julia Lara
Deputy Executive Director of State Services and Technical Assistance
Council of Chief State School Officers
Washington, DC

Mary Yakimowski
Director of Assessment
Council of Chief State School Officers
Washington, DC

Contributing Groups

In developing the *Framework* and *Specifications*, the National Assessment Governing Board (NAGB) benefited from the extraordinary efforts of hundreds of individuals and organizations across the nation. Although these individuals and organizations provided valuable comments and feedback on draft documents, they were not asked to endorse the final versions of the *Framework* and *Specifications*. NAGB wishes to acknowledge their contributions by listing them below. While every effort has been made to ensure the accuracy and comprehensiveness of this list, we apologize for any omissions or errors.

External Framework Review Panel

Penny J. Gilmer, Professor of Chemistry & Biochemistry, Florida State University
Eric Pyle, Associate Professor of Geology & Environmental Sciences, James Madison University
Gail Richmond, Associate Professor of Biology & Science Education, Michigan State University

Presentations and/or Feedback Sessions

American Association for the Advancement of Science
- Section Q meeting

Center for Assessment and Evaluation of Student Learning
- Annual conference

Council of Chief State School Officers
- Education Information Management Advisory Consortium (EIMAC)
- National Conference on Large-scale Assessment
- Mega-SCASS Conference

Council of State Science Supervisors
- Thirteen regional feedback meetings with at least 368 total participants representing 44 states, DC, the U.S. Department of Education, and the Department of Defense
- National feedback meeting
- Annual meeting

InterAcademy Panel (I.A.P.)
- Workshop on the Evaluation of Inquiry-Based Science Education Programs

NAEP State Coordinators and State Science Supervisors
- Three online WebEx sessions
- Poll on technological design

National Research Council
- Board on Science Education
- Committee on Science Learning K-8

National Science Teachers Association
- National and regional conventions
- Online survey with 1,769 total responses, representing all 50 states and DC

Smithsonian Institution

- Presentation to education staff

Triangle Coalition for Science and Technology Education

- Annual conference

U.S.-China Education Leaders Forum on Math and Science Education

Testimony at NAGB Public Hearing, October 2005, Washington, DC

Achieve (represented by Jean Slattery)
American Association for the Advancement of Science (represented by Alan Leshner)
Intel Foundation (represented by Wendy Hawkins)
International Technology Education Association (represented by Kendall Starkweather)
Museum of Science, Boston (represented by Patti Curtis)
National Academy of Engineering (represented by William Wulf)
National Science Teachers Association (represented by Gerald Wheeler)
Thomas B. Fordham Foundation (represented by Michael Petrilli)

NAGB Invitational Science and Industry Discussion Forum, May 2005, Atlanta, GA

Lela Blackburn, Atlanta Public Schools
Pete Briggs, Scientific Atlanta
Robert DeHaan, Emory University
Victor De Jesus, Georgia Tech Research Institute
Maynard Hammond, Scientific Atlanta
Stewart Holm, Georgia Pacific
Lamar May, Lockheed Martin (retired)
Seth Marder, Georgia Tech
Tom McKlin, Center for Education Integrating Science, Mathematics, and Computing (CEISMC)
Chris Miller, Cytec Surface Specialties
Bob Moore, Southern Regional Education Board
Paul Ohme, CEISMC
Rick Paulus, EMS Technologies
Randy Rowe, GE Energy
Ken Stefanini, Georgia Power
Marion Usselman, CEISMC
Donna Whiting, CEISMC

In addition, NAGB received feedback on the draft *Framework* and *Specifications* via hundreds of individual letters and e-mails. The authors of these letters and e-mails included scientists; teachers; university administrators and professors; education researchers; curriculum developers; assessment and measurement specialists; as well as representatives from science industry, museums and science centers, community interest groups, professional organizations, and government agencies at the federal, state, and district levels.

APPENDIX B

STEERING COMMITTEE GUIDELINES

The *Framework* is Informed by the *National Standards* and *Benchmarks*.

The *Framework* should reflect the nation's best thinking in science instruction and thus be guided by two national documents: *National Science Education Standards* (NRC, 1996) and *Benchmarks for Science Literacy* (AAAS, 1993). Both of these documents were subject to extensive internal and external reviews during their development.

Informed by the *National Standards* and *Benchmarks*, the *Framework* should emphasize knowledge and use of science concepts, appropriate linking of science facts to concepts, relationships among concepts, and major themes unifying the sciences. The *Framework* should also incorporate investigative skills.

The *Framework* Reflects the Nature and Practice of Science.

The *National Standards* and *Benchmarks* include standards addressing science as inquiry, nature of science, history of science, and the designed world. The *Framework* should emphasize the importance of these aspects of science education and should include the expectation that students will understand the nature and practice of science. Science is a self-correcting process, a way of knowing where theories are continually modified and refined based on new research findings. Students should demonstrate the ability to accomplish the following:

- Make warranted inferences from evidence
- Use evidence to justify conclusions based on scientific investigations
- Demonstrate reasoning skills in the application of science content and in understanding the connections between science concepts
- Exercise skepticism when evaluating, using, and discarding data
- Understand and use models to describe and do science
- Apply content knowledge and skills to solve problems as they occur in the natural world
- Understand and apply knowledge of links and commonalities of science across fields

The scientific disciplines are no longer practiced in isolation, and research that cuts across discipline boundaries is common. The *Framework* should

- Identify some of the science concepts and skills that cut across the assessed content areas;
- Address science in both the natural and designed world; and
- Clearly define and identify commonalities and differences between "science" and "technology" or "technological design."

The *Framework* should also address social and historical contexts, which are keys to understanding how the scientific community has arrived at its current body of knowledge.

The *Framework* Incorporates Key Attributes of Effective Assessment.

The *Framework* should use assessment formats that are consistent with the objectives being assessed. It should be guided by the best available research on assessment item design and delivery.

The *Framework* should be inclusive of student diversity as reflected in gender, geographic location, language proficiency, race/ethnicity, socio-economic status, and disability condition. The assessment should be designed and written to be accessible by the majority of students and to minimize the need for special accommodations for both students with disabilities and English language learners. Students with special needs should be provided accommodations to allow them to participate in the assessment.

The *Framework* should reflect knowledge about the acquisition of key science concepts over time, based on research about how students learn. The existing research findings should make clear, when possible, what the progression of science knowledge looks like across the grade levels. Concepts should be represented in a manner that reflects how students progress through a discipline and across disciplines. Assessment items should reflect students' potential for applying concepts and more varied and complex situations over time.

Critical content and skills should be articulated and assessed across grades 4, 8, and 12 (vertically), as well as across the fields of science (horizontally), by creating items that are deliberately layered to achieve these goals. An example of measuring similar constructs within and across subjects is the progression of increasingly sophisticated understanding about energy from elementary to middle to high school in the content areas of biology, chemistry, Earth science, and physics.

A variety of assessment formats should continue to be used in the NAEP assessment, including well-constructed multiple-choice and open-ended items as well as performance tasks. In addition, multiple methods of assessment delivery should be considered, including the appropriate uses of digital-based technology. The *Framework* should consider use of digital delivery systems for the assessment including Web-based or CD-ROM formats. The use of embedded simulations that can represent scientific phenomena such as data, representations, and factors captured within laboratory experiments and use of an adaptively designed series of assessment items should also be considered. Advances in machine scoring of text should provide the opportunity for increased use of open-response format questions. The assessment format and delivery system employed should offer accessibility to the widest range of students.

Each achievement level—Basic, Proficient, and Advanced—should include a range of items assessing various levels of cognitive knowledge that is broad enough to ensure each is measured with the same degree of accuracy. Descriptions of Basic, Proficient, and Advanced must be as clear as possible.

The Assessment Provides Data for Research.

NAEP assessment results are increasingly being used to review state student assessments and compare student achievement across states. The *Framework* should address the important uses of assessment data both to conduct research to better understand science learning and to improve science achievement. Data from the assessment should be collected in such a way as to provide information that accomplishes the following:

- Supplies details of the attributes (e.g., race/ethnicity, gender) of the students being assessed;
- Provides results by student gender, race/ethnicity, and socio-economic level;
- Describes the academic preparation of the teachers of the students being assessed;
- Describes the nature of the educational system of the students being assessed;
- Relates the instructional delivery and materials, professional development of the teachers, and the learning environment to the results from assessment; and
- Provides feedback to educators for improving science instruction and learning.

The *Specifications* Document is Closely Aligned with the *Framework*.

The connections among the *Framework*, the *Specifications*, and the assessment items themselves, should be transparent, have a consistent level of specificity, and be coherent.

The *Specifications* should be written with consistent detail across all fields, domains, and expectations of the *Framework*:

- The *Specifications* should have a consistent structure across all areas.
- The expected science knowledge that represents the target for assessment should be described in a clear and consistent format. The content addressed in the *Specifications* should reflect the standards and focus on the significant information and knowledge that students should retain (e.g., big ideas, fundamental understandings) over time.
- The verbs used in the *Specifications* should describe the expected action to be taken in the assessment (e.g., identify, describe, evaluate, relate, analyze, demonstrate).
- Expectations across the content areas should match in level of specificity and scope.
- The *Specifications* should follow the idea of learning trajectories. To assess overarching concepts or themes, the assessment specifications should reflect a scaffolded or layered understanding of growth in knowledge of the concepts.

APPENDIX C

NAEP SCIENCE
PRELIMINARY ACHIEVEMENT LEVEL DESCRIPTIONS

Table 24. Grade 4 Preliminary Achievement Level Descriptions

	BASIC	PROFICIENT	ADVANCED
	Identifying Science Principles	**Identifying Science Principles**	**Identifying Science Principles**
Physical Science	Compare properties of solids and liquids	Describe the changes in physical properties that result when substances are heated and cooled	Identify properties of gases, liquids, and solids (e.g., whether made of one or more substances)
Life Science	Identify the foods that animals eat as sources of energy and building blocks for growth and repair	Describe life cycles of familiar plants and animals	Relate an organism's survival with conditions in the environment that meet the organism's basic needs
Earth and Space Science	Describe changes in the apparent shape of the moon over a month's time	Describe the change in the apparent path of the sun, as viewed in the Northern Hemisphere between June and December	Relate the changes in the location of sunrise with time of year
	Using Science Principles	**Using Science Principles**	**Using Science Principles**
Physical Science	Describe what happens when a given amount of liquid is poured from a tall and narrow container into a broad and shallow container	Explain why a metal container filled with a hot liquid feels hotter to the touch than a foam container filled with the same hot liquid	Explain what happens to the gas in a helium balloon when it is punctured in a closed room
Life Science	Describe how familiar animals meet their basic needs for food, air, water, and shelter	Predict how a change in a plant's environment will affect the plant's survival	Explain why animals need food and plants need nutrients
Earth and Space Science	Explain why some rivers become muddy during or after heavy rains	Provide three examples of weathering and erosion	Relate periods of erosion to weather events or seasons

Table 24. Grade 4 Preliminary Achievement Level Descriptions (cont.)

	BASIC	PROFICIENT	ADVANCED
	Using Scientific Inquiry	**Using Scientific Inquiry**	**Using Scientific Inquiry**
Physical Science	Order magnets by strength based on data describing the number of objects (e.g., paper clips) attracted to each of the magnets	Critique several proposed investigations comparing the heat produced by burning different quantities of wax	Design an investigation to demonstrate the relationship between the length of a vibrating string and the waves produced
Life Science	Construct a bar graph from numerical data showing changes in the height of a plant over time	Select the best designed investigation from descriptions of several different ways to investigate the effects of light intensity on a plant	Design an investigation to demonstrate how a change in an environment changes the number of a familiar kind of animal in the environment
Earth and Space Science	Identify a thermometer as a tool for measuring temperature	Construct a bar graph from data showing average monthly temperatures over a twelve month time period	Design an investigation to relate weather data to the changes in seasons
	Using Technological Design	**Using Technological Design**	**Using Technological Design**
Physical Science	Choose materials that are identified as good thermal insulators for use in containers to keep hot beverages from cooling off	Propose a method for testing materials to determine which are the best thermal insulators	Apply information about the thermal properties of materials to design a beverage container which keeps hot beverages hot and cold beverages cold for as long as possible
Life Science	Select from three drawings of shelters for an animal, the one that will meet most of the animal's basic needs	Apply information about an animal's basic needs to design a shelter for the animal	Analyze and explain how the design of a shelter or den, as constructed by an animal, meets its needs
Earth and Space Science	Explain how to use the motion of the sun to determine if it is morning or afternoon	Choose from among three different drawings the best method of using the motion of the sun to tell the approximate time of day	Explain how to safely observe and record the motion of the sun during the day so as to construct a means of telling the approximate time of day

Table 25. Grade 8 Preliminary Achievement Level Descriptions

	BASIC	PROFICIENT	ADVANCED
	Identifying Science Principles	**Identifying Science Principles**	**Identifying Science Principles**
Physical Science	Recognize that acids are a class of compounds that exhibit common chemical properties	Describe properties of acids such as color changes with acid/base indicators and the tendency to react with bases	Relate the properties of elements that form acids with their position in the Periodic Table
Life Science	Identify producers and consumers as components of living systems	Describe the functions of consumers in ecological systems	Relate the functions of consumers to the energy flow in ecological systems
Earth and Space Science	Describe the location of Earth in the solar system	Relate the phases of the moon and the length of a day and a year to the motions of Earth, the moon, and the sun	Relate the force of gravity to the regular motion of bodies in the solar system
	Using Science Principles	**Using Science Principles**	**Using Science Principles**
Physical Science	Explain the physical properties of solids, liquids, and gases, using the idea that matter is composed of tiny particles in motion	Explain chemical properties of metals using the structure of atoms	Predict the properties of an element based on its position in the Periodic Table
Life Science	Give examples of producers and consumers in aquatic ecosystems	Predict the effect of a reduction in the population of predator species on the population of a species on which it preys	Evaluate an alternative explanation for patterns observed in an ecosystem's population data
Earth and Space Science	Predict the effect of a reduction of the amount of light from the sun reaching Earth	Explain how the tilt of Earth's rotation axis produces annual variation in the intensity of sunlight on the Earth's surface	Predict how changes in axial tilt would affect annual variations in the intensity of sunlight on the Earth's surface at different latitudes

Table 25. Grade 8 Preliminary Achievement Level Descriptions (cont.)

	BASIC	PROFICIENT	ADVANCED
	Using Scientific Inquiry	**Using Scientific Inquiry**	**Using Scientific Inquiry**
Physical Science	Use physical and chemical properties to classify substances as metals or non-metals	Select the best designed investigation from descriptions of several different ways to demonstrate that water does not change chemically when it changes state	Design an investigation to measure the temperature of water when it changes state (phase diagrams)
Life Science	Describe simple patterns in population data	Conduct a survey of an ecosystem's population data and propose an explanation for patterns observed in the data	Design a survey of an ecosystem's consumer populations and propose an explanation for patterns in the data based on energy flow through the system
Earth and Space Science	Apply a framework for analyzing the constituent minerals and texture of a rock	Analyze properties of a rock formation to draw valid conclusions about the conditions under which it was formed	Propose a framework for classifying rocks, based on constituent minerals, texture, and environment of formation
	Using Technological Design	**Using Technological Design**	**Using Technological Design**
Physical Science	Interpret a distance versus time graph of two different types of parachutes (showing how far they have fallen in equal time intervals) to determine which is better at slowing the jumper's fall	Construct a graph based on a table showing how far two different types of parachutes have fallen in equal time intervals. Determine which parachute is better at slowing the jumper's fall and describe the evidence that supports that determination	Design a method for measuring and comparing the effectiveness of two new types of parachutes to see which is best at slowing the fall of a jumper
Life Science	Apply information about consumers and producers to critique the design of a self-sustaining terrarium	Apply information about consumers and producers to design a self-sustaining terrarium	Apply information about energy flow to critique a plan for managing the deer population in a forest ecosystem located near fields where corn is grown
Earth and Space Science	Identify the possible effects on the environment of cutting down a forested area to establish a new field to grow corn	List the scientific trade-offs that need to be taken into account when cutting down a forested area to establish a new field to grow corn	Taking into account the scientific trade-offs involved in converting forested lands to fields for growing corn, propose what can be done to reduce some of the negative environmental impacts while allowing for increased food production

Table 26. Grade 12 Preliminary Achievement Level Descriptions

	BASIC	PROFICIENT	ADVANCED
	Identifying Science Principles	**Identifying Science Principles**	**Identifying Science Principles**
Physical Science	State the difference between exothermic and endothermic chemical reactions	Describe the increases in kinetic, rotational, and vibrational energy that result when a substance is heated	Explain the melting of a solid crystal in terms of the energy of the constituent molecules, ions, or atoms
Life Science	State modern scientific ideas about evolution such as natural selection and common descent	Describe fossil, anatomical, and molecular evidence for biological evolution	Make connections among the following related science principles: the potential of a species to increase its numbers; the genetic variability of its offspring; limitations on the resources required for life; and the ensuing selection of those organisms better able to survive and leave offspring
Earth and Space Science	State the law of superposition in an undisturbed sequence of rock layers—that younger rock layers sit atop older rock layers below	Classify some geological processes as happening on a human time scale, such as earthquakes, and other processes occurring on a geological time scale, such as mountain building	Use index fossils and type sections to assign sequences of rocks to geological eras
	Using Science Principles	**Using Science Principles**	**Using Science Principles**
Physical Science	Give an example of an element that has an isotope	Predict some common chemical reactions, given a choice of reactants (e.g., metals and non-metals, acids and bases)	Explain the difference between ionic and covalent bonding
Life Science	Use antibiotic resistance as an example of principles of biological evolution	Predict the spread of infectious disease based on basic concepts of evolution	Use basic concepts of evolution to explain antibiotic resistance and invasive species
Earth and Space Science	Explain the process of mountain building using the theory of plate tectonics	Explain the location of deep sea trenches as an outcome of geologic processes	Propose geologic processes that explain structures found on a geologic map

Table 26. Grade 12 Preliminary Achievement Level Descriptions (cont.)

	BASIC	PROFICIENT	ADVANCED
	Using Scientific Inquiry	**Using Scientific Inquiry**	**Using Scientific Inquiry**
Physical Science	Describe patterns of physical and chemical properties within rows of the Periodic Table	Select the best-designed demonstration among descriptions of how to illustrate the translational, vibrational, and rotational motion of molecules	Design an investigation to determine the effect of surface area on evaporation rate
Life Science	Use information contained in a food web to illustrate energy conservation in an ecosystem	Select the best-designed investigation among descriptions of ways to determine the effect of fertilizers on the growth of plants	Design a strategy for estimating the quantity of plant material required to produce a kilogram of beef
Earth and Space Science	Analyze photographs of rock layers to determine the order in which the layers were deposited	Describe how radioactive dating is used to estimate the age of rock formations and sequence data to provide relative age information	Integrate fossil, stratigraphic, structural, and rock-type information to identify past spatial relationships between environments
	Using Technological Design	**Using Technological Design**	**Using Technological Design**
Physical Science	A railway company wants to purchase the fastest bullet train on the market. Select the fastest train based on data about distance traveled by each of two trains and the time it took to travel that far	A railway company wants to purchase the fastest bullet train on the market. Design a test of two trains (consider what is to be measured, instruments to be used, and how data are to be collected)	A railway company wants to purchase the fastest bullet train on the market. List science-based criteria and constraints that might be used to select the best train and propose a method to compare two trains on these criteria and constraints
Life Science	Given a description of plants and animals in a stable ecosystem, identify from a list of possibilities what changes might occur if people migrate into the area	Given a description of plants and animals in a stable ecosystem, list some of the changes that might occur if people migrate into the area; and describe how those changes could affect the entire food web	Given a description of plants and animals in a stable ecosystem, list some of the changes that might occur if people migrate into the area; and describe what the new residents could do to limit those changes and avoid affecting the entire food web
Earth and Space Science	Given data on increasing levels of carbon dioxide in the atmosphere, select from a list some of the anticipated effects of this carbon dioxide buildup	Given data on increasing levels of carbon dioxide in the atmosphere, list both positive and negative effects of this carbon dioxide buildup	Given data on increasing levels of carbon dioxide in the atmosphere, list positive and negative effects of this buildup and science-based actions people might take to reduce or prepare for any negative effects

APPENDIX D

SAMPLE ITEMS AND SCORING GUIDES

Table 27 lists all of the illustrative items that appear in the *Specifications*. Provided in the table for each item are page number location(s) in the *Specifications*, science content statement designation, science practice designation, and source. The original identifying code is also provided for all TIMSS items. Finally, a "Yes" in the "Appendix D?" column indicates that a scoring rubric or additional description for the item can be found in this appendix; specific page number locations of these rubrics and descriptions are provided in parentheses.

Table 27. Illustrative Items Appearing in the *Specifications*

Page	Question	Content	Practice	Source	Identifier	Appendix D?
113	The Earth's Moon is…	E8.1	Identifying Science Principles	NAEP 2000, Grade 8		
114	Animals and plants are made…	L8.5	Identifying Science Principles	TIMSS 2003, Grade 8	S032682	
115	Look at the food web above…	L8.6	Using Science Principles	TIMSS 1999, Grade 8	N02	Yes (251-252)
116	The pictures show a lightbulb…	P4.11	Using Science Principles	TIMSS 2003, Grade 4	S031038	
118	The graph below shows the…	P8.14	Items 7 and 8: Identifying Science Principles Using Science Principles	Colorado Department of Education, 2002, Grade 8		Yes (253)
121	(Lynx and Hare ICT)	L8.6	Using Scientific Inquiry	Quellmalz, Griffin, Hurst, Kreikemeier, Rosenquist, and Zalles (2004)		Yes (254)
125	Occasionally, a fire will…	E8.15	Using Technological Design	Washington Assessment of Student Learning, 2004, Grade 8		Yes (255-258)
130	What digestive substance is…	L8.1	Identifying Science Principles	TIMSS 1995, Grade 8	P6	Yes (259)
131	You have three bags of white…	P8.4	Using Scientific Inquiry	Ruiz-Primo, Shavelson, Hamilton, and Klein (2002)		Yes (259-262)
132	The graph shows the progress…	P8.14	Using Science Principles	TIMSS 1995, Grade 8	P1	
134	A rocket is moving along…	P12.19	Using Science Principles	Clement (1982)		
135	Which statement explains…	E8.2	Using Science Principles	TIMSS 1995, Grade 8	Q11	
148, 320	What force keeps the planets…	E8.2	Identifying Science Principles	Adapted from Massachusetts Department of Education, MCAS, 2000, Grade 8		

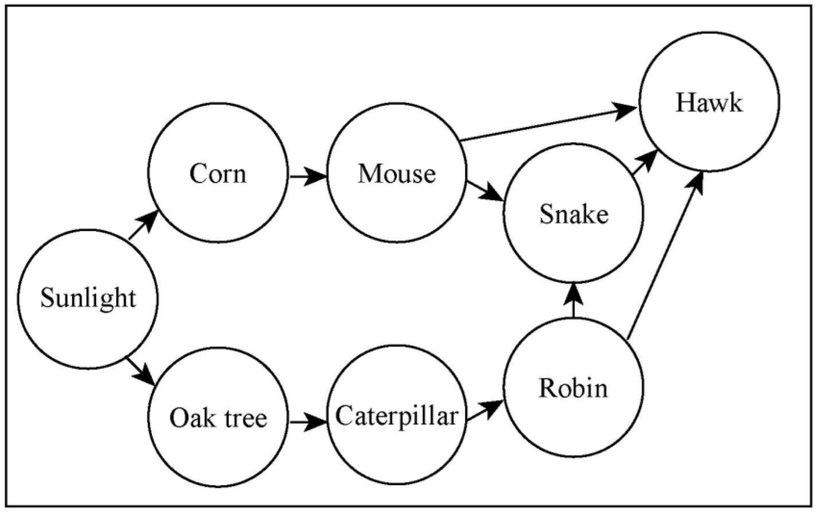
Table 27. Illustrative Items Appearing in the *Specifications* (cont.)

Page	Question	Content	Practice	Source	Identifier	Appendix D?
148, 321	The drawings show a rocket…	E8.2	Identifying Science Principles	TIMSS 1999, Grade 8	J05	
149, 321	A space station is to be…	E8.2	Using Science Principles	NAEP 1996, Grade 8		Yes (262-263)
162	Air is made up of many…	E8.7	Identifying Science Principles	TIMSS 1995, Grade 8	O12	
162	The diagram above shows a…	E8.12	Identifying Science Principles	TIMSS 2003, Grade 8	S032652	
166	On a hot, humid day the air…	P4.6	Identifying Science Principles	TIMSS 2003, Grade 4	S031382	Yes (264)
166	When a population of mice is…	L8.6	Using Technological Design	NAEP 1996, Grade 8		Yes (265)
166	One day when the temperature…	P8.6	Using Science Principles	TIMSS 1995, Grade 8	Y2	Yes (266)
167	The biosphere (living…	E12.12	Using Science Principles	New Standards Spring Field Test 1999 for High School		Yes (267)
168	Look at the picture of a roller…	P8.16 (a), P8.12 (b and c)	Identifying Science Principles (a and b), Using Science Principles (c)	Colorado State Assessment Program, 2001, Grade 8 Science		Yes (268)
175	What causes day and night? The main reason for its…	E8.2, E8.12	Using Science Principles	Sadler (1998, pp. 274-276)		
177	Rich cut block A into two…	P8.4	Using Science Principles	Adapted from Shavelson (2006) and Shavelson et al. (2005)		
178	Here is a cross-section of a…	P12.1	Using Science Principles	TIMSS 1995, Grade 12	H14	Yes (269)
179	Anna and Uri had identical…	P4.2	Using Science Principles	TIMSS 1995, Grade 4	X1	Yes (270)
182	(Electric Mysteries hands-on performance task)	P4.11	Using Scientific Inquiry	Shavelson et al. (1991)		Yes (271)
182	(Sow Bugs task)	L4.2	Using Scientific Inquiry	Shavelson et al. (1991); Shavelson, Gao, and Baxter (1995)		Yes (272)
188	(NAEP TRE Online ICT)	P8.16	Using Scientific Inquiry	Persky et al. (2005)		Yes (273)

Table 27. Illustrative Items Appearing in the *Specifications* (cont.)

Page	Question	Content	Practice	Source	Identifier	Appendix D?
189	(Electric Mysteries ICT)	P4.11	Using Scientific Inquiry	Shavelson et al. (1991, p. 357)		Yes (271)
190	(Solar Power ICT)	E12.10	Using Technological Design	Quellmalz et al. (2004)		Yes (273-274)
191	(Concept Mapper ICT)	L12.5	Identifying Science Principles	Adapted from Herl et al. (1999)		Yes (274)
284	Which of the boxes X, Y, or Z…	P4.1	Identifying Science Principles	Adapted from TIMSS 1995, Grade 4	N9	
284	The picture below shows two…	P4.1	Using Science Principles	Adapted from WestEd, RISSA, 2001		
286	When most substances change…	P12.12	Identifying Science Principles	WestEd, 2005		
286	The random molecular motion…	P12.1	Identifying Science Principles	California Department of Education, California Standards Test, 2004, Grade 12 Chemistry		
287	Which of the following graphs…	P12.5	Identifying Science Principles	Adapted from NAEP 1996, Grade 12		
287	If the attractive forces among…	P12.1	Using Science Principles	California Department of Education, California Standards Test, 2004, Grade 12 Chemistry		
288	Which of the following is…	P4.6	Identifying Science Principles	Massachusetts Department of Education, MCAS, 2002, Grade 5		
289	What change of state is shown?	P4.6	Identifying Science Principles	Oregon Department of Education, 2003-2005 Sample Test, Grade 5		
314	Animals usually have physical…	L4.7	Identifying Science Principles	TIMSS 2003, Grade 4	S031284	
314	A girl found the skull of...	L4.7	Using Science Principles	Illinois Science Sample Test, 2006, Grade 4		
314	The Pacific Tree Frog lives in…	L4.7	Using Science Principles	Adapted from Washington Assessment of Student Learning, 2004, Grade 5		Yes (275-277)
332	The burning of fossil fuels has…	E8.15	Identifying Science Principles	TIMSS 2003, Grade 8	S012017	

Item Source: TIMSS 1999, Grade 8 (*Specifications*, p. 115)



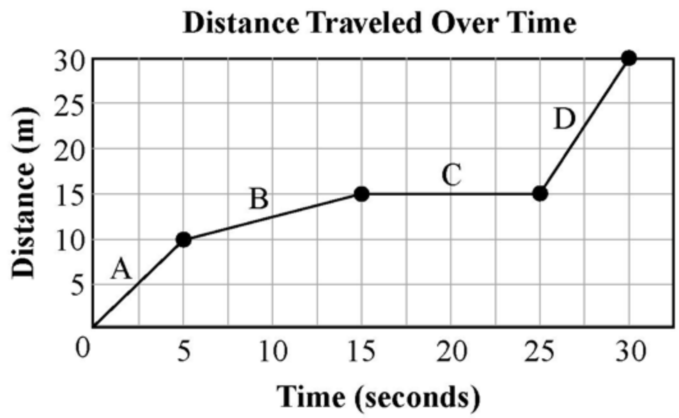
Look at the food web above. If the corn crop failed one year what would most likely happen to the robin population? Explain your answer.

Note: A correct response must include a feasible explanation directly relating the predicted change in robin population to the effect of corn crop failure on prey/predator relationships indicated in the food web. Responses do not have to use the specific terms **decrease**, **increase**, and **same**, as long as the explanation is clear with respect to the effect on the robin population. If more than one effect is given, assign the code corresponding to the first correct explanation.

Code	Response	Item: S022141
	Correct Response	
10	Robin population may **decrease**. Explanation based on predators (snakes/hawks) eating more robins if mice die. *Examples:* *Goes down. The mice would starve and the snake would eat the robins.* *There would be less robins because the mouse population would decrease and the snakes (and/or hawks) would eat more of the robins.*	
11	Robin population may **increase**. Explanation based on predators (snakes/hawks) dying due to lack of food (mice). *Examples:* *It would go up because the snakes die if the mouse starves.* *There could be more robins because there are fewer snakes (and/or hawks) to eat them.*	
12	Robin population would stay the **same** with a feasible explanation. *Example:* *It would not change because the mouse would find other grain to eat so the snake would be unaffected.*	
19	Other acceptable explanation.	
	Incorrect Response	
70	Robin population would **decrease**. Incorrect explanation based on robins starving if snakes die (confuses prey/predator relationship). *Examples:* *Decreases because there are less snakes to eat.* *When corn dies, then snakes, then robins starve.*	
71	Robin population would **decrease**. Incorrect explanation based on the robin needing corn to survive. *Example:* *Decrease because they need the corn.*	
72	Robin population would stay the **same.** Incorrect explanation based on the robins not needing corn to survive or not being connected to corn in the food web. (Does not consider the effect of predators.) *Examples:* *Nothing because the robin only eats insects.* *Nothing would happen. The corn is on a different chain in the food web.*	
73	Mentions **only** that the whole food web will be upset and/or all the animals will die. *Example:* *The whole food web would erupt and everything would die.*	
79	Other incorrect (including crossed out/erased, stray marks, illegible, or off task).	
	Nonresponse	
99	BLANK	

Item Source: Colorado Department of Education, 2002, Grade 8
(*Specifications*, p. 118)

The graph below shows the distance traveled over time by a student walking down a hall. Use the information shown on the graph to do Numbers 7 and 8.



7. During which time interval was the student moving the fastest?

   ○ A

   ○ B

   ○ C

   ○ D

Key: D

8. What was the average speed of the student from 0 seconds to 5 seconds?

   Average speed: ___________________________________

Scoring: Average speed is 2 m/s. (The student traveled 10 meters in 5 seconds.)

Item Source: Quellmalz et al., 2004 (*Specifications*, p. 121)

Lynx/Hare Task

This is an interactive computer task in which students are expected to conduct a scientific investigation regarding the question of whether or not lynx should be introduced into a national park in order to reduce the abiding overpopulation of hares. Students are directed to complete six modules, which make use of different computer programs in order to determine the best solution for the proposed question:

Module 1 asks the student to access, organize, analyze, and interpret data that they are given about the populations of hares over the past four years, using Word processor, Spreadsheet, or Presentation software.

Module 2 asks the student to determine a better way to analyze and display some disorganized data that show how many lynx and hares were present each year over the past 25 years.

Module 3 first asks the student to submit a web search that will give insight into the relationship between lynx and hare populations. It subsequently asks the student to critically evaluate the relevance of several given web searches.

Module 4 asks the student to collect information on specific questions regarding the lynx/hare question, to take notes on the information given on the web sites, and to include citations for each site.

Module 5 asks the student to use a modeling program to predict the results of adding more lynx to the parks through viewing population trends over several years. Students are then asked questions based on what they have observed in the modeling tool regarding increases or decreases in the hare population if lynx are or are not added to the park.

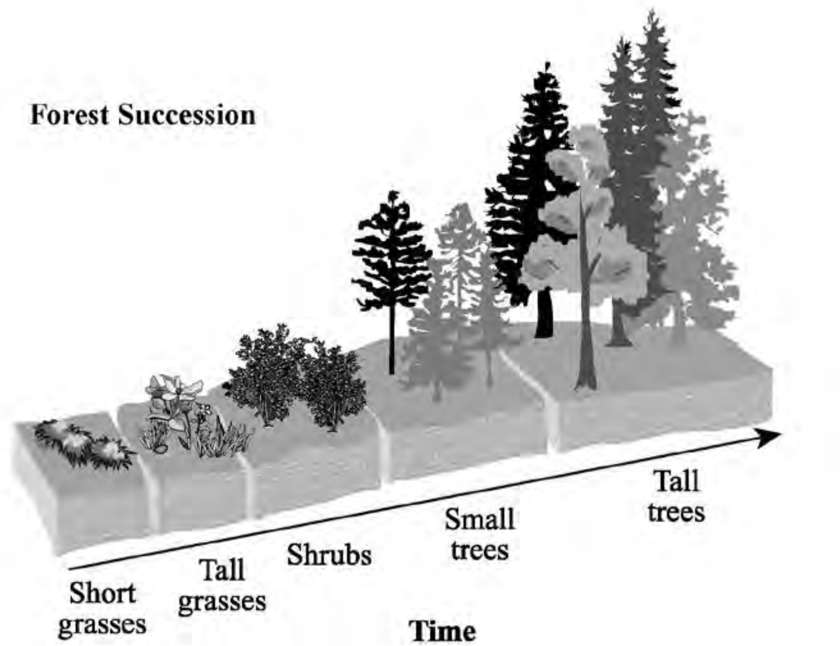
Module 6 asks the student to create a presentation with Word processor or Presentation software in order to communicate the problem, the findings, and any recommendations that resulted from the newly completed research.

Scores for this task are given for inquiry skills and technology use, along with the appropriate use of concepts within their explanations and recommendations.

For more information, see http://ipat.sri.com/tasks/pred_prey/subtasks/taskstud.html

Item Source: Washington Assessment of Student Learning, 2004, Grade 8 (*Specifications*, p. 125)

Occasionally, a fire will destroy a forest, burning down trees and pushing wildlife out of their forest homes. However, the forest will grow back. Eventually, through the process of forest succession as shown below, short grasses and flowers begin to grow and animals make new homes.



Over time, shrubs and trees begin to grow. The forest returns to a lush habitat for the wildlife listed in the chart below.

Forest Wildlife

Ground-dwelling	Worms, beetles
Reptiles and amphibians	American toads, wood frogs, snakes, Eastern box turtles
Small animals	Squirrels, chipmunks
Medium to large animals	Opossums, raccoons, white-tailed deer, black bears
Airborne	Butterflies, moths, bees, wild turkeys, red-tailed hawks, bald eagles

A power company owns part of a forest that was destroyed by a fire. The forest could take decades to rebuild on its own. The company's department of environmental studies suggests planting new trees to help the forest rebuild.

Using the information in the scenario:

- Explain how planting trees could **benefit** the natural ecosystem.
- Explain how planting trees could **harm** the natural ecosystem.

Scoring Rubric

2-point response: The response demonstrates that the student can analyze how human societies' use of natural resources affects the quality of life and the health of ecosystems.

The student explains one reasonable way planting the trees could **benefit** the natural ecosystem.
AND
The student explains one reasonable way planting the trees could **harm** the natural ecosystem.

Example responses:

Benefits of planting the trees include, but are not limited to—

- providing a habitat for animals,
- providing a canopy, which would help to prevent soil erosion,
- creating root systems, which would anchor soil in place, and
- creating shade, which would help maintain sunlight levels and inhibit the introduction of nonnative plant species.

Harms of planting the trees include, but are not limited to—

- disrupting the natural flow of animals re-entering the forest,
- inhibiting the growth of other plants,
- decreasing the diversity of tree species growing in the forest, and
- introducing foreign species into an area, which may affect native species of plants and animals.

1-point response: The response demonstrates that the student can partially analyze how human societies' use of natural resources affects the quality of life and the health of ecosystems.

The student explains one reasonable way planting trees could **benefit** the environment.
OR
The student explains one reasonable way planting trees could **harm** the environment.

0-point response: The response demonstrates that the student can do little or no analysis of how human societies' use of natural resources affects the quality of life and the health of ecosystems.

Note: Benefits/harms to the natural ecosystem that only relate to humans shall not be credited score points.

Annotated Example of a 2-point response:

Explanation of how planting trees could benefit the natural ecosystem:
The ecosystem will rebuild more quickly, and that would give the
animals a new habitat more quickly.
Explanation of how planting trees could harm the natural ecosystem:
It could disrupt the natural order of the ecosystem. If some
smaller plants don't get a chance to grow first, the trees might
push them out.

Annotation:

The response demonstrates that the student can analyze how human societies' use of natural resources affects the quality of life and the health of ecosystems.

The student explains one reasonable way planting the trees could **benefit** the natural ecosystem, "The ecosystem will rebuild more quickly, and that would give the animals a new habitat more quickly." (1 point)

The student explains one reasonable way planting the trees could **harm** the natural ecosystem, "If some smaller plants don't get a chance to grow first, the trees might push them out." (1 point)

Annotated Example of a 1-point response:

Explanation of how planting trees could benefit the natural ecosystem:
More plants and trees would mean giving off more oxygen. more shade.
more food for forest animals.
Explanation of how planting trees could harm the natural ecosystem:
There could be a very big storm and trees fall an houses.

Annotation:

The response demonstrates that the student can partially analyze how human societies' use of natural resources affects the quality of life and the health of ecosystems.

The student explains one reasonable way planting the trees could **benefit** the natural ecosystem: "More food for forest animals." (1 point)

The student explains one way planting the trees could **harm** the human structure: "Trees falling on houses" but not the natural ecosystem. (0 points)

Annotated Example of a 0-point response:

Explanation of how planting trees could benefit the natural ecosystem:
have more trees
Explanation of how planting trees could harm the natural ecosystem:
half the plants will die

Annotation:

The response demonstrates that the student can do little to no analysis of how human societies' use of natural resources affects the quality of life and the health of ecosystems.

The student states one factor, "have more trees," but does not explain one reasonable way planting the trees could **benefit** the natural ecosystem. (0 points)

The student states a possible harm "half the plants will die" but does not explain how planting the trees could cause this **harm** to the natural ecosystem. (0 points)

Item Source: TIMSS 1995, Grade 8 (*Specifications*, p. 130)

What digestive substance is found in the mouth? What does it do?

Code	Response
Correct Response	
20	Names saliva and explains that it makes the food moist or soft [mechanical process]. *Example: Saliva. It helps us swallow.*
21	Names saliva and explains that it breaks down the starch or food. [Chemical process].
22	Names enzymes and explains that they break down the starch or food. [Chemical process].
29	Other correct: Names a substance and provides a reasonable explanation.
Partial Response	
10	Names saliva but with no description or with an incorrect description of what it does. *Example: Saliva. Contains acid which helps digesting the food.*
11	Names enzymes but with no description or with an incorrect description of what they do, such as it digests starch.
19	Other partially correct.
Incorrect Response	
70	Acid. With or without description.
71	Teeth, tongue, etc. With or without description.
79	Other incorrect.
Nonresponse	
90	Crossed out/erased, illegible, or impossible to interpret.
99	BLANK

Item Source: Ruiz-Primo et al., 2002 (*Specifications*, p. 131)

Saturated Solutions

Students are asked to find out which of three powders are the most and least soluble in 20 mL of water. Students are asked to provide information in assessment notebooks about how they conducted the investigation, the results they obtained, and how they dissolved the maximum possible amount of powder in a solution.

The following materials are provided: 3 labeled bags of white powder (baking soda, salt, and baking powder); 3 labeled plastic mixing cups; 1 measuring cup; 1 gram measuring spoon; 1 blue cup with water; 1 blue cup for liquid waste; 1 paper towel for spills; 1 tray with black paper; 6 wooden stirring sticks; and 1 pencil.

The following is an excerpt from the main task instructions given to students:

You have three bags of white powder in front of you:

- Bag A: baking soda
- Bag B: salt
- Bag C: baking powder

Find out how many grams of each powder are needed to saturate 20 mL of water. Of the three powders, which one do you have to use the most to saturate the liquid? Which one do you have to use the least?

In your investigation:

- You may use any of the materials in front of you.
- Be sure to slide a wooden stirring stick across the top of the 1 gram spoon each time you take a scoop of powder to make it level.
- Keep track of the number of scoops of powder you add to the water.
- Remember that one spoonful = one gram.

The following is an excerpt from the student response form:

Recording and Interpreting Your Results

Record the number of grams of powder that is required to saturate 20 mL of water in the chart below:

Baking Soda	Salt	Baking Powder

Of the three powders, which one did you have to use the most to saturate 20 mL of water? Which one did you have to use least?

Least Grams	Most Grams

List the powders in order of increasing solubility in water:

Least Soluble		Most Soluble

Describing your Investigation

Describe in detail what you did at each step of your investigation so one of your classmates can do exactly the same investigation you did to find out how many grams were needed to saturate 20 mL of water:

Steps in your investigation.
Please number each step in order.
Step
1. ____

[Blank lines are provided for the student to write a description of the investigation.]

Answering Some Questions

1. How could you tell when each solution was saturated?
2. What are some things you can do to help dissolve a powder when it is mixed with water?
3. Remember that you did your investigation with 20 ml. of water in each cup. How many grams of <u>salt</u> (Powder B) could you dissolve if you did your investigation with twice as much water (40 ml.) in each cup?

<u>The following is a brief description of the scoring system:</u>

The scoring system focuses on the accuracy of the results and the quality of the procedures used to solve the problem. For example, students are not given credit for "Care in Manipulation" if they do not write explicitly about carefully using only 20 mL of water, leveling the measuring scoop, or stirring the solution after each scoop. If students mentioned that they filled the cup "up to the top," this was interpreted to mean that they used more than 20 mL of water, and therefore, no credit was given for care in manipulating the amount of water used.

<u>The scoring form is reproduced below and on the next page:</u>

Place a check in the box that applies.

Accuracy of Results--Page 2

Recording Results

A Baking Soda	B Salt	C Baking Powder
2 grams 2	5-6 grams 2	1 gram 2

Sum the scores of the boxes checked | Max. 6

Interpreting Results

Least Grams	Most Grams
C Baking Powder 1	B Salt 1

Sum the scores of the boxes checked | Max. 2

Comparing Solubility (If least and most are correct, assume A correct.)

Least Soluble		Most Soluble
C Baking Powder 1	A Baking Soda 1	B Salt 1

Sum the scores of the boxes checked | Max. 3

Describing the Investigation--Page 3 and Question 1, Page 4

Control of Manipulation

A. Use 20 mL in each plastic cup OR Fill meas. cup up to 20 mL for e/solution	1
B. Add 1 g at a time	1
C. Stir/Mix the solution	1

Care in Manipulation (Assume NO if not explicitly written.)

A. Fill the measuring cup exactly 20 mL	1
B. Level the 1 gram scoop, measure the 1 gram exactly/carefully	1
C. Stir/Mix the solution after each scoop	1

Determine the Result (Read Question #1 on page 4)

A. Undissolved powder/material stays/accumulates on the bottom of the cup	1

Replication (Extra pt. if student did a replication; not considered in Max. Score)

A. Replicate the investigation or at least one of the solutions	1	
	Sum the scores of the boxes checked	Max. 7

Helping Dissolve A Powder--Question 2, Page 4

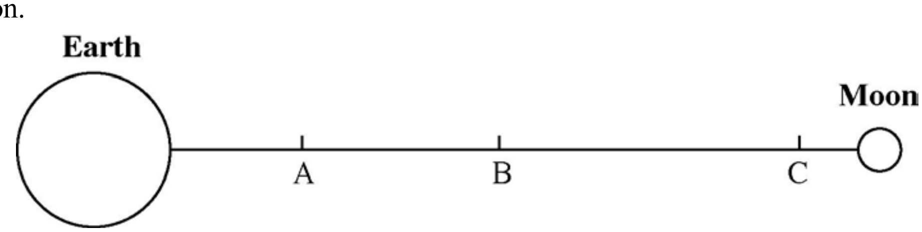
A. Increase the temperature/heat water	1	
B. Add water	1	
C. Stir the solution	1	
	Sum the scores of the boxes checked	Max. 3

Increasing The Amount of Water--Question 3, Page 4

A. Must be twice the amount reported in Question 1 for salt	1	
	Sum the scores of the boxes checked	Max. 1
	TOTAL	Max. 22

Item Source: NAEP 1996, Grade 8 (*Specifications*, pp. 149, 321)

A space station is to be located between the Earth and the Moon at the place where the Earth's gravitational pull is equal to the Moon's gravitational pull.
On the diagram below, circle the letter indicating the approximate location of the space station.



Explain your answer.

Scoring Rationale: Student demonstrates ability to explain the role of gravity in a man-made satellite and relates the force of gravity to the mass (size) of the object pulling it.

- **3 = Complete -** Student circles point C and gives a correct explanation that gravitational pull depends on mass and distance, thus the station must be closer to the Moon because the Moon's mass is less than that of the Earth.
- **2 = Partial -** Student circles point C and explains that the moon has less gravity than the Earth but does not link it to mass.
- **1 = Unsatisfactory/Incorrect -** Student circles A, B, or C and gives an incorrect explanation or no explanation.

Sample Student Responses

Complete (Level 3)

Explain your answer.

Point C because the earth has a stronger gravitational pull because of its size so the station would have to located nearer to the moon to equal pulls

Partial (Level 2)

Explain your answer.

The Earth has a greater gravitational pull than the moon so it needs to be closer to the moon.

Unsatisfactory/Incorrect (Level 1)

Explain your answer.

It would be right in the middle due to gravitational focus

Item Source: TIMSS 2003, Grade 4 (*Specifications*, p. 166)

On a hot, humid day the air contains a lot of water vapor. What happens to the water vapor in the air when the air becomes very cold?

Note: Priority should be given to Code 10. If a response mentions condensation or freezing, then Code 10 should be given even if other correct codes apply. Responses that mention ONLY that the water vapor becomes cold or rises without any mention of a change of state (explicitly or implicitly) are scored as incorrect (Code 70 or 71).

Code	Response	Item: S031382
	Correct Response	
10	Refers to **condensation** or **freezing** (or equivalent). *Examples: It freezes.* *It condenses.* *Condensation.* *It condenses and turns into rain.*	
11	Mentions cloud formation or a form of precipitation (e.g., rain, snow, fog, etc.) *Examples: The water vapor changes to rain.* *It changes to snow.* *Water vapor turns into clouds.* *It rises into the clouds and becomes rain droplets.* *It turns foggy.* *It rains.*	
19	Other correct *Examples: It falls to the ground.*	
	Incorrect Response	
70	Mentions only that the water becomes cold. [No mention of a change of state or precipitation.] *Examples: The water vapor becomes cold.* *Its temperature drops.*	
71	Mentions only that water vapor rises (or similar). [No mention of condensation or precipitation.] *Examples: The water vapor will rise on a hot day.*	
79	Other incorrect (including crossed out/erased, stray marks, illegible, or off task) *Examples: It disappears.*	
	Nonresponse	
99	Blank	

Item Source: NAEP 1996, Grade 8 (*Specifications*, p. 166)

When a population of mice is infected with parasites, many of the mice die from the parasitic infection, but some mice appear as healthy as they were before being infected. Some people are considering using these parasites to control the mouse population in people's homes.

Give one advantage and one disadvantage of using these parasites instead of mouse traps or poisons to limit the population of mice.

Scoring Rationale: Student demonstrates understanding of an advantage and a disadvantage of using parasites to control the mouse population in people's homes.

2 = Complete - Student response demonstrates understanding of an advantage and a disadvantage of using parasites to control the mouse population in people's homes.
1 = Partial - Student response demonstrates understanding of an advantage or a disadvantage of using parasites to control the mouse population in people's homes.
0 = Unsatisfactory/Incorrect - Student response demonstrates no understanding of the advantages and disadvantages of using parasites to control the mouse population in people's homes.

Credited responses include:

Advantages: mice will spread it to other mice; cheaper; more specific and directed than a trap or poison

Disadvantages: some mice are resistant; resistant mice will be selected for; the parasites may infect pets or people.

Item Source: TIMSS 1995, Grade 8 (*Specifications*, p. 166)

One day when the temperature was just below 0°C, Peter and Ann made snowballs. They put a thermometer into one of the snowballs and it showed 0°C. They tried to make the snowball warmer by holding it in their hands. What do you think the thermometer showed after two minutes? Explain your answer.

Code	**Response**
Correct Response	
20	Reports 0 degrees or mentions "the same temperature". The explanation includes: Snow cannot be warmer than 0 degrees. *Example: The melting point of snow is 0 degrees.*
29	Other correct.
Partial Response	
10	0 degrees or "the same temperature". No explanation or an incorrect explanation.
19	0 degrees or "the same temperature" *Example: Some snow melts, but it will not be warmer.* Other partially correct.
Incorrect Response	
70	Above 0 degrees, because the hands are warm.
71	Above 0 degrees, because the snow melts.
72	Above 0 degrees: No explanation.
79	Other incorrect.
Nonresponse	
90	Crossed out/erased, illegible, or impossible to interpret.
99	BLANK

Item Source: New Standards Spring Field Test 1999 for High School
(*Specifications*, p. 167)

The biosphere (living organisms), the lithosphere (rocks and soils of the Earth's crust), and the atmosphere are all involved in the cycling of carbon atoms. Describe the role that each plays in the carbon cycle.

Scoring guide:

Score point	Description
4	The response describes the cycling of carbon among all three reservoirs. The response is complete and detailed, showing evidence of logical reasoning. There is no evidence of misconceptions.
3	The response describes the cycling of carbon between two of the reservoirs. The response may contain omissions or minor errors.
2	The response describes the presence of carbon compounds in two or more reservoirs but may not link the reservoirs. The response may contain errors or misconceptions.
1	The response is largely incomplete, lacks detail, and contains errors of fact and reasoning.
0	The answer may contain words from the question, but does not add any information that might answer the question.
Off topic	Off-topic response.
Blank	No marks were made in the student response section.

Background information for appropriate response:

Carbon forms a variety of compounds. Carbon monoxide (CO) and carbon dioxide (CO_2) are carbon compounds that exist as gases in the atmosphere. Organic carbon compounds come in a wide variety of forms in living organisms: sugars, carbohydrates, cellulose, starches, collagen, chitin, amino acids, proteins, lipids, and nucleic acids. These organic compounds are also to be found in soil due to decaying and decomposition. The carbonate ion (CO_2^{-2}) combines with different positive ions to form carbonate compounds found in rocks and soil.

Producers (green plants) take carbon dioxide from the air and convert it to sugar, using solar energy (photosynthesis). Additional nutrients (including carbonate ions from the soil) are taken into the plant via its root system. These nutrients are rearranged and turned into the wide array of carbon compounds listed above, along with the sugar compounds formed via photosynthesis. As consumers eat producers, the carbon compounds are broken down and rearranged into carbon compounds used by the consumer organism.

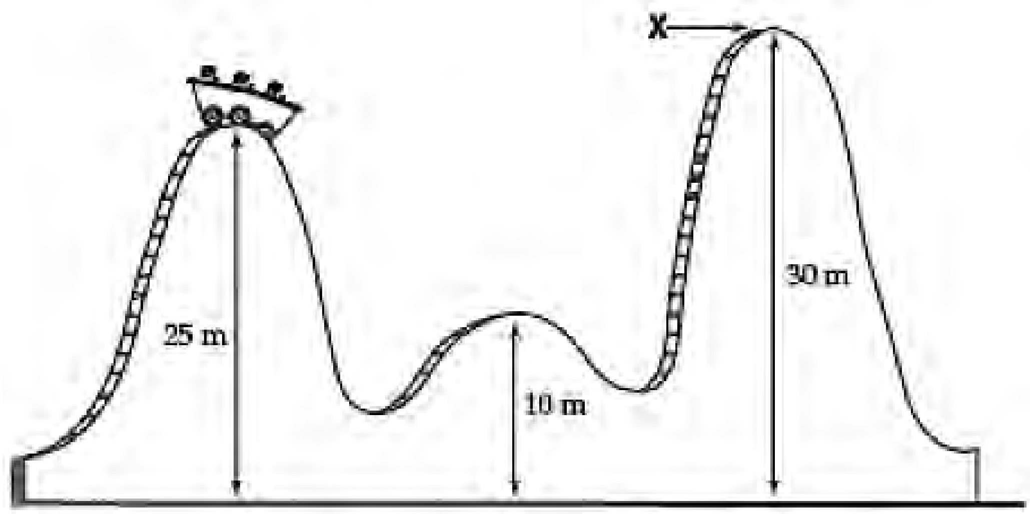
When any living organism dies, the carbon compounds are broken down and returned to the soil by decomposer organisms. Huge sediment layers of dead marine organisms, which once had chitin cell walls or shells, compress under the weight of continuing sediments and the ocean. Uplifting of these layers form sedimentary cliffs and rocks. Weathering and erosion continue to change the carbon compounds into forms that will again be incorporated into living organisms.

Producer, consumer, and decomposer organisms harness the chemical bond energy stored in food via cellular respiration. Cellular respiration releases carbon dioxide back into the atmosphere.

The score is based on evidence of: Geochemical cycling of carbon.

Item Source: Colorado State Assessment Program, 2001, Grade 8 Science (*Specifications*, p. 168)

Look at the picture of a roller coaster below.



The car on the roller coaster is released from the position shown and allowed to roll freely.

a. Name **two** of the forces that affect the motion of the car while it moves on the roller coaster.
b. Describe how the potential and kinetic energy of the car change as the car rolls downhill.
c. Explain why the car **cannot** reach point **X** on the third hill, as shown in the picture.

<u>Scoring Rubric</u>

Key elements:

a) Any two of the following:

- gravity
- air resistance/air/wind
- friction

b) The kinetic energy increases and the potential energy decreases.

c) One of the following:

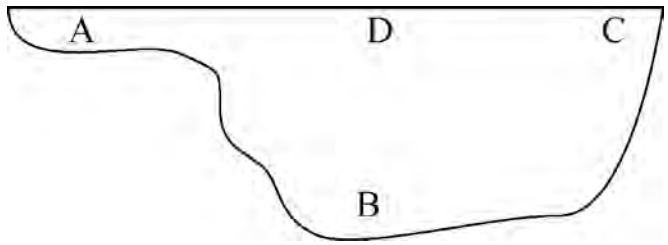
- The car loses energy due to friction/air resistance.
- The car started from a point that is lower than **X**; hence it does not have the energy to reach **X**.
- Any answer indicating that energy is lost **or** that the energy the car has at the starting point is less than the energy needed to reach **X**.

Score Points:

3 points=three key elements
2 points=two key elements
1 point=one key element
0 points=other

Item Source: TIMSS 1995, Grade 12 (*Specifications*, p. 178)

Here is a cross-section of a lake in the mountains. The air temperature gets below freezing in the winter and stays below freezing for 3 months.

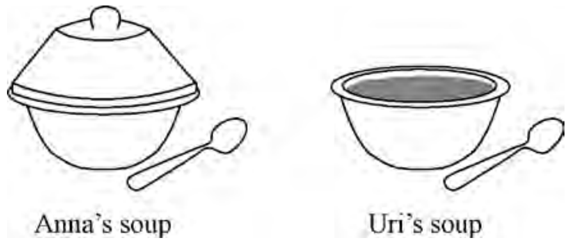


Not all of the water in the lake freezes. Which part of the lake will remain the warmest? Explain.

Code	Response
Correct Response	
20	B. Response refers to the maximum density of water (or the water is heaviest) at 4 degrees Celsius. *Example: Warmest at B because water has greatest density at 4º C so this water will stay there.*
29	Other acceptable responses.
Partial Response	
10	B. Refers to the fact that the water is 4 degrees Celsius at B without mentioning density.
11	B. Refers to the fact that ice will insulate this part of the water and/or that water is a bad heat conductor. *Examples: a) The surface will freeze first and then downwards.* *b) It takes time for heat and cold to get there.*
19	Other partially correct responses.
Incorrect Response	
70	B. No explanation.
71	B. Incorrect explanation referring to the heat from the earth (closer to the earth's center). *Example: The heat from the Earth will give heat to the water.*
72	B. Refers to the fact that hot water is heavier than cold water.
73	A/D/C with or without explanation.
76	Merely repeats information in the stem. *Example: B is the deepest point of the lake.*
79	Other unacceptable responses.
Nonresponse	
90	Crossed-out/erased, illegible, or impossible to interpret.
99	BLANK

Item Source: TIMSS 1995, Grade 4 (*Specifications*, p. 179)

Anna and Uri had identical bowls of soup, both at the same temperature. Anna put a cover on her bowl.



Whose soup do you think would stay hot longer? _____________________

Give a reason for your answer.

Code	**Response**
Correct Response	
20	Anna's soup. Mentions that heat or hot air does not escape Anna's soup or cold air does not enter it OR that heat (vapor, steam, smoke, etc.) disappears from Uri's soup or cold air enters it (or some combination).
29	Anna's soup. Other correct explanations.
Partial Response	
10	Anna's soup. Explanation refers to the cover. *Example: The soup with the cover.*
11	Anna's soup. Incomplete or incorrect explanation.
12	Anna's soup. No explanation.
19	Other partially correct.
Incorrect Response	
70	Uri's soup. Explanation is inadequate.
71	Uri's soup. No explanation.
79	Other incorrect.
Nonresponse	
90	Crossed out/erased, illegible, or impossible to interpret.
99	BLANK

Item Source: Shavelson et al., 1991 (*Specifications*, pp. 182, 189)

Electric Mysteries

The following is a brief description of two warm-up tasks:

1. Students are asked to connect one battery, one bulb, and wires so the bulb lights. They are then asked to draw a picture of this simple circuit.
2. Given mystery box "?," students are asked to identify whether it contains a battery or a wire. They are told that they can determine the contents of the mystery box by connecting it in a circuit with a bulb.

The following is an excerpt from the main task instructions given to students:

Find out what is in the six mystery boxes A, B, C, D, E, and F. They have five different things inside, shown below. Two of the boxes will have the same thing. All of the others will have something different inside.

[The five options—two batteries, a wire, a bulb, a battery and a bulb, nothing at all—are presented in words and drawings. Drawings are not provided here.]

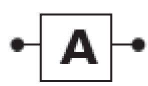
For each box, connect it in a circuit to help you figure out what is inside. You can use your bulbs, batteries, and wires in any way you like.

When you find out what is in a box, fill in the spaces on the following pages.

The following is an example of the student response format:

Box A: Has ____ inside.

Draw a picture of the circuit that told you what was inside **Box A**:



The following is a brief description of the scoring system:

For each of the six boxes (A-F), students' responses are scored on two components: (1) identification of the contents of the box and (2) the circuit used to make the conclusion. For each box, if both components are correct, the student receives 1 point; if one or both components are incorrect, the student receives 0 points. Total maximum score is 6 points.

Item Source: Shavelson et al., 1991; Shavelson et al., 1995
(*Specifications*, p. 182)

Students' performance on the sow bugs task was scored using the following rubric:

Hands-On Sow Bugs Score Form Experiment #2

Student __________________________ Observer __________________________ Score #2 ________

1. Method

A. Two conditions one dish **B.** One condition at a time

2. Control of Manipulation

A.	Equal area for each condition	Yes		No
B.	Number of bugs	1	2-3	4-5
C.	Starting location of bugs	One side	Middle	Half each side
D.	Sufficient time	Yes		No

3. Care in handling of bugs *(i.e. fry, drown, or crush)* Yes No

4. Determine Result

A. Count the number of bugs each location

B. Observe how busy the bugs are

C. Other ________________________________

5. Result *(logically follows from 4A)* Yes No

Grade	Method	Control of Manipulation				Care	Determine Result	Result
		Area	# Bugs	Location	Time			
A	1A	Yes	4-5	Middle/half	Yes	Yes	4A	Yes
B	1A	Yes	>1	Middle/half	Yes	Yes	4A	Yes
C	1A	Yes	>1	Any	Yes/No	Yes	4A	Yes
D	1A	No	1	Any	Yes/No	Yes	4B	Yes
F	1B	Yes	Any	Any	Yes/No	Yes/No	4A *or* 4B	No

Item Source: Persky et al., 2005 (*Specifications*, p. 188)

TRE Web Search Task

In this task, students are asked to use a search engine to determine why scientists use large helium gas balloons to explore outer space and the atmosphere instead of using satellites, rockets, or other such tools. One open-ended question and several multiple-choice questions are presented to students. Students are scored on how well they performed the search, the quality of the bookmarked pages, and how well the questions were answered.

A complex scoring framework is used to assess students' proficiency on the task. A "student model" is created which describes the "theory" of how different skills are linked together. For this task, there are five component skills: problem-solving in technology-rich environments, computer skill, scientific inquiry skill, exploration, and synthesis. An "evidence model" is created that describes how the students' responses are connected to each of these skills. Once students begin the task, every action is recorded and connected to one or more skills in the student model. Then, a three-step process is used to evaluate this record. The first step is "feature extraction," which shows what action the student took, when the action was taken, and any value that was associated with that action. The second step is "feature evaluation," which gives the scores for the actions taken based on developed rules that show the best way to complete the task. The third step is "evidence accumulation," which systematically combines responses into summary scores that detail the inferences that can be made from the students' responses.

This task incorporates a multifaceted evaluation process that can uncover the many skills involved in a single task or module. Thus, this module can provide a more comprehensive evaluation of students' skills and aptitudes than traditional test questions.

Item Source: Quellmalz et al., 2004 (*Specifications*, p. 190)

Solar Power Task

This is an interactive computer task in which students act as energy consultants to identify which two states will generate the most amount of electricity from photovoltaic cells. There are four modules in which students use a series of map visualizations and other data to reach and to present their final conclusions. Ultimately, students are asked to create a presentation to one state, recommending that the state apply for federal funds marked for solar energy use.

<u>Module 1</u> asks the student to review and apply background information on the conditions that both optimize and reduce solar energy production. They are also asked to conduct simple analysis using the "ArcView" map visualization program.

<u>Module 2</u> asks the student to explore several datasets to identify states with high incoming solar radiation and to manipulate the ArcView program and its "Map Calculator" tool. The student uses the ArcView program to generate visualizations and to calculate which states will generate the best monthly and annual electricity yields from solar panels.

<u>Module 3</u> asks the student to determine what other data may be necessary in order to create the most compelling recommendation to the states.

Module 4 asks the student to create a presentation to one of the two states determined to have the highest capacity for solar energy production. The student uses the newly gathered data to support the recommendation to the state.

Both generic and item-specific rubrics are used for scoring, which include scores for content (math and science), the use of problem solving or inquiry strategies (planning and thorough communication), and usage of technological tools.

For more information, see
http://ipat.sri.com/tasks/solarpower/subtasks/solar_tasks.html

Item Source: Adapted from Herl et al., 1999 (*Specifications*, p. 191)

Concept-Mapping Task

In this task, students use a custom software program to create a concept, or knowledge, map. Students are given eighteen environmental science terms: atmosphere, bacteria, carbon dioxide, climate, consumer, decomposition, evaporation, food chain, greenhouse gases, nutrients, oceans, oxygen, photosynthesis, producer, respiration, sunlight, waste, and water cycle. They are also given seven link labels: causes, influences, part of, produces, requires, used for, and uses. Students can then drag and drop these concepts onto the grid space of the mapping program and add, erase, and link the items in their newly constructed maps. The concept maps are scored based on semantic content, organizational structure, number of terms used, and number of links made.

Additionally, students explore a simulated World Wide Web space, which allows them to search for relevant environmental science information to improve their concept maps. Students can bookmark Web pages they believe to be helpful to the construction of their concept maps. This portion of the task is scored based on relevant information found, the hypertext links that were selected (browsing), keyword searching, and the accessing of three or more highly relevant Web pages for a single concept (focused browsing).

While performing the concept-mapping task, students are able to access real-time feedback, which compares students' maps to experts' maps and gives corresponding feedback as to which items are correct and which need improvement. A score is also assigned to how often a student accessed feedback, a measure of monitoring one's learning.

This computer-based assessment provides a detailed view into not only the ultimate performance of the student but also the steps of the thought processes that were employed to generate the ultimate product or answer (in this case, the concept map).

Item Source: Adapted from Washington Assessment of Student Learning, 2004, Grade 5 (*Specifications*, p. 314)

The Pacific Tree Frog lives in the Olympic rainforest in the state of Washington. This frog can change its color from brown to green very quickly.

Explain how this color change would help the frog survive longer.

In your explanation, be sure to:
Describe how changing color would help the frog get more food.
Describe how changing color would help the frog get away from its predators.

Use words, labeled diagrams, and/or labeled pictures in your answer.

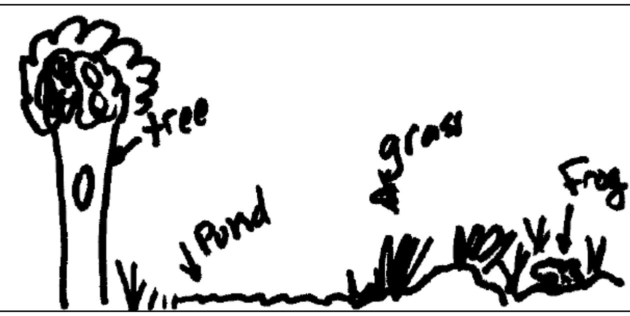
Scoring Rubric

2-point response: The response demonstrates that the student understands that an organism's ability to survive is influenced by the organism's behavior and by the ecosystem in which it lives.

The student explains how the color changes in the Pacific Tree frog help the frog survive longer. In the explanation, the student describes how changing colors would help the frog get more food.
AND
The student describes how changing colors would help the frog get away from its predators.

Example:
The frog can get more food because it blends in with its environment and its prey cannot see the frog as well. The frog can get away from its predators as the predators cannot see the frog as well.

Annotated Example of a 2-point response:



Color would help the frog get food and stay away form its
predators becaus brown and green will blend in with the dirt,.
trees and grass so its predators will not see the frog and things
that it eats will not see him.

Annotation:

The response demonstrates that the student understands that an organism's ability to survive is influenced by the organism's behavior and by the ecosystem in which it lives.

The response gives a correct description of how changing color would help the frog get more food, "becaus brown and green will blend in with the dirt,. trees and grass so its predators will not see the frog" (1 point)

The response gives a correct description of how changing color would help the frog get away from its predators, "…things that it eats will not see him."

Note: Using the "It" Rule, all 3 of the "its" in the response are replaced with "the frog" and the response still makes sense.

1-point response: The response demonstrates that the student has partial understanding that an organism's ability to survive is influenced by the organism's behavior and by the ecosystem in which it lives.

The student explains how the color changes in the Pacific Tree frog help the frog survive longer. In the explanation, the student describes how changing colors would help the frog get more food.
OR
The student describes how changing colors would help the frog get away from the frogs predators.

Annotated Example of a 1-point response:

When the frog is on a fern it will chang its color to green so it
can blend in and when the bug comes to close, the frog will eat
the bug. When the frog is on a tree and it sees an owl it will
turn brown to blend in and not become a meal.

Annotation:

The response demonstrates that the student partially understands that an organism's ability to survive is influenced by the organism's behavior and by the ecosystem in which it lives.

The response gives a correct description of how changing color would help the frog get more food, "…so it can blend in and when the bug comes to close, the frog will eat the bug." (1 point)

The response gives a correct description of how changing color would help the frog get away from its predators, "…it will turn brown to blend in and not become a meal." However, the misconception that the frog changes color in response to a predator is a minus point (see rubric Note #2) (0 point).

0-point response: The response demonstrates that the student has little to no understanding that an organism's ability to survive is influenced by the organism's behavior and by the ecosystem in which it lives.

Annotated Example of a 0-point response:

It probly helps the frog so that other animals cant hert it.

Annotation:

The response demonstrates that the student has little to no understanding that an organism's ability to survive is influenced by the organism's behavior and by the ecosystem in which it lives.

Using the "It" Rule, the first "It…" can be translated for the prompt to mean "changing colors." However, the second "…it" must be read as "changing colors" so the response makes no sense (0 points).

<u>Notes:</u>

1. The response needs to use the concept of hiding, blending in, or being invisible in relation to predators/prey.
2. The misconception that a frog changes color as response to the presence of a predator or prey will reduce the score by one point.

APPENDIX E

GROUP 2 SMALL-SCALE SPECIAL STUDIES

Knowing What Students Know about Technological Design

Knowledge about technology and the technological design process are prominent in both the *National Standards* and *Benchmarks*. The *National Standards* states, “Although these are science education standards, the relationship between science and technology is so close that any presentation of science without developing an understanding of technology would portray an inaccurate picture of science” (p. 190). Despite taking some time for schools to include technology in the curriculum for all students, there is growing recognition that technology should be an important component. The number of states that include technology in their standards is increasing. In 2001, 30 states included technology in their state standards; by 2004, the number had increased to 38 (73%) (Meade & Dugger, 2004). Consequently, the process of technological design is being included in NAEP as parallel to—though with less emphasis than—the process of scientific inquiry.

Since relatively few questions on NAEP will probe Using Technological Design, this study proposes the development of an additional set of questions to probe in depth students’ understanding of this practice.

Specifically, this study would address the following research question:

- What do students know about *technological design* in the contexts of agricultural technologies, energy generation technologies, and technologies related to Earth materials and resources?

Extended Investigations by Students

Science education standards nationally and locally emphasize scientific inquiry. In many states, this goal requires student engagement in projects that can take days, weeks, and even months as they undertake genuine investigations. Important outcomes of these projects include a range of skills that are a crucial feature of high quality science education but that cannot be assessed adequately in a 50-minute assessment (NRC, 2005a). They include, for example, gauging the quality of students’ (a) reasoning while framing their research questions, (b) planning for data collection and the execution of that plan, (c) ability to meet unpredictable challenges that arise during any actual, ongoing scientific investigation, (d) persistence in seeking productive explanations for their observations and revising plans for the investigation, (e) lines of argument in deciding how to alter their experimental approach in the light of new evidence, (f) engagement with fellow students and/or the teacher in interpreting an observation or result and deciding what to do about it, and (g) deliberations when settling on the defensible conclusions that might be drawn from their work.

In many countries, teachers are the ones expected to make assessments of student work during extended projects. Often their judgments of student achievement are made during ongoing classroom activities that are part of the regular curriculum. The assessments provided by the teachers are incorporated into an overall score that also includes results of the short, timed tests. In some places, a defined percentage of the total score is based on teachers' judgments about achievement associated with investigative projects.

This study, then, might include both a national sample of students and an exploration of what other countries do under similar circumstances. Specifically, the study would address the following research questions:

- What methods can be or have been developed to assess student achievement with respect to the ability to conduct extended scientific investigations?
- To what extent are shorter investigations interchangeable for the extended investigation and to what extent are they not?

APPENDIX F

PHYSICAL SCIENCE EXAMPLES OF GENERATING AND INTERPRETING ITEMS

Appendix F provides Physical Science examples of generating and interpreting items. Examples, which differ in level of completeness, are grouped by subtopic, and within each subtopic are sequenced by content statement code. Performance expectations given within each example are illustrative, not exhaustive. Table 28 summarizes the Physical Science content statements that are represented in this appendix.

Table 28. Physical Science Content Statements Represented in Appendix F

Properties of Matter
P4.1: Objects and substances have properties. Weight (mass) and volume are… **P4.2:** Objects vary in the extent to which they absorb and reflect light and conduct…[a] **P12.1:** Differences in the physical properties of solids, liquids, and gases are…[b]
Changes in Matter
P4.6: One way to change matter from one state to another and back again is by heating... **P8.6:** Changes of state are explained by a model of matter composed of tiny particles… **P12.5:** Changes of state require a transfer of energy. …[b]
Forms of Energy
P4.8: Heat (thermal energy) results when substances burn, when certain kinds of…[a] **P8.8:** Objects and substances in motion have kinetic energy. For example, a moving…[c] **P8.9:** Three forms of potential energy are gravitational, elastic, and chemical. …[c] **P8.10:** Energy is transferred from place to place. Light energy from the sun travels… **P12.8:** Atoms and molecules that compose matter are in constant motion (translational…[d]
Energy Transfer and Conservation
P12.12: Heating increases the translational, rotational, and vibrational energy of the…[b] **P12.16:** Total energy is conserved in a closed system.
Motion at the Macroscopic Level
P4.12: An object's position can be described by locating the object relative to other… **P4.13:** An object is in motion when its position is changing. The speed of an object is… **P8.14:** An object's motion can be described by its speed and the direction in which it is… **P12.18:** Objects undergo different kinds of motion—translational, rotational, and…[d]
Forces Affecting Motion
P12.19: The motion of an object changes only when a net force is applied.

[a]P4.2 and P4.8 are grouped together to generate a single example. See "Forms of Energy" section for this example.

[b]P12.1, P12.5, and P12.12 are grouped together to generate a single example. See "Properties of Matter" section for this example.

[c]P8.8 and P8.9 are grouped together to generate a single "Forms of Energy" example.

[d]P12.8 and P12.18 are grouped together to generate a single example. See "Forms of Energy" section for this example.

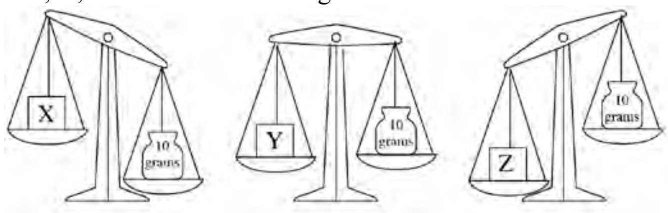
Properties of Matter

Grade 4: Matter—Properties of Matter
Content Statement **P4.1:** Objects and substances have properties. Weight (mass) and volume are properties that can be measured using appropriate tools.
Commentary Measurement involves comparison of a property with a standard measure of that property. Length and volume are examples of properties. The meter and liter are standard units of measure in the metric system. Rulers and meter sticks are tools used to measure length. Graduated cylinders, measuring cups, and repeated cubes are tools used to measure volume. Measurement of volume using water displacement is a sophisticated technique that is appropriate for the upper end of the 4th grade achievement scale. The property of volume is defined by the way in which volume is measured and the intended use of the volume measure. Consider, for example, a glass jam jar. The jam purveyor wants to know how much jam the jar holds—a volume measure; the maker of glass jars wants to know how much glass it will take to make the jar—a volume measure; the manufacturer of the shipping boxes wants to know how much space the jars take up—a volume measure. The volume of a liquid is conserved when it is poured from one container to another. From an assessment perspective, students' understanding of these aspects of measuring volume can be probed in two ways: (1) asking students to measure volume or (2) describing how volume was measured and asking students to interpret the measurement. See textbox on p. 26 for more detail on the distinction between weight and mass.
Examples of Performance Expectations Identifying Science Principles. Students can: • Recognize that when water is poured from one glass to another, volume is conserved (does not change). • Describe ways of measuring volume of solids and liquids and appropriate units of measurement (e.g., volume of water is measured in mL using a measuring cup). • Order a group of irregularly shaped objects according to weight (mass) and/or volume. Using Science Principles. Students can: • Interpret descriptions of procedures for measuring volume to identify the specific properties being measured. Using Technological Design. Students can: • Design a container that can hold a given volume of liquid. *Performance expectations for Using Scientific Inquiry are not provided.*

Item to Assess Identifying Science Principles

<u>Illustrative Item</u>

Which of the boxes X, Y, or Z has the least weight?



A. X
B. Y
C. Z
D. All three boxes have the same weight.

Key: A
Source: Adapted from TIMSS 1995, Grade 4

Items to Assess Using Science Principles

<u>Illustrative Item</u>

The picture below shows two cylinders that are the same size and shape, and contain the same amount of water. Next to each cylinder is a marble. The marbles are the same size. One is made of glass and the other is made of steel.

If a student puts a marble into each of the cylinders, what will happen to the water inside?

A. The water level will go up in both cylinders, but will go up higher in the cylinder with the steel marble.
B. The water level will go up the same in both cylinders.
C. The water level will not change in either cylinder.
D. The water level will go down in both cylinders, but will go down lower in the cylinder with the steel marble.

Key: B
Source: Adapted from WestEd, RISSA, 2001

<u>Item Suggestion 1</u>

A cylindrical glass jar is filled with water, and the quantity of water contained in the jar is measured with a graduated cylinder. What is being measured?

Key: This measures the capacity of the jar or the volume that it can hold.

Items to Assess Using Science Principles (cont.)

Item Suggestion 2

> A cylindrical glass jar with a metal screw-on lid is immersed in water, and the change in water level is converted to the volume of displaced water. What is being measured?

Key: This measures the space occupied by the jar with its lid on.

Item Suggestion 3

> An empty and uncovered cylindrical glass jar is immersed in water, and the change in water level is converted to the volume of displaced water. What is being measured?

Key: This measures the volume of the glass.

Item Suggestion 4

> The metal lid of a glass jar is immersed in water, and the change in water level is converted to the volume of displaced water. What is being measured?

Key: This measures the volume of the metal of which the cover is composed.

Interpretation: Before students can do Item Suggestions 2, 3, and 4, they need to know how to measure volume by water displacement. These items probe the upper end of the 4th grade achievement scale.

Items to assess Using Scientific Inquiry and Using Technological Design are not provided.

Grade 12: Matter—Properties of Matter and Changes in Matter and Energy—Energy Transfer and Conservation

Content Statements

P12.1: Differences in the physical properties of solids, liquids, and gases are explained by the ways in which the atoms, ions, or molecules of the substances are arranged and the strength of the forces of attraction between the atoms, ions, or molecules.

P12.5: Changes of state require a transfer of energy. …

P12.12: Heating increases the translational, rotational, and vibrational energy of the atoms composing elements and the molecules or ions composing compounds. As the translational energy of the atoms, molecules, or ions increases, the temperature of the matter increases. Heating a sample of a crystalline solid increases the vibrational energy of the atoms, molecules, or ions. When the vibrational energy becomes great enough, the crystalline structure breaks down and the solid melts.

Commentary

Taken together, the content statements above describe changes of state in atomic-molecular terms. A series of items that probes students' understanding of atomic molecular theory as it applies to these three content statements is a recommended approach. Individual items should still be given primary content statement designations in order to maintain an appropriate balance across subtopics.

Examples of Performance Expectations

Identifying Science Principles. Students can:
- Describe how atoms, ions, or molecules are arranged in solids, liquids, and gases.
- Describe the motion of atoms, ions, or molecules in a solid.
- Recognize that the quantity of heat required to melt different solids is different (e.g., it takes more heat to melt a mass of solid lead than to melt an equal mass of solid water).
- Recognize the effects of increased temperature on rates of changes of state.

Using Science Principles. Students can:
- Explain the melting of a solid in atomic-molecular terms, that is, in terms of the strength of intermolecular forces and increases in kinetic energy.
- Use the particulate model of matter to explain observations of evaporation, sublimation, and solubility.

Using Scientific Inquiry. Students can:
- Given data, find the relationship between temperature and pressure of a gas.

Performance expectations for Using Technological Design are not provided.

Items to Assess Identifying Science Principles

Illustrative Item 1

When most substances change from a solid to a liquid state, the particles of the substance

A. slow down.
B. move farther apart.
C. lose energy.
D. move closer together.

Key: B
Source: WestEd, 2005

Illustrative Item 2

The random molecular motion of a substance is greatest when the substance is

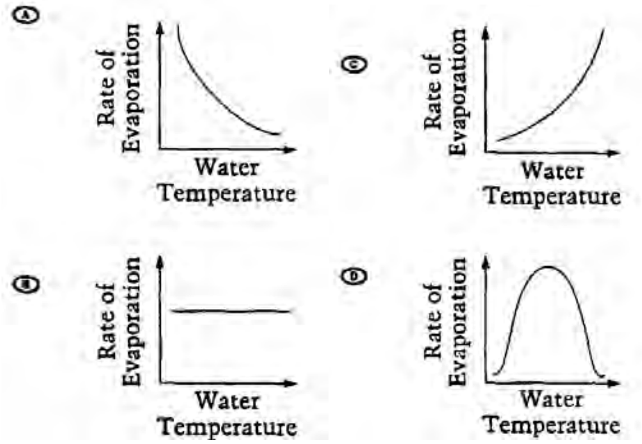
A. condensed.
B. a liquid.
C. frozen.
D. a gas.

Key: D
Source: California Department of Education, California Standards Test, 2004, Grade 12 Chemistry

Items to Assess Identifying Science Principles (cont.)

Illustrative Item 3

Which of the following graphs shows how the rate of evaporation changes with changes in water temperature?



Key: C
Source: Adapted from NAEP 1996, Grade 12

Item Suggestion

Which statement about the particles in ice and the particles in liquid water is correct?

A. The particles in ice have more energy than the particles in liquid water.
B. The particles in ice contain different atoms than the particles in liquid water.
C. The particles in ice have more electric charge than the particles in liquid water.
D. The particles in ice are less free to move than the particles in liquid water.

Key: D

Item to Assess Using Science Principles

Illustrative Item

If the attractive forces among solid particles are less than the attractive forces between the solid and a liquid, the solid will

A. probably form a new precipitate as its crystal lattice is broken and re-formed.
B. be unaffected because attractive forces within the crystal lattice are too strong for the dissolution to occur.
C. begin the process of melting to form a liquid.
D. dissolve as particles are pulled away from the crystal lattice by the liquid molecules.

Key: D
Source: California Department of Education, California Standards Test, 2004, Grade 12 Chemistry

Items to assess Using Scientific Inquiry and Using Technological Design are not provided.

Changes in Matter

Grade 4: Matter—Changes in Matter
Content Statement **P4.6:** One way to change matter from one state to another and back again is by heating and cooling.
Commentary Evaporation is difficult to understand at grade 4. Student understanding should deal with change from liquid to gas by the addition of heat (thermal energy), as in boiling, rather than the idea that, at room temperature, some molecules have very high kinetic energy. Appropriate substances for melting, freezing, and boiling are those that can be observed at temperatures that can be reproduced in a classroom (e.g., water). Students are expected to have a qualitative understanding of heat; that is, they are not expected to quantify the amount of energy required for changes of state. Condensation is an advanced concept for 4th graders.
Examples of Performance Expectations <u>Identifying Science Principles.</u> Students can: • Recall that matter exists as solids, liquids, and gases. • Recognize that changes of state are reversible and involve adding or removing heat (thermal energy) from a sample of substance. • Identify adding heat as a way to melt ice or evaporate liquids. • Recognize or recall names for changes of state: melting, freezing, evaporation, and condensation. <u>Using Science Principles.</u> Students can: • Cite examples of melting, freezing, evaporation, and condensation. <u>Using Scientific Inquiry.</u> Students can: • Describe procedures for adding or removing heat from a sample of a substance. *Performance expectations for Using Technological Design are not provided.*
Items to Assess Identifying Science Principles <u>Illustrative Item 1</u> Which of the following is most likely to cause water to change from one state of matter to another state of matter? A. increased precipitation B. the weathering of rocks C. a decrease in wind speed D. a change in air temperature Key: D Source: Massachusetts Department of Education, MCAS, 2002, Grade 5

Items to Assess Identifying Science Principles (cont.)

Illustrative Item 2

What change of state is shown?

A. Liquid to gas
B. Solid to gas
C. Gas to liquid
D. Solid to liquid

Key: A
Source: Oregon Department of Education, 2003-2005 Sample Test, Grade 5

Items to Assess Using Science Principles

Item Suggestion 1

In order to cool off, a student puts a piece of ice in her mouth. What causes the ice to melt?

Item Suggestion 2

People going on a trip may pack their lunches and drinks in a container with ice. Explain why the ice will keep the food cold for only one day.

Items to assess Using Scientific Inquiry and Using Technological Design are not provided.

Grade 8: Matter—Changes in Matter

Content Statement

P8.6: Changes of state are explained by a model of matter composed of tiny particles that are in motion. When substances undergo changes of state, neither atoms nor molecules themselves are changed in structure. Mass is conserved when substances undergo changes of state.

Commentary

Matter is composed of tiny particles that are in motion; they are held together by forces of attraction between and among the particles. Adding energy to a sample of matter increases the kinetic energy of the particles. As the kinetic energy of the particles in a sample of matter increases, the particles move faster and (on average) farther apart, and the force of attraction holding the particles together becomes weaker. A sufficient increase in kinetic energy causes the matter to change in state from solid to liquid or from liquid to gas. Conversely, removing energy from a sample of matter decreases the kinetic energy and slows down the motion of the particles, eventually causing the matter to change in state from gas to liquid or from liquid to solid.

Commentary (cont.)

Changes of state either require energy from the surroundings (evaporation and melting) or require energy to be added to the surroundings (condensation and freezing). Note that 8th graders should not be expected to explain the nature of the attractive forces that hold the particles together; quantify the amount of energy change required to cause a substance to change state; or explain why ice floats on liquid water (i.e., distance between water molecules explains different densities of ice and water).

Sublimation should be included at 8th grade. A common example of sublimation is dry ice becoming CO_2 gas without first melting.

High altitude cooking is an appropriate context for assessment items.

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Recall that matter is composed of tiny particles that are in motion.
- Define condensation, evaporation, freezing, and melting.
- Connect closely related examples of condensation.

Using Science Principles. Students can:

- Explain observations of changes of state including evaporation, condensation, and sublimation (e.g., seeing a person's "breath" on a cold day, fog on a bathroom mirror).
- Explain how energy from the sun drives changes of state.

Performance expectations for Using Scientific Inquiry and Using Technological Design are not provided.

Item to Assess Identifying Science Principles

Item Suggestion

> A student observes that there is "fog" on the bathroom mirror when a hot shower is running with the door and window closed and that, on a hot day, there are drops of water on the outside of a glass of cold water. How are these two observations similar?

Items to Assess Using Science Principles

Item Suggestion 1

> A student observed a puddle that forms in the schoolyard every time it rains. During the summer, the student noticed that it takes only one day for the water to disappear from the puddle if the sun is shining but three days for the water to disappear if the days are cloudy. Explain in detail how the puddle disappears and why it takes longer to disappear on cloudy days than on sunny days.

Interpretation: Students must draw on "declarative knowledge" about evaporation: when water seems to disappear, it is evaporating; evaporation is a change from the liquid to gas state; the sun shining warms the water; warming a sample of water increases the rate of evaporation. Students should also use the particulate model of matter to explain the observations described in the items, for example how heating a substance is related to increasing the kinetic energy of the particles composing the substance (thus drawing on "schematic knowledge"). To probe performance at the higher end of the achievement scale, students might be asked to use the particulate model to qualitatively discuss how forces of attraction between particles are related to the structure of the particles and the distance between them; and how the kinetic energy of the particles composing a substance is related to the distance between the particles.

Items to Assess Using Science Principles (cont.)

Item Suggestion 2

> A student observed that a puddle forms in the schoolyard when it rains during the winter. Overnight, the temperature dropped to below freezing, and the water in the puddle froze. It stayed cloudy and the outside temperature stayed below freezing for five days after the puddle froze. The student was very surprised to observe that the ice had disappeared at the end of those five days. How do you explain the student's observations?

Interpretation: Students should conclude that the principles explaining the evaporation of water (see interpretation for Item Suggestion 1 above) apply to the observation that ice "disappears" (sublimates) at temperatures below freezing.

Item Suggestion 3

> Why can you see a person's "breath" (or "cloud" or "puff") when he exhales outdoors on a cold day?

Key: The cold air outside causes the water in the person's breath to condense or even freeze.

Item Suggestion 4

> How does the water from a hot shower get onto the bathroom mirror (assuming that the door and window are closed)? What can you infer about the temperature of the bathroom mirror?

Key: The water evaporates and then condenses on the bathroom mirror. The mirror must be cooler than the steam.

Items to assess Using Scientific Inquiry and Using Technological Design are not provided.

Forms of Energy

Grade 4: Matter—Properties of Matter and Energy—Forms of Energy

Content Statements

P4.2: Objects vary in the extent to which they absorb and reflect light and conduct heat (thermal energy) and electricity.

P4.8: Heat (thermal energy) results when substances burn, when certain kinds of materials rub against each other, and when electricity flows though wires. Metals are good conductors of heat (thermal energy) and electricity. Increasing the temperature of any substance requires the addition of energy.

Commentary

Taken together, the content statements above concern the property of temperature of a substance. A series of items that probes students' understanding of temperature as it applies to these two content statements is a recommended approach. Individual items should still be given primary content statement designations in order to maintain an appropriate balance across subtopics.

Examples of Performance Expectations

Using Science Principles. Students can:

- Predict changes in temperature of a sample of substance.

Using Scientific Inquiry. Students can:

- Interpret a data table that shows the temperature of a volume of water as it comes to a particular temperature and answer such questions as "How fast did the water come to room temperature?"
- Design an investigation to determine changes in temperature of a sample of substance.

Using Technological Design. Students can:

- Design a container to keep substances hot (or cold) for a long period of time.

Performance expectations for Identifying Science Principles are not provided.

Item to Assess Using Science Principles

Item Suggestion

Predict the temperature of a glass of (ice or boiling) water that is left at room temperature for a day. Explain your prediction.

Item to Assess Using Scientific Inquiry

Item Suggestion

Hot water at 80ºC is mixed with an equal volume of cold water at 10ºC. Design an investigation to find the temperature of the mixture.

Item to Assess Using Technological Design

Item Suggestion

Determine which one of the following three cups works best to keep water cold for the longest time. Describe the criteria you used in your selection. Design another cup that will also keep water cold for a long time.

Items to assess Identifying Science Principles are not provided.

Grade 8: Energy—Forms of Energy

Content Statements

P8.8: Objects and substances in motion have kinetic energy. For example, a moving baseball can break a window; water flowing down a stream moves pebbles and floating objects along with it.

P8.9: Three forms of potential energy are gravitational, elastic, and chemical. Gravitational potential energy changes in a system as the relative positions of objects are changed. Objects can have elastic potential energy due to their compression, or chemical potential energy due to the nature and arrangement of the atoms.

Commentary

From a physical science perspective, kinetic and potential are two forms of energy. It is easy to understand that a rapidly flowing stream has energy, and thus, kinetic energy is a more intuitive concept than potential energy.

However, gravitational potential energy is also part of students' common experiences (e.g., a boulder perched on the edge of a cliff). For a more complete description of the kinetic and potential energy transfers that occur as an object falls, see content statement P8.12 and the description of "Energy Sources and Transfer" in the Crosscutting Content section of Chapter Two.

Chemical potential energy is a subtler concept. The "appearance" of thermal energy, light, and sound when wood burns can be explained by assuming that energy is stored in the reacting wood and oxygen; that is, wood and oxygen have chemical potential energy.

Performance expectations and items for Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Grade 8: Energy—Forms of Energy

Content Statement

P8.10: Energy is transferred from place to place. Light energy from the sun travels through space to Earth (radiation). Thermal energy travels from a flame through the metal of a cooking pan to the water in the pan (conduction). Air warmed by a fireplace moves around a room (convection). Waves—including sound and seismic waves, waves on water, and light waves—have energy and transfer energy when they interact with matter.

Commentary

Wave principles recommended for assessment at grade 8:

- Waves involve transfer of energy without a transfer of matter.
- Waves are caused by disturbances and are also themselves disturbances. Some of the energy of these disturbances is transmitted by the wave.
- Water, sound, and seismic waves transfer energy through a material.

Wave principles that are related but not recommended for assessment at grade 8:

- Some waves are transverse (water, seismic), and other waves are longitudinal (sound, seismic).
- In transverse waves, the direction of the motion is perpendicular to the disturbance.
- In longitudinal waves, the direction of motion is parallel to the disturbance.
- Waves (e.g., light waves) traveling from one material to another undergo transmission, reflection, and/or changes in speed.
- Waves can be described by their wavelength, amplitude, frequency, and speed (speed is frequency multiplied by wavelength; energy is a function of the amplitude for non-electromagnetic waves).
- Light has dual wave-particle properties.

Energy and refraction calculations are also not recommended for assessment at grade 8. Note that a quantitative understanding of electromagnetic waves is expected at grade 12. See P12.10.

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Identify disturbances that create sound, water, and seismic wave energy.

Using Science Principles. Students can:

- Describe the energy transfer of a thrown object in contrast to a wave moving across the room.
- Describe the transfer of energy from the sun to Earth (radiation).

Using Scientific Inquiry. Students can:

- Investigate how larger energy disturbances change the properties of a wave.
- Read and interpret a data table that shows the temperature of different colored solutions after exposure to a constant light source.

Using Technological Design. Students can:

- Critique the insulating capacity of containers (e.g., Styrofoam, plastic, metal).

Item to Assess Using Science Principles

Item Suggestion

The sun is 150 million kilometers from Earth. Explain in detail how the sun is able to melt an ice cube on Earth.

Item to Assess Using Scientific Inquiry

Item Suggestion

The following are data obtained in a student's experiment using equal volumes of three different colored solutions, thermometers, and a light bulb. Explain in detail how the temperature changed and why the solutions are at different temperatures.

Color of Solution	Distance from Light Bulb (cm)	Temperature after 15 minutes (^{o}C)
Dark Blue	15	30
Red	15	28
Clear	15	24

Items to Assess Using Technological Design

Item Suggestion 1

How can the same Styrofoam container be used to keep cold things cold and hot things hot?

Key: Styrofoam containers keep cold things cold and hot things hot by decreasing the heat exchange with the environment. (This is largely "declarative knowledge").

Item Suggestion 2

Why can't the same container be used to keep hot coffee hot and cold milk cold at the same time?

Key: The container can't be used to keep hot coffee hot and cold milk cold at the same time, because heat exchange will occur between the two liquids.

Items to Assess Using Technological Design (cont.)

Item Suggestion 3

Can you modify the container so that it could keep hot coffee hot and cold milk cold at the same time?

Items to assess Identifying Science Principles are not provided.

Grade 12: Energy—Forms of Energy and Motion—Motion at the Macroscopic Level

Content Statements

P12.8: Atoms and molecules that compose matter are in constant motion (translational, rotational, or vibrational).

P12.18: Objects undergo different kinds of motion—translational, rotational, and vibrational.

Commentary

The following are examples of translational, rotational, and vibrational motion at the macroscopic level. A car moving along a straight segment of a highway is undergoing translational motion, changing its position each second. A carousel rotating about a central point that never changes position is undergoing rotational motion, and thus, a horse on the carousel returns to its original position with each turn. The planets in our solar system turning on their axes and orbiting the sun are also undergoing rotational motion. A violin string moving rapidly back and forth is undergoing vibrational motion, and at the end of a single vibration, the string has not changed its position.

Molecules translate, moving randomly in straight lines from one position to another. Molecules rotate or turn on their axes much as a curve ball rotates as it flies from the pitcher's hand toward the batter. Molecules in solids undergo vibration, moving rapidly back and forth with respect to a central position.

Performance expectations and items for Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Energy Transfer and Conservation

Grade 12: Energy—Energy Transfer and Conservation

Content Statement

P12.16: Total energy is conserved in a closed system.

Commentary

The principle that energy is conserved is a powerful physical principle. When one form of energy seems to "disappear" from a closed system, knowledge of this principle motivates the search for other forms of energy in the system. The converse is also true: the "appearance" of energy in a closed system motivates the search for the energy source. For example, coal interacts with oxygen to produce thermal energy and light. While thermal energy and light are easily identified as forms of energy, the idea of conservation "suggests" that energy was originally contained in the coal and oxygen—hence, the notion of stored or chemical potential energy.

Commentary (cont.)

Application of the energy conservation principle prompts both qualitative and quantitative accounting. For example, if the gain in kinetic energy of a falling object is less than the loss of the object's gravitational potential energy, one might ask, "Where is the 'lost' energy?" Very often, the "lost" energy is found in the form of thermal energy. The falling object's surface rubs against the air through which it falls, increasing the temperature of the object and the air. The conservation principle prompts one to account for energy changes in a system that might otherwise be "missed" or attributed to measurement error.

Qualitative relationships observed in systems where energy transfers occur are pre-principles leading to the idea of energy conservation; that is, these relationships provide a foundation on which to build a quantitative understanding. Examples of such pre-principles include: as one form of energy "disappears" from a system, other forms of energy "appear"; and the quantity of one form of energy in a system decreases as the quantity of other forms of energy increases.

See the description of "Energy Sources and Transfer" in the Crosscutting Content section of Chapter Two.

Examples of Performance Expectations

<u>Using Science Principles.</u> Students can:

- Calculate the gravitational potential energy and kinetic energy of a falling object very close to Earth's surface.
- Calculate changes in the kinetic energy of a falling object based on changes in its gravitational potential energy.

Performance expectations for Identifying Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Items to Assess Using Science Principles

<u>Item Suggestion 1</u>

> A large boulder slides down a mountain and into a lake. What are the energy transfers that take place from the time the boulder begins to slide until it comes to rest on the bottom of the lake? How does the magnitude of the energy transfers compare with the boulder's gravitational potential energy before it began to fall?

Interpretation: Students at the upper end of the achievement scale could be expected to discuss the transfer of thermal energy to the ground and air because of friction.

<u>Item Suggestion 2</u>

> A 1000kg mass falls 1m. Ignoring the effects of air resistance, by how much does the mass's gravitational potential energy decrease? What is the mass's kinetic energy after falling 1m? What is its speed after having fallen 1m?

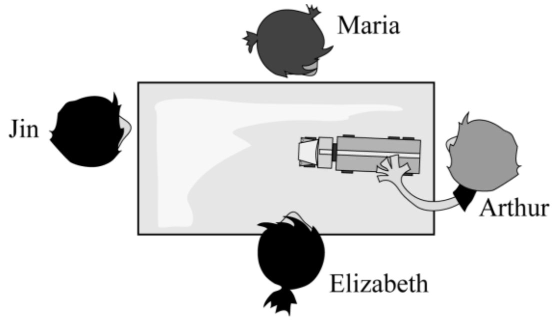
Items to assess Identifying Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Motion at the Macroscopic Level

Grade 4: Motion—Motion at the Macroscopic Level (The following example also appears in Chapter Four, Table 18.)
Content statement **P4.12:** An object's position can be described by locating the object relative to other objects or a background. The description of an object's motion from one observer's view may be different from that reported from a different observer's view.
Commentary The above content statement has the following implications: it is confusing when individuals observing the same motion from different positions describe it differently; using an external frame of reference results in observers describing the motion in the same way; and there is scientific value in consistent descriptions of an object's motion. A person observing the motion of a car from one side of the road might say that the car is moving from left to right, while a person on the other side of the road might say that the car is moving from right to left. Each person is using his/her body as the reference point. If the same observers each use a compass, they will observe that the car is moving from north to south. Their observations will be the same because they are using the same frame of reference, the cardinal points of the compass. An external reference point needs to be made clear when describing motion.
Examples of Performance Expectations <u>Identifying Science Principles.</u> Students can: • Recognize that in order to consistently describe an object's motion, an external reference point (e.g., North, South, East, West) is needed. <u>Using Science Principles.</u> Students can: • Interpret a situation in which more than one person observes an object in motion from different positions and describe or select descriptions of observations made by individuals from different positions. *Performance expectations for Using Scientific Inquiry and Using Technological Design are not provided.*

Items to Assess Using Science Principles

Item Suggestion 1



Arthur, Maria, Jin, and Elizabeth are sitting around a table. Arthur rolls a toy truck toward Jin. Arthur observes that the truck moves toward Jin. Maria observes that the truck moves from left to right.

a. What does Jin observe about the motion of the truck?
b. What does Elizabeth observe about the motion of the truck?
c. How do you explain differences in what each person observed about the truck's motion?

Item Suggestion 2

If Maria observes that the truck moves from left to right, what should Elizabeth observe about the motion of the truck?

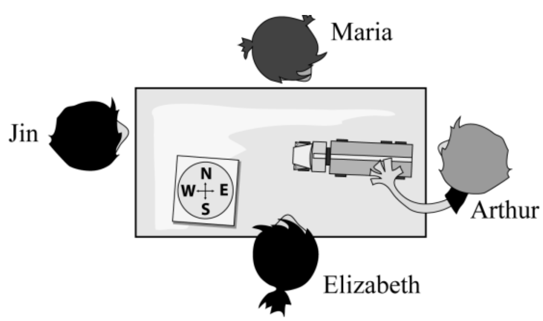
Elizabeth should observe that the truck moves

A. left to right.
B. right to left.
C. toward Arthur.
D. away from Jin.

Key: B

Item Suggestion 3

After the students have made their observations of the truck's motion, Ms. Fu, the science teacher, places a compass on the table and asks the students to use the compass to make their observations.



Items to Assess Using Science Principles (cont.)

The students use the compass to describe the truck's motion after Arthur rolls it toward Jin. Now, what will each student observe about the motion of the truck?

Student	Observation
Arthur	
Maria	
Jin	
Elizabeth	

Item Suggestion 4

The students described the motion of the truck with and without the compass. Which is the better way of describing the motion of the truck? Why do you think this way is better?

Interpretation: Students who respond correctly to the above assessment items have demonstrated knowledge about "left" versus "right." They are able both to describe the motion of an object using his/her body as the reference point and to mentally place herself/himself in the position of another person and use that perspective as a reference point. Students can distinguish the contexts in which it is or is not appropriate to use one's body as a reference point for describing motion; that is, using a compass is more appropriate for describing motion in a scientific context. Students know how to interpret a compass to determine direction. Advanced students also know why, when using the compass, the description of an object's motion will be independent of the observer's position and why the cardinal points as references for motion are useful only for motion in a single plane.

Items to assess Identifying Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Grade 4: Motion—Motion at the Macroscopic Level

Content Statement

P4.13: An object is in motion when its position is changing. The speed of an object is defined by how far it travels divided by the amount of time it took to travel that far.

No commentary on this content statement.

Examples of Performance Expectations

Using Scientific Inquiry. Students can:

- Given start and end positions for one object and different start and end positions for a second object (for the same time interval), determine which object went faster and explain why.
- Given a table of position versus time, find an object's position at a particular time (clock-reading).
- Given a table of position versus time, find an object's average speed during a given time interval.

Performance expectations for Identifying Science Principles, Using Science Principles, and Using Technological Design are not provided.

Items to assess Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Grade 8: Motion—Motion at the Macroscopic Level

Content Statement

P8.14: An object's motion can be described by its speed and the direction in which it is moving. An object's position can be measured and graphed as a function of time. An object's speed can be measured and graphed as a function of time.

No commentary on this content statement.

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Identify the units for speed.

Using Scientific Inquiry. Students can:

- Given a description of procedures and tools used by a person who claims an object moved with a particular motion, analyze and describe limitations to the methods used.
- Given a particular description of motion, draw representative distance versus time and speed versus time graphs.
- Describe the motion depicted by position versus time and speed versus time graphs.
- Calculate an object's speed given the distance it traveled and the amount of time taken to travel that distance.

Performance expectations for Using Science Principles and Using Technological Design are not provided.

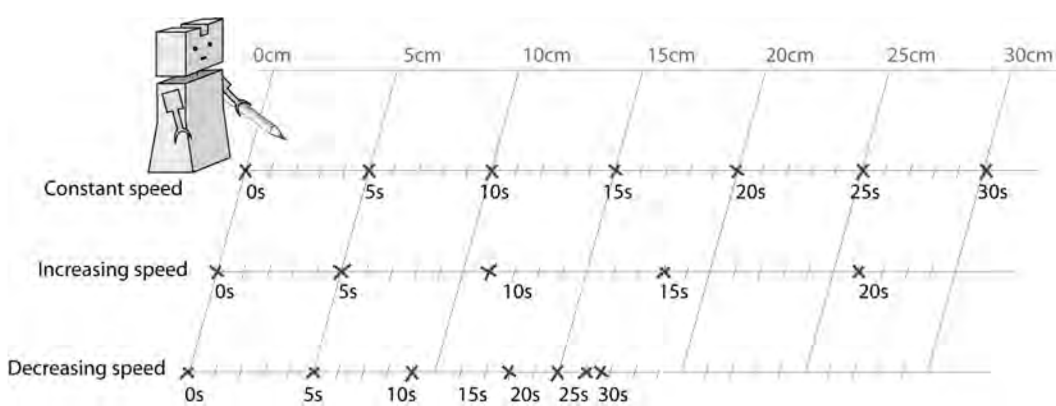
Items to Assess Using Scientific Inquiry

Item Suggestion 1

A position versus time graph is shown for a car going at a constant speed. If the car returned at constant speed to its start point, construct a position versus time graph.

Item Suggestion 2

A toy robot leaves a mark every 5 seconds on the floor.



a. Draw a distance versus time graph for the robot's motion at all three speeds.
b. Draw a speed versus time graph for the robot's motion at all three speeds.
c. Draw similar graphs for a toy moving faster than this robot.

Item to Assess Using Scientific Inquiry (cont.) Item Suggestion 3 Describe the motion of three toy cars given their position versus time and speed versus time graphs. Item Suggestion 4 A graph shows that an insect travels 10cm in 2 seconds. A second insect travels 10cm in 4 seconds. A third insect travels 20cm in 3 seconds. Which insect has the greatest speed? Explain your answer. Item Suggestion 5 A manufacturer states that an object moves with a constant speed along a straight line. Describe the tools and procedures that might be used to check this claim. What are the limitations of the experiment that you designed?
Items to assess Identifying Science Principles, Using Science Principles, and Using Technological Design are not provided.

Forces Affecting Motion

Grade 12: Motion—Forces Affecting Motion
Content Statement **P12.19:** The motion of an object changes only when a net force is applied.
Commentary The following describes a context in which the motion of an object can be changed due to the application of net forces. It takes a net force to change the direction of a baseball. The pitcher applies a force to the ball and sends it through the air. The batter applies a force to the bat, and the bat applies a force to the ball. The magnitude and direction of the forces on the ball ultimately determine the magnitude and direction of the change in motion of the ball; that is, determining the net force on the ball requires analysis of all of the forces acting on it (e.g., force applied by the pitcher, force of the bat, force of gravity).
Performance expectations and items for Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

APPENDIX G

LIFE SCIENCE EXAMPLES OF GENERATING AND INTERPRETING ITEMS

Appendix G provides Life Science examples of generating and interpreting items. Examples, which differ in level of completeness, are grouped by subtopic, and within each subtopic are sequenced by content statement code. Performance expectations given within each example are illustrative, not exhaustive. Table 29 summarizes the Life Science content statements that are represented in this appendix.

Table 29. Life Science Content Statements Represented in Appendix G

Organization and Development
L8.1: All organisms are composed of cells, from just one cell to many cells. About… **L12.2:** Cellular processes are carried out by many different types of molecules…[a]
Matter and Energy Transformations
L8.4: Plants are producers—they use the energy from light to make sugar molecules… **L12.6:** As matter cycles and energy flows through different levels of organization of…
Interdependence
L8.7: The number of organisms and populations an ecosystem can support depends on… **L12.7:** Although the interrelationships and interdependence of organisms may generate…
Heredity and Reproduction
L12.8: Hereditary information is contained in genes, located in the chromosomes of… [a] **L12.9:** The genetic information encoded in DNA molecules provides instructions for…[a]
Evolution and Diversity
L4.7: Different kinds of organisms have characteristics that enable them to survive in…

[a]L12.2, L12.8, and L12.9 are grouped together to generate a single example. See "Heredity and Reproduction" section for this example.

Organization and Development

Grade 8: Structures and Functions of Living Systems— Organization and Development
Content Statement **L8.1:** All organisms are composed of cells, from just one cell to many cells. About two-thirds of the weight of cells is accounted for by water, which gives cells many of their properties. In multicellular organisms, specialized cells perform specialized functions. Organs and organ systems are composed of cells and function to serve the needs of cells for food, air, and waste removal. The way in which cells function is similar in all living organisms.
Commentary For example, to burn food for the release of energy stored in it, oxygen must be supplied to cells, and carbon dioxide removed. In vertebrates, lungs take in oxygen for the combustion of food, and they eliminate the carbon dioxide produced. The urinary system disposes of dissolved waste molecules, the intestinal tract removes solid waste, and the skin and lungs rid the body of heat energy. The circulatory system moves all these substances to or from cells where they are needed or produced, responding to changing demands (quoted from *Benchmarks*, p. 137). Human organs and organ systems are subsumed under this content statement. See details in Table 10, Life Science Content Boundaries.
Performance expectations and items for Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Matter and Energy Transformations

Grade 8: Structures and Functions of Living Systems— Matter and Energy Transformations
Content Statement **L8.4:** Plants are producers—they use the energy from light to make sugar molecules from the atoms of carbon dioxide and water. Plants use these sugars along with minerals from the soil to form fats, proteins, and carbohydrates. These products can be used immediately, incorporated into the plant's cells as the plant grows, or stored for later use.
Commentary This content statement describes the basic matter transformations that occur in plants, making three key points: • Plants "use" light energy to make "food" from substances that are not food: carbon dioxide and water. (Note that the statement "they use the energy from light" does not imply that energy is converted into matter or that energy is lost.) • "Making food" is a chemical process. The atoms that make up molecules of carbon dioxide and water are rearranged to make sugar molecules. • Plants use the food they make for two purposes: to provide materials for their growth and energy for their functions.

Commentary (cont.)

Plants use sugars (and oxygen from the air), along with minerals from the soil, to form fats, proteins, and carbohydrates.

Students are not expected to know the terms photosynthesis or cellular respiration or to know the chemical intermediates in these processes. Students should be familiar with structural formulas for carbon dioxide, water, and oxygen.

For 8th grade, terms such as "food making," "breaking down food," and "oxidizing/burning fuel" are acceptable. See p. 53 for textbox on "food."

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Identify the raw materials that plants use to make sugars.
- Identify the reactants and products of plants' food-making processes (photosynthesis).
- Describe two possible ways that plants use the sugars they make.

Using Science Principles. Students can:

- Explain why many plants (e.g., grass, irises) grow better in the sunlight than in the shade.
- Predict what will happen to seeds (e.g., radish, grass) that sprout in the dark.
- Explain why sugars are found to move primarily down the stem of a growing plant (e.g., tree, bean plant).
- Explain why water is found to move primarily up the stem of a growing plant (e.g., tree, bean plant).
- Account for a plant's (e.g., maple tree, wheat plant, daisy) increase in mass from the molecular building blocks it makes.

Using Scientific Inquiry. Students can:

- Identify or state patterns in results from experiments on plant growth.
- Design experiments to assess factors affecting plant growth.
- Assess whether results of an investigation of plant growth are consistent with theoretical models.
- Critique or identify limitations of studies or investigations of plant growth reported in newspaper articles (e.g., investigations of the effects of carbon dioxide on plant growth).
- Criticize reasoning in arguments about claims that do not follow logically from data about plant growth under various conditions.

Performance expectations for Using Technological Design are not provided.

Item to Assess Identifying Science Principles

Item Suggestion

As a plant grows, what is the source of its food?

Interpretation: An acceptable answer should indicate that the student understands that the energy rich material made by a plant through photosynthesis is its food. For most students, plant nutrition is a major area of conceptual confusion, and food production through photosynthesis is a complex, abstract, and counterintuitive concept. As a result, students often have alternative conceptions of plant nutrition and growth that are not scientifically justifiable.

Item to Assess Identifying Science Principles (cont.)

For instance, if students mention that material absorbed from the soil by the plant is its food, then these students demonstrate that they are not able to link photosynthesis with food production in plants and are unclear of the distinction between plants' "food" and the nutrients that plants absorb from the soil. Similar confusion is evident in students that see sunlight as food for plants. These students incorrectly think that the sun's energy somehow gets converted into food during photosynthesis and thus are failing to conserve both mass and energy.

This item suggestion draws primarily on "declarative knowledge."

Items to Assess Using Science Principles

Item Suggestion 1

Where does a plant's increase in mass come from?

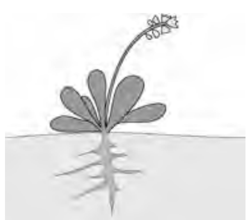
Interpretation: The acceptable response mentions that increase in weight comes primarily from conversion of carbon dioxide and water into plant matter as a result of photosynthesis (largely "schematic knowledge"). According to research, few students, even at the high school level, consider the conservation of matter and energy when thinking about plant growth and nutrition. Thus, if a student mentions that increase in weight of the plant comes from the material (nutrients) and water absorbed by the plant from the soil, then evidently the student is not linking photosynthesis with food production, is confusing nutrients with food, and is seeing food as providing only energy for living and not material for growth. This item suggestion draws primarily on "schematic knowledge."

Item Suggestion 2

Explain why water is found to move primarily up the stem of a growing plant (e.g., tree, bean plant).

Item Suggestion 3

Draw arrows to explain where water moves in a green plant. Explain why it needs to go there.



Interpretation: Most students should correctly draw arrows showing water going up the stem from the roots. Their reasons for why the water needs to go there, however, will be varied and revealing. Students who mention that plants need water for photosynthesis or for "making food" clearly understand a key idea in the content statement. Many students are likely to suggest that water *is* food for the plant or to describe the need of the leaves for water in vague terms (e.g., the leaves need water to live or to grow). Students who say that water is food for the plant are revealing a common naïve conception that does not recognize the role of plants as producers of sugar and other organic substances, i.e., food.

Items to Assess Using Science Principles (cont.)

Item Suggestion 4

> Draw arrows to explain how food moves through a green plant. Explain why it needs to travel this way.

____

____

____

Interpretation: The students' arrows and reasons reveal how they think about the origins of "food" for plants. Students who understand the content statement will identify food (or sugar) as a substance that is made in the leaves. Many students will reveal that they think of plants as animals—organisms that take in food from outside their bodies. The accepted answer to this question would be drawing of arrows going from leaves to other parts of the plant along with explanation that food is manufactured only in leaves and that is the reason why it moves from leaves to other parts of the plant. Students who think that plants get their food from soil are likely to draw arrows in the other direction (i.e., from roots to other parts of the plant). Some students may draw some arrows outside the plant that point towards the leaves. Such students may perhaps think of sunlight as food for plants. Students' reasons will often identify the substances that students consider to be food for plants. Here, again, there is an important distinction between substances/forms of energy that the plant takes in (water, sunlight, minerals, or "nutrients") and sugar or other substances made in the plant.

In combination, Item Suggestions 3 and 4 assess students' understanding of how plants obtain and use food and water ("schematic knowledge"). Both the arrows and the reasons that students give will reveal their underlying reasoning.

Items to Assess Using Scientific Inquiry

Item Suggestion 1

> Design an experiment to see how well plants grow under three different conditions—full sun, full shade, and partial sun/shade.
>
> What would be the best measure of plant growth?

Or, alternatively, in selected-response format:

> Which of the following would be the best measure of plant growth?
>
> A. Height of plant
> B. Circumference of plant
> C. Depth of plant roots
> D. Mass of plant

Items to Assess Using Scientific Inquiry (cont.)

Item Suggestion 2

> Design an experiment to see how well plants grow under three different conditions—full sun, full shade, and partial sun/shade. List three variables that need to be controlled in this experiment.

Key: Any three of the following: temperature, amount of water, amount and type of fertilizer, soil type, time of day of watering, type and size of initial plant

Item Suggestion 3

> A student carries out an experiment to see if the amount of fertilizer (0g, 5g, or 10g in 1 L of water) in the soil affects plant growth. The student plants one plant under each condition and measures their growth. There is a problem with this experiment. How would you fix it?

Items to assess Using Technological Design are not provided.

Grade 12: Structures and Functions of Living Systems—Matter and Energy Transformations

Content Statement

L12.6: As matter cycles and energy flows through different levels of organization of living systems—cells, organs, organisms, communities—and between living systems and the physical environment, chemical elements are recombined in different ways. Each recombination results in storage and dissipation of energy into the environment as heat. Matter and energy are conserved in each change.

Commentary

By 12th grade, an understanding of the way in which the cycling of matter and the flow of energy proceed on a larger time scale can and should be addressed. By this grade, students' appreciation of the atomic and molecular nature of matter makes it possible for them to understand change and constancy of ecosystems as well as appreciate how these processes contribute to the formation of one of the most important data sources available—the fossil record.

Note connections between this content statement and the "History of Earth" subtopic in Earth and Space Science.

Performance expectations and items for Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Interdependence

Grade 8: Structures and Functions of Living Systems—Interdependence

Content Statement

L8.7: The number of organisms and populations an ecosystem can support depends on the biotic resources available and abiotic factors, such as quantity of light and water, range of temperatures, and soil composition.

Commentary

Given adequate biotic and abiotic resources and no disease or predators, populations (including humans) increase at rapid rates. Lack of resources and the presence of factors such as predation and harsh climate limit growth of populations.

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Describe the concept of carrying capacity (though they are not expected to use this term).

Using Science Principles. Students can:

- Describe the interdependence of abiotic and biotic elements in a specific ecosystem (e.g., specialized ecosystem dependent on abiotic factors, such as a microclimate created by a cliff wall).

Performance expectations for Using Scientific Inquiry and Using Technological Design are not provided.

Items to assess Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Grade 12: Structures and Functions of Living Systems—Interdependence

Content Statement

L12.7: Although the interrelationships and interdependence of organisms may generate biological communities in ecosystems that are stable for hundreds or thousands of years, ecosystems always change when climate changes or when one or more new species appear as a result of migration or local evolution. The impact of the human species has major consequences for other species.

Commentary

Examples of human impact on other species include reducing the amount of Earth's surface available as habitats, interfering with food sources, changing the temperature and chemical composition of habitats, introducing foreign species into ecosystems, and altering organisms directly through selective breeding and genetic engineering. Other examples of the interdependence of organisms include relationships between the environment and public health and between migration and the potential spread of diseases.

Examples of Performance Expectations

Using Science Principles. Students can:

- Provide examples of "goods and services" provided by natural ecosystems (e.g., purification of air, climate regulation, flood control, soil generation, nutrient cycling, biodiversity).
- Predict positive and negative effects of various human activities on natural "goods and services."

Using Technological Design. Students can:

- Describe examples of the use of technology to restore ecosystems and hypothesize various science-based solutions to local, regional, and worldwide environmental issues. Examples of tools that could be used in this context include computer modeling and bioengineering (e.g., installing plants to prevent erosion).

Performance expectations for Identifying Science Principles and Using Scientific Inquiry are not provided.

Items to assess Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Heredity and Reproduction

Grade 12: Structures and Functions of Living Systems—Organization and Development and Changes in Living Systems—Heredity and Reproduction (The following example also appears in Chapter Four, Table 19.)
Content Statements The content statements below describe two main functions of DNA: (1) providing instructions for assembling protein molecules, which then specify the characteristics of organisms and (2) carrying the hereditary information from generation to generation. Here also are the foundations for the genetic basis of evolution. **L12.2:** Cellular processes are carried out by many different types of molecules, mostly proteins. Protein molecules are long, usually folded chains made from combinations of amino-acid molecules. Protein molecules assemble fats and carbohydrates and carry out other cellular functions. The function of each protein molecule depends on its specific sequence of amino acids and the shape of the molecule. **L12.8:** Hereditary information is contained in genes, located in the chromosomes of each cell. A human cell contains many thousands of different genes. One or many genes can determine an inherited trait of an individual, and a single gene can influence more than one trait. **L12.9:** The genetic information encoded in DNA molecules provides instructions for assembling protein molecules. Genes are segments of DNA molecules. Inserting, deleting, or substituting DNA segments can alter genes. An altered gene may be passed on to every cell that develops from it. The resulting features may help, harm, or have little or no effect on the offspring's success in its environment.
Commentary Relationships among proteins, DNA, genes, and traits can be briefly summarized: • Proteins carry out the work of cells. • Proteins are long, usually folded, chains of amino acids. • Protein function depends on its amino acid sequence. • DNA provides instructions for assembling proteins. • Genes are segments of DNA, which carry hereditary information and are located in the chromosomes of each cell. • One or more genes can determine an inherited trait. • Genes can be altered (e.g., inserting, deleting, and substituting DNA), and altered genes can be passed on to every cell that develops from them. • Resulting features from gene alteration may help, harm, or have little effect on offspring's success in its environment. For L12.2: Why does this content statement focus on proteins? From the perspective of biochemistry, proteins are the workhorses of cells. They are responsible for cell motility, movement of chromosomes during cell division, membrane transport, biosynthesis, and more. For example, nucleic acids assembled by proteins are biologically inert without proteins, which is why viruses aren't considered to be living things; phospholipids (themselves made by proteins) are passive barriers; and proteins transport ions and maintain the electrochemical gradients needed for such functions as nerve conduction. High school students are expected to understand specific examples of protein function as well as the generalization that proteins carry out the work of cells. (However, see exclusions in grade 12 content boundaries for "Organization and Development" and "Heredity and Reproduction.") For L12.8: Chromosomal abnormalities can be used as contexts for assessment items (e.g., explaining Down's Syndrome).

Commentary (cont.)

The following are not recommended for assessment at grade 12:

- Names and structures of nucleotides
- Nature of bonding between DNA strands
- Details of DNA transcription and translation
- Mendelian genetic details

Examples of Performance Expectations

Identifying Science Principles. Students can:
- List three ways in which genes can be altered.

Using Science Principles. Students can:
- Use knowledge of DNA and proteins to explain the basis for differences between two organisms of the same species.
- Explain the basis for particular hereditary disorders, given information about the associated DNA alteration or about critical proteins.
- Predict possible consequences to offspring of an altered DNA sequence in a parent organism.
- Explain possible consequences to an organism's proteins and cells resulting from a change in its DNA.
- Describe the relationship between any two of the following: DNA, genes, proteins, and traits.

Using Scientific Inquiry. Students can:
- Given data about particular genetic traits, critique possible explanations for them.
- Use pedigrees to establish when a mutation occurred in a family's history.
- Describe ways to determine possible causes of differences between organisms of the same species.

Using Technological Design. Students can:
- Given information about the relationships between particular DNA sequences, proteins, and traits of various plants within the same species, design the DNA of a plant to exhibit particular traits.
- Describe how to breed a model plant to exhibit a particular trait.

Items to Assess Identifying Science Principles

Item Suggestion 1

List two ways genes can be altered.

Key: Any two of the following: inserting, deleting, or substituting DNA

Item Suggestion 2

Which of the following is **not** a way genes can be altered?

A. Inserting DNA segments
B. Deleting DNA segments
C. Substituting DNA segments
D. Duplicating DNA segments

Key: D

Interpretation: Students answering these items correctly have recalled or recognized the "declarative knowledge" asked for in these items.

Item to Assess Using Science Principles

Item Suggestion

> Suppose you observe 1000 flowers of the same species blooming. You observe that all flowers have four petals except for one flower, which has three petals. Assume the DNA of the parent flower had been altered. Describe the sequence of events that could have occurred to result in the three-petal flower.

Key: The DNA segment (gene) for number of petals was altered in some way in the parent flower. That altered DNA segment was passed from parent to offspring through reproduction. The altered DNA segment results in the offspring having three petals, rather than four petals.

Item to Assess Using Scientific Inquiry

Item Suggestion

> Suppose you observe 1000 flowers of the same species blooming. You observe that all flowers have four petals except for one flower, which has three petals. Describe two ways in which you could investigate whether the three-petal trait is due to a recessive gene in normal inheritance or a manifestation of an altered DNA sequence.

Key: Trace the lineage of the flower as far back as possible; compare the molecular sequence of the four-petal flower and the three-petal flower.

Interpretation: This item probes the upper end of the 12^{th} grade achievement scale.

Item to Assess Using Technological Design

Item Suggestion

> Suppose you observe 1000 flowers of the same species blooming. You observe that all flowers have four petals except for one flower, which has three petals. You want to produce the three-petal flower because you believe it has aesthetic and economic value. Applying your understanding of genetics and the life cycle of organisms, how do you make certain you get a consistent set of three-petal flowers?

Key: Successively breed generations of three-petal flowers until they breed true.

Evolution and Diversity

Grade 4: Changes in Living Systems—Evolution and Diversity
Content Statement **L4.7:** Different kinds of organisms have characteristics that enable them to survive in different environments. Individuals of the same kind differ in their characteristics, and sometimes the differences give individuals an advantage in surviving and reproducing.

Commentary

This content statement describes what students should know about similarities and differences among organisms:

- Organisms of the same kind have some characteristics that enable them to survive and reproduce in the environment(s) where they live. (Note that this is circular: if they did not have these helpful characteristics, then they would not be found living in these environments.)
- Kinds of organisms living and reproducing in particular environments may differ from those living and reproducing in different environments.
- Organisms of the same kind vary in some of their characteristics, and some of these differences may help them find food and/or reproduce better than other members of their same kind.

Students should know that characteristics are those features that are held by all members of a kind of organism and can be used as a basis for recognizing them and sorting them into groups. The focus is on external characteristics that help organisms obtain food (e.g., talons for holding on to prey, beaks for cracking seeds or scooping up fish), make food (e.g., wide leaves that display a large surface area to sunlight), and reproduce (e.g., ways of distributing seeds, ways of protecting or nourishing young).

Students are not expected to be familiar with the term “species” or to be able to distinguish closely related kinds of organisms.

Examples of Performance Expectations

<u>Identifying Science Principles.</u> Students can:

- Based on looking at pictures, identify characteristic(s) that organisms of the same family have that may help them survive and/or reproduce in a particular environment.
- Based on looking at pictures, identify characteristics of the same kind of organism that vary.

<u>Using Science Principles.</u> Students can:

- Predict what might happen to plants and animals with particular characteristics (in terms of survival) if features of an environment changed.
- Given pictures or descriptions of environments, explain why kinds of organisms with particular (poorly adapted) characteristics might not survive.
- Based on descriptions of characteristics of an organism, explain why these characteristics may offer advantages in survival or reproduction.

<u>Using Scientific Inquiry.</u> Students can:

- Identify which tools would be most useful to collect data about particular characteristics of organisms and what measurements would be most useful to make.
- Design a fair test to identify necessary conditions for survival and reproduction of particular organisms in particular environments.

<u>Using Technological Design.</u> Students can:

- Given information about the survival and reproductive needs of particular endangered animals and alternative designs of zoo habitats, choose plans most likely to increase their numbers.
- Given information about the survival and reproductive needs of particular animals, identify possible consequences of proceeding with flawed designs of zoo habitats.
- Given information about the survival and reproductive needs of particular animals, suggest how to improve flawed designs of zoo habitats.
- Given information about the survival and reproductive needs of various plants, design a garden that takes into account the needs of these organisms.

Item to Assess Identifying Science Principles

<u>Illustrative Item</u>

Animals usually have physical features that help them live in certain places. Look at the picture shown above. This animal lives in a hot desert.

What physical feature does this animal have that helps it lose heat?

A. a thick coat of fur
B. a bushy tail
C. small eyes
D. large ears

Key: D
Source: TIMSS 2003, Grade 4

Items to Assess Using Science Principles

<u>Illustrative Item 1</u>

A girl found the skull of an animal. She did not know what the animal was, but she was sure that it preyed on other animals for its food. Which clue led to her conclusion?

A. The eye sockets faced sideways.
B. The skull was much longer than it was wide.
C. There was a projecting ridge on the front of the skull.
D. Four of the teeth were long and pointed.

Key: D
Source: Illinois Science Sample Test, 2006, Grade 4[47]

<u>Illustrative Item 2</u>

The Pacific Tree Frog lives in the Olympic rainforest in the state of Washington. This frog can change its color from brown to green very quickly.

Explain how this color change would help the frog survive longer.

In your explanation, be sure to:
Describe how changing color would help the frog get more food.
Describe how changing color would help the frog get away from its predators.

Use words, labeled diagrams, and/or labeled pictures in your answer.

(See Appendix D for item scoring guides.)
Source: Adapted from Washington Assessment of Student Learning, 2004, Grade 5

Items to assess Using Scientific Inquiry and Using Technological Design are not provided.

APPENDIX H

EARTH AND SPACE SCIENCE EXAMPLES OF GENERATING AND INTERPRETING ITEMS

Appendix H provides Earth and Space Science examples of generating and interpreting items. Examples, which differ in level of completeness, are grouped by subtopic, and within each subtopic are sequenced by content statement code. Performance expectations given within each example are illustrative, not exhaustive. Table 30 summarizes the Earth and Space Science content statements that are represented in this appendix.

Table 30. Earth and Space Science Content Statements Represented in Appendix H

Objects in the Universe
E4.1: Objects in the sky have patterns of movement. The sun, for example, appears to… [a] **E4.2:** The observable shape of the moon changes from day to day in a cycle that lasts…[a] **E8.2:** Gravity is the force that keeps most objects in the solar system in regular and…
History of Earth
E4.3: The surface of Earth changes. Some changes are due to slow processes, such as…
Properties of Earth Materials
E4.4: Earth materials that occur in nature include rocks, minerals, soils, water, and… [b] **E4.5:** Natural materials have different properties, which sustain plant and animal life.[b] **E4.6:** Some Earth materials have properties that make them useful either in their… [b]
Tectonics
E8.9: Lithospheric plates on the scale of continents and oceans constantly move at… [c] **E12.8:** Mapping of the Mid-Atlantic Ridge, evidence of sea floor spreading, and…[c]
Energy in Earth Systems
E8.11: The sun is the major source of energy for phenomena on Earth's surface. The… [d] **E8.12:** Seasons result from annual variations in the intensity of sunlight and length of…[d]
Climate and Weather
No content statements from this subtopic are represented in Appendix H.
Biogeochemical Cycles
E8.15: Human activities, such as reducing the amount of forest cover, increasing the… **E12.13:** Natural ecosystems provide an array of basic processes that affect humans. …

[a]E4.1 and E4.2 are grouped together to generate a single "Objects in the Universe" example.
[b]E4.4, E4.5, and E4.6 are grouped together to generate a single "Properties of Earth Materials" example.
[c]E8.9 and E12.8 are grouped together to generate a single "Tectonics" example.
[d]E8.11 and E8.12 are grouped together to generate a single "Energy in Earth Systems" example.

Objects in the Universe

Grade 4: Earth in Space and Time—Objects in the Universe
Content Statements **E4.1:** Objects in the sky have patterns of movement. The sun, for example, appears to move across the sky in the same way every day, but its path changes slowly over the seasons. The moon appears to move across the sky on a daily basis much like the sun. **E4.2:** The observable shape of the moon changes from day to day in a cycle that lasts about a month.
Commentary Young children are capable of observing patterns in the natural world. In the area of astronomy, the most important patterns involve the movements of the sun and moon, both of which can be seen during the day. As with the sun, the moon appears to move from the eastern side of the sky in the morning to the south and the western side in the afternoon. These patterns provide evidence for the modern model of the solar system.
Examples of Performance Expectations <u>Identifying Science Principles.</u> Students can: • Shown a sequence of shadows drawn and times marked during the day, pick out the shadow that was marked incorrectly. • Given several images of the moon in different phases, place them in order to show the cycle of moon phases. <u>Using Science Principles.</u> Students can: • Explain how a sundial works. • Given several images of the moon in different phases, estimate the period of time over which those images were taken. <u>Using Scientific Inquiry.</u> Students can: • Given a debate between two people about how long it takes to complete a full cycle of moon phases, describe a procedure that will settle the debate. • Given a debate between two people about the changing pattern of shadows cast by a flagpole during the day, describe a procedure that will settle the debate. <u>Using Technological Design.</u> Students can: • Apply their understanding of the motions of the sun to design a sun clock to tell time. • Apply their understanding of moon phases to design a calendar to track the passage of time.

Item to Assess Using Scientific Inquiry

Item Suggestion

> Two students were arguing about the changing pattern of shadows cast by a flagpole. Both students agreed that the shadow would move during the day. However, the first student thought that the tip of the shadow would move equal distances in equal times, while the second student thought the tip of the shadow would move slowly when the sun rose in the morning, faster during the middle part of the day, and then more slowly later in the afternoon.
>
> Plan a series of observations that would allow you to find out who is right, and explain how you would go about carrying out the plan.

Interpretation: A good plan will show evidence that the student has a reasonable idea of how the shadow will move and what observations will answer the research question. The simplest investigation would be to mark the position of the tip of the shadow with a piece of chalk or crayon during equal time intervals (e.g., once per hour or every 15 minutes) and then to measure the distances between the marks. A longer distance between the marks will indicate that the shadow was moving faster during that time period. The results can be kept in a table or graphed. Other features of a sound investigation include the use of an accurate watch or clock with a second hand and participation of several students to ensure accuracy. (Note that neither of the first two students is correct. The top of the shadow will move fastest in the morning, slower during midday, and then faster again in the late afternoon.)

Item to Assess Using Technological Design

Item Suggestion

> Imagine that you were born before there were printed calendars. You need to know when to plant crops, so you want to make a calendar to keep track of time. Describe what observations of the moon you would make to create a calendar.

Interpretation: This item provides an opportunity for students to show how they would apply their knowledge of the changing movements and appearance of the moon in the sky to meet the practical human need of devising a calendar to track the passage of time.

Students' responses should include evidence that they understand the cycle of moon phases, using the major phases to mark weeks and complete cycles to mark months. Students are not expected to use technical terms related to moon phases.

Items to assess Identifying Science Principles and Using Science Principles are not provided.

Grade 8: Earth in Space and Time—Objects in the Universe
(The following example also appears in Chapter Four, Table 20.)

Content Statement

E8.2: Gravity is the force that keeps most objects in the solar system in regular and predictable motion. Those motions explain such phenomena as the day, the year, phases of the moon, and eclipses.

Commentary

This content statement encompasses two interrelated sets of concepts:

1. Gravity acts between and among all objects in the solar system, and it plays an essential role in the regular and predictable motions of planets around the sun and satellites around planets.

    - On Earth, gravity is experienced as a force that pulls everything “down” towards the center of the Earth. (A common naïve conception is that the atmosphere “pushing” things down causes gravity.)
    - Gravity is a force of attraction that is exerted by every object on every other object.
    - Gravity exists in space and on other planets. (A common naïve conception among students is that there is no gravity in space because space has no air.)
    - The almost circular motion of planets and satellites results from the force of gravity and the tendency of a body to continue moving through space in a straight line unless acted upon by a net force.

2. The regular and predictable motions of the Earth, the sun, and the moon cause the cyclic phenomena that can be observed in the sky.

    - The day-night cycle results from Earth’s rotation on its axis once in 24 hours.
    - Annual changes in the visible constellations and the seasons result from Earth’s revolution around the sun once every 365-1/4 days.
    - Moon phases result from the moon’s orbit around the Earth about once a month, which changes what part of the moon is lighted by the sun and how much of the lighted part can be seen from Earth.

Note connection between this content statement and the Physical Science subtopic, “Forces Affecting Motion.”

Students are not expected to use the inverse square relationship of gravitational force and distance to find the strength of the gravitational force between two objects.

Students need not know that the motion of planets and satellites is elliptical and not circular.

Examples of Performance Expectations

<u>Identifying Science Principles.</u> Students can:

- Identify gravity as the force exerted by every object in the solar system on every other object.
- Identify gravity as the force that keeps the moon circling Earth, rather than flying off into space.
- Describe the regular motions of Earth through space, including its daily rotation on its axis, and its yearly motion around the sun.

<u>Using Science Principles.</u> Students can:

- Explain that the orbit of one object around another is due to the tendency of an object to move in a straight line through space and the force of gravity between the two objects.
- Explain how the monthly pattern of moon phases observed from a point on Earth results from the moon’s orbit around Earth, which changes what part of the moon is lighted by the sun and what portion of the lighted part can be seen from Earth.
- Distinguish between explanations for lunar (moon) phases and lunar eclipses.
- Explain that astronauts and other objects in orbit seem to “float” because they are in free fall, under the influence of gravity.

Examples of Performance Expectations (cont.)

Using Scientific Inquiry. Students can:

- Arrange a set of photographs of the moon taken over a month's time in chronological order and explain the order in terms of a model of the Earth-sun-moon system.
- Design a plan for observing the sun over a year's time to find out how the length of the day is related to the rising and setting point of the sun on the horizon.
- Design a series of observations or measurements to determine why some objects—such as certain asteroids or comets—visit the solar system just once, never to return.

Using Technological Design. Students can:

- Choose among several (qualitative) methods for aiming a rocket so that it reaches the planet Mars and give a rationale that shows understanding of orbital motion.
- Use scientific trade-offs in deciding whether or not to support a plan to observe and predict orbits of asteroids that enter the inner solar system.
- Given a scenario in which a person is shipwrecked on an island in the ocean, critique plans to create a calendar to keep track of the passage of time.

Items to Assess Identifying Science Principles

Illustrative Item 1

What force keeps the planets in our solar system in orbit around the Sun?

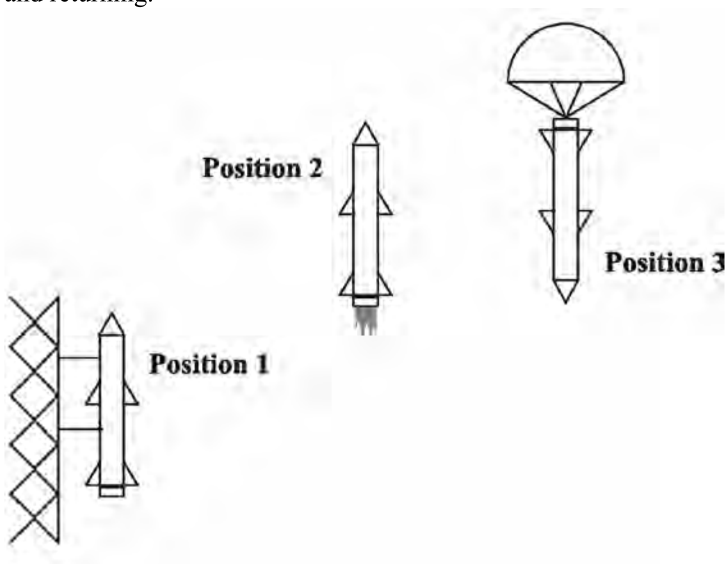
A. gravitational
B. magnetic
C. electrical
D. nuclear

Key: A
Source: Adapted from Massachusetts Department of Education, MCAS, 2000, Grade 8

Items to Assess Identifying Science Principles (cont.)

Illustrative Item 2

The drawings show a rocket being launched from Earth and returning.



In which of these positions does gravity act on the rocket?

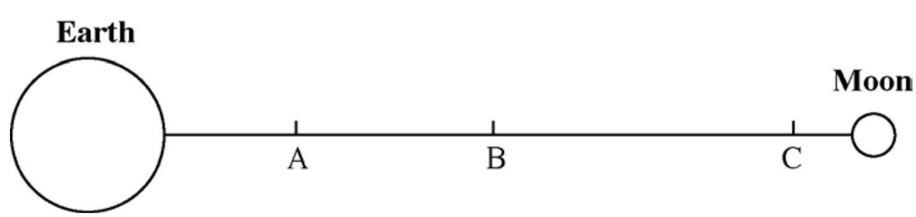
A. 3 only
B. 1 and 2 only
C. 2 and 3 only
D. 1, 2, and 3

Key: D
Source: TIMSS 1999, Grade 8

Items to Assess Using Science Principles

Illustrative Item

A space station is to be located between the Earth and the Moon at the place where the Earth's gravitational pull is equal to the Moon's gravitational pull.
On the diagram below, circle the letter indicating the approximate location of the space station.



Explain your answer.

Source: NAEP 1996, Grade 8

Items to Assess Using Science Principles (cont.)

Interpretation: The correct answer is C. Since the moon has 1/6 the amount of gravity as Earth, a body that experiences an equal gravitational force from Earth and the moon should be closer to the moon. Point C is the only point that is closer to the moon. Note: Point C is about 1/12 of the way between the moon and Earth; it should be 1/6 of the distance. (See Appendix D for the item scoring guide.)

Item Suggestion

Is there gravity in space? Which of the following gives the best response to this question?

A. No. You can see that astronauts float around weightless in their cabin.
B. No. There is no air in space, so gravity cannot exist there.
C. Yes. There must be gravity since planets keep circling the sun.
D. Some. The moon has one-sixth as much gravity as Earth, so we know there is some gravity in space.

Key: C

Interpretation: The correct answer is C. This question is drawn from a series of studies that show a common naïve conception—that there is no gravity in space because space has no air. The distractors are drawn from student interviews. It is likely that these naïve conceptions stem from images that students have seen of astronauts floating around in a "weightless" environment while in orbit. This item probes "schematic knowledge."

Item to Assess Using Scientific Inquiry

Item Suggestion

A student is presented with a set of photographs of the moon taken over a month's time. The photos are not presented in chronological order. The student is asked to arrange them in the order in which they were taken and explain the reason for moon phases.

Interpretation: This suggestion reflects items used frequently in curricular materials (e.g., Schatz & Cooper, 1994). Students are asked to find patterns in the data. First, they should be sufficiently familiar with the lunar cycle to arrange the photographs in order, either in a line to represent a chronology or in a circle to represent a cycle (tapping "declarative knowledge" and "procedural knowledge" to a lesser extent). Then, students should be able to explain moon phases in terms of the moon circling Earth and the changing angle between the sun and moon as observed from Earth. This part of the item probes "schematic knowledge." This is a challenging question that many educated adults fail. However, studies show that middle school students can learn to do this by observing lunar phases and explaining them using a model of the Earth-sun-moon system (Barnett & Morran, 2002; Kavanagh, Agan, & Sneider, 2005).

Items to Assess Using Technological Design

Item Suggestion 1

NASA wants to launch a spacecraft with rockets from Earth so that it will reach and orbit Mars. Which of the following statements about this flight is WRONG?

A. In the first phase of its flight, the forces acting on the spacecraft are the thrust of the rocket engine, gravity, and friction from the Earth's atmosphere.
B. When the rocket engine shuts off, the only force acting on the spacecraft is the force of gravity.
C. Once the spacecraft is above the Earth's atmosphere and the rocket engine is off, it will travel at a constant speed since there is no gravity in space.
D. If the spacecraft is aimed correctly and has the proper speed, the spacecraft will reach Mars and require only engine braking to attain orbit.

Key: C

Interpretation: The correct answer is C, since there is gravity in space and planning for such a rocket flight would need to take into account the gravity from Earth, Mars, and the sun ("declarative knowledge"). This question is drawn from a series of studies that show that the following naïve conceptions about gravity are common among many students at the middle school, high school, and even college levels: If a body is moving, there is a force acting on it in the direction of motion (Finegold & Gorsky, 1991; Gunstone & Watts, 1985; Sequeira & Leite, 1991); there is no gravity in space (Bar, Zinn, Goldmuntz, & Sneider, 1994; Chandler, 1991; Morrison, 1999); and gravity cannot act in space because there is no air in space (Bar & Zinn, 1998). One study showed that, with effective instruction, middle school students can overcome these naïve conceptions and learn that gravity does, in fact, act in space, where it keeps satellites and planets in their orbits (Bar, Sneider, & Martimbeau, 1997).

Item Suggestion 2

Decisions about whether or not to develop new technologies always concern trade-offs. For example, most scientists today believe that the extinction of the dinosaurs and many other species was caused by the collision of a large asteroid with Earth 65 million years ago. As a result, there is a proposal to develop two new technologies: (1) the detection and tracking of all asteroids large enough to do considerable damage if they should strike Earth, and (2) the development of means of sending a spacecraft to meet the asteroid in space and change its path. Write a paragraph describing:

a. whether or not you think it would be possible to develop these technologies based on your knowledge of science; and
b. some of the scientific trade-offs that should be considered in deciding whether or not to develop these new technologies.

Interpretation: Look for evidence that the students understand what asteroids are, that scientists have observed asteroids, and that observations taken at several points in time allow for the prediction of an asteroid's path. Also look for evidence that students understand that spacecraft can be built, launched, and navigated to intercept solar system bodies. (In fact, several spacecraft have intercepted asteroids and comets.) Regarding scientific trade-offs, look for evidence that students recognize the advantages of the proposed technologies (e.g., avoid a catastrophic collision in which billions of people and animals could die) as well as possible negative effects (e.g., break-up of the asteroid so there are many collisions rather than one, accidents on launch).

History of Earth

Grade 4: Earth in Space and Time—History of Earth
Content Statement **E4.3:** The surface of Earth changes. Some changes are due to slow processes, such as erosion and weathering, and some changes are due to rapid processes, such as landslides, volcanic eruptions, and earthquakes.
No commentary on this content statement.
Examples of Performance Expectations Identifying Science Principles. Students can: • Identify wind and water movement as causes of erosion. • Identify causes of weathering (e.g., rocks broken apart by plant growth, water expansion and contraction). • Describe what happens over time to soil or soft rock as water flows rapidly over it. • Given descriptions of weathering and earthquakes, classify them as causing either dramatic quick changes or slow gradual changes to the Earth's surface. Using Science Principles. Students can: • Analyze alternative methods for preventing erosion and select the best one for a given scenario. • Given two pictures showing a change on the surface of Earth, describe potential causes of the change (e.g., erosion, landslide, earthquake). • Given descriptions of a site (e.g., widths of four cracks in the earth near a mound of soil), evaluate predictions about what is likely to happen at this site in the future (e.g., landslide). Using Scientific Inquiry. Students can: • Find patterns in data comparing erosion of different soil types (e.g., sod-covered soil, bare and loose soil, prairie-grass-covered soil). Relate the pattern to ways to prevent erosion. Using Technological Design. Students can: • Critique the plans (e.g., wood barriers, planting flowers) of a ditch contractor to stop or control water erosion on a ditch having loose, bare soil. • Predict the effects (e.g., erosion, flooding) of various activities (e.g., building construction, road paving, land clearing) on Earth surfaces. • Predict the effects of natural erosion on various land uses (e.g., where houses are built, where crops are planted).
Item to Assess Using Scientific Inquiry Item Suggestion (Students are shown a mound of non-oily clay [not modeling clay] formed into a mass and having pebbles pressed into its surface.) Design an investigation to determine what happens when water runs over the clay.
Items to assess Identifying Science Principles, Using Science Principles, and Using Technological Design are not provided.

Properties of Earth Materials

Grade 4: Earth Structures—Properties of Earth Materials
Content Statements **E4.4:** Earth materials that occur in nature include rocks, minerals, soils, water, and the gases of the atmosphere. **E4.5:** Natural materials have different properties, which sustain plant and animal life. **E4.6:** Some Earth materials have properties that make them useful either in their present form or designed and modified to solve human problems and enhance the quality of life, as in the case of materials used for building or fuels used for heating and transportation.
No commentary on these content statements.
Examples of Performance Expectations Identifying Science Principles. Students can: • Classify examples of materials as either natural Earth materials or human-made materials. • Identify Earth materials that are used to grow food (e.g., air, water, soil). • Recognize that materials used for building, heating, and transportation come originally from Earth materials. • State how Earth materials help plants and animals to live (e.g., provide sources of food, shelter, defense). Using Science Principles. Students can: • Explain why certain Earth materials are used for building houses by citing their specific properties. • Given a picture of a home or garden, describe ways that natural materials have been adapted for human use. • Give one example of how rocks and minerals are used in people’s homes. Using Scientific Inquiry. Students can: • Given data on properties of materials, identify those data consistent with particular conclusions about the materials. Using Technological Design. Students can: • Given information on materials used by Native American peoples to build structures, define the specific problem that each material was designed to solve. • Given a picture of a tree, identify three things a human could do to modify the tree in order to increase its utility to humans. • Compare the advantages and disadvantages of gasoline-powered cars and electric cars.
Items to assess Identifying Science Principles, Using Science Principles, Using Scientific Inquiry, and Using Technological Design are not provided.

Tectonics

Grade 8: Earth Structures—Tectonics

Content Statements

E8.9: Lithospheric plates on the scale of continents and oceans constantly move at rates of centimeters per year in response to movements in the mantle. Major geological events, such as earthquakes, volcanic eruptions, and mountain building, result from these plate motions.

E12.8: Mapping of the Mid-Atlantic Ridge, evidence of sea floor spreading, and subduction provided crucial evidence in support of the theory of plate tectonics. The theory currently explains plate motion as follows: the outward transfer of Earth's internal heat propels the plates comprising Earth's surface across the face of the globe. Plates are pushed apart where magma rises to form mid-ocean ridges, and the edges of plates are pulled back down where Earth materials sink into the crust at deep trenches.

Commentary

Although this sequence is suggested for grade 8, the knowledge expected and the examples of items also address the relevant content statement for grade 12, as provided above. Hence, this sequence should be viewed as illustrative of the upper end of the 8th grade achievement scale as well as being appropriate for the high school grades, including grade 12.

The content statements address evidence and explanations that comprise the theory of plate tectonics. Lithospheric plates are bounded by constructive and destructive boundaries. Lithospheric plates are comprised of either continental crust material or oceanic crust material, the latter being more dense. There are four types of plate boundaries: spreading along oceanic plates; collision between two continental crust plates; collision between a continental crust plate and an oceanic crust plate; and a slip/slide fault collision between two continental crust plates. Each type of plate boundary has signature geographical features and signature patterns of earthquake foci and epicenters.

Students will be expected to know about continental movements and that drift has occurred in the past. Students will not be assessed on their knowledge of how continents looked in the past.

Examples of Performance Expectations

<u>Identifying Science Principles.</u> Students can:

- State the types of plate boundaries associated with given patterns of earthquake foci.

<u>Using Scientific Inquiry.</u> Students can:

- Find patterns in data about a set of earthquakes. Relate these patterns to plate boundary models.

Performance expectations for Using Science Principles and Using Technological Design are not provided.

Item to Assess Identifying Science Principles

Item Suggestion

Which of the following plate boundaries is most likely to be the location of shallow focus only earthquakes?

A. Collision between two continental plates
B. Collision between a continental plate and an oceanic plate
C. Spreading of two oceanic plates
D. Island hot spots

Key: C

Interpretation: The student needs prior knowledge that spreading (Sp) or slip-slide (SS) fault boundaries are characterized by shallow focus earthquakes. Possible naïve conceptions that students might hold include (1) “epicenter” and “focus” are two labels for the location of an earthquake, and (2) earthquakes only occur on land.

Items to Assess Using Scientific Inquiry

Item Suggestion 1

The earthquake data in the table below were reported by the U.S. Department of the Interior for one location over a two-week period.

Magnitude Richter Scale	Foci Shallow (0-70km) Medium (70-350km) Deep (350-700km)	Epicenter Latitude/ Longitude (Numbers for this column would need to be provided in an actual assessment item.)
5.6	66	
4.5	120	
3.4	223	
4.2	55	
6.0	467	
3.5	309	
4.0	49	
3.8	79	
4.3	99	

Items to Assess Using Scientific Inquiry (cont.)

Based on the above data, which is the most likely location for this pattern of earthquakes?

A. Mid-oceanic ridge
B. Island arc with volcanic activity
C. High altitude mountains
D. Trailing edge of a lithospheric plate

Key: B

Interpretation:

Prior knowledge that students need:

- Spreading (Sp) or slip-slide(SS) fault boundaries are characterized by shallow focus earthquakes with epicenters occurring along linear patterns.
- Continental crust/continental crust (CC) collision boundaries are characterized by shallow and medium focus earthquakes with epicenters occurring in a scattered pattern.
- Continental crust/oceanic crust (CO) collision boundaries are characterized by shallow, medium, and deep focus earthquakes with epicenters occurring in linear patterns and with a sequence pattern of shallow, medium, and deep.
  - Students need to understand the distinction between the focus and epicenter of an earthquake.
  - Students need to understand that a less dense continental crust plate overrides the more dense oceanic crust plate.
  - The descending oceanic crust plate causes medium and deep focus earthquakes and friction heat to produce magma.
  - Deep sea trenches and volcanoes are physiographic features of CO plate boundaries.
  - Mountains are physiographic features of CC plate boundaries.

Possible naïve conception that students might hold:

- Lithosopheric plates all have the same composition and density.
- Epicenter and focus are two labels for the location of an earthquake.
- Earthquakes only occur on land.

<u>Item Suggestion 2</u>

Carefully examine the map projection below and the accompanying key. Based on the information presented, provide an explanation for the pattern of earthquakes and the presence of volcanoes. You may write your answer or do a labeled drawing.

Include in the stem of the question a map projection of a region of the world that experiences both earthquake and volcanic activity. The projection would be one that includes physiographic features of the bottom of the ocean floor (e.g., trenches, mid-ocean ridges, transform faults) and/or continental physiographic features (e.g., island arcs, volcanoes). Plate boundaries with both earthquake and volcanic features include oceanic/oceanic spreading plate boundaries (e.g., Iceland along the mid-ocean ridge) and oceanic/continental collision plate boundaries (e.g., island arcs that comprise the Pacific 'ring of fire'; Mediterranean region). An alternative option would be to show two projections—one reflecting the ascending constructive plate boundary at the mid-ocean spreading ridges and one reflecting the descending destructive plate boundary at deep ocean trenches—and ask students to explain the differences in kinds of earthquakes and physiographic features at each location.

Items to Assess Using Scientific Inquiry (cont.)

Interpretation for Mid-Ocean Ridge Map Projection: The key is to recognize the mid-ocean ridge plate boundary as a constructive boundary, where rising magma moves up and out, away from the spreading zone. Earthquakes are shallow focus and the epicenters occur along the axis of the mid-ocean ridges. The rising magma causes the volcanic eruptions that form the mid-ocean ridges. Students at the upper end of the achievement scale would need to identify the linear earthquakes occurring along the spreading ridges and identify the earthquakes as being shallow focus earthquakes. Both plates would be identified as oceanic plates moving apart from one another. Magma from the ascending or rising convection currents in the mantle and crust would explain the volcanic activity along the mid-ocean ridge. The formation of the mid-ocean ridges splitting and pushing oceanic plates on the ocean floor would be another element of an answer at the upper end of the achievement scale. Mention of transform faults and ocean-floor magnetic reversals would be indication of a superior response.

Interpretation for Island Arc Map Projection: The key is to recognize the island arc boundary as a destructive interaction between a continental plate and an oceanic plate. The oceanic plate descends because it is denser. The descending oceanic plate is what causes the earthquake. Earthquakes progress from shallow to deep focus, with shallow focus earthquakes located adjacent to the trench and deep focus earthquakes typically located adjacent to the island arc. The friction caused by the descending oceanic plate produces magma that rises to form volcanic eruptions or granitic mountains. Students at the upper end of the achievement scale would need to identify the earthquakes occurring along the island arcs as shallow and deep and relate the type of earthquakes to the colliding and descending oceanic plate. The descending plate would also explain deep-sea trenches. The friction caused by the descending oceanic plates would be identified as the source of magma for volcanic activity on island arcs. Recognition of the oceanic plate as being denser would be evidence of a superior response.

See also "Interpretation" for Using Scientific Inquiry, Item Suggestion 1.

Item Suggestion 3

> Examine the Earth cross-section below. Each dot represents the location of an earthquake along an east-west section of the Earth 50 miles wide. How does the theory of plate tectonics account for this pattern of earthquakes? Explain in writing.

(This example is the sideways projection of the island arc projection. Earthquakes would be shown in a pattern like that below—such patterns can be found on the Internet—and labeled with scales to show depth and horizontal scale.)

```
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ surface of Earth
O
         O  O
                  OO
         O                 O                  depth
                  O                 O  O
                              O     O     O
                           O
                                          O    O
```

Interpretation: See "Interpretation" for Using Scientific Inquiry, Item Suggestions 1 and 2 above.

Illustrative items to assess Using Science Principles and Using Technological Design are not provided.

Energy in Earth Systems

Grade 8: Earth Systems—Energy in Earth Systems

Content Statements

E8.11: The sun is the major source of energy for phenomena on Earth's surface. The sun provides energy for plants to grow and drives convection within the atmosphere and oceans, producing winds, ocean currents, and the water cycle.

E8.12: Seasons result from annual variations in the intensity of sunlight and length of day, due to the tilt of Earth's rotation axis relative to the plane of its yearly orbit around the sun.

No commentary on these content statements.

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Identify regions of the Earth's surface (e.g., land, ocean, ice caps) that are better/worse at absorbing and releasing the sun's energy.
- Define convection and its role in producing winds and ocean currents.

Using Science Principles. Students can:

- Given a map with geographical features, predict where convection cells would occur.
- Explain the surface circulation patterns of ocean currents.
- Explain the cause of seasons.
- Predict effects of changes in the Earth's tilt.

Using Scientific Inquiry. Students can:

- Design or critique investigations to determine how the height of the sun at different times during the day changes from season to season.
- Given data on annual trends of incoming solar radiation for several cities, determine whether the cities are located in the Northern Hemisphere or Southern Hemisphere.

Using Technological Design. Students can:

- Given a map of terrain in a specific area, design a house that would take advantage of sunlight so it is warm in the winter (saving on heating bills) and cool in the summer (saving on air conditioning bills).

Items to Assess Using Science Principles

Item Suggestion 1

Predict how seasons would be affected if Earth's axis were perpendicular to the plane of its orbit around the sun.

Item Suggestion 2

Explain why the amount of sunlight changes from one season to another.

Item Suggestion 3

Explain why seasons are opposite in the Northern and Southern Hemispheres.

Item to Assess Using Scientific Inquiry Item Suggestion Design an investigation to determine how the height of the sun at noon changes from day to day.
Items to assess Identifying Science Principles and Using Technological Design are not provided.

Biogeochemical Cycles

Grade 8: Earth Systems—Biogeochemical Cycles
Content Statement **E8.15:** Human activities, such as reducing the amount of forest cover, increasing the amount and variety of chemicals released into the atmosphere, and intensive farming, have changed Earth's land, oceans, and atmosphere. Studies of plant and animal populations have shown that such activities can reduce the number and variety of wild plants and animals and sometimes result in the extinction of species.
Commentary The following sequence illustrates one aspect of human activity related to changes in Earth's land, oceans, and atmosphere, namely, global warming. The sun's energy reaches Earth and warms the land, water, and air. That energy is trapped in the atmosphere. When there is a change in the percentage of certain gases in the atmosphere, such as carbon dioxide, it changes the amount of energy trapped. Burning more fossil fuels releases not just carbon dioxide but also particulate matter into the atmosphere. Particles also trap some of the radiation that comes from Earth. Carbon dioxide transmits visible light from the sun, but absorbs infrared radiation that rises from the ground and holds this heat in the atmosphere instead of reflecting it back to space. As Earth warms, glaciers recede and pack ice melts. Ocean temperatures and currents change, altering habitats around the world. The recent rapid rate of change in the warming of Earth leads scientists to hypothesize that global warming is correlated with human activity. As climates and conditions change, ecological niches (terrestrial or aquatic) change. Plants and animals are affected by these changes. Some plant and/or animal populations are forced to migrate to other ecosystems, adapt, or die. Within the altered ecosystems, plants and animals are forced to compete for resources. Thus, as habitats change, certain species flourish and others can no longer survive.
Examples of Performance Expectations Identifying Science Principles. Students can: • Identify some of the causes of global warming. • Describe some of the changes that occur in ecosystems as climates and conditions change. • State some of the changes that may occur in the oceans as Earth warms. Using Science Principles. Students can: • Predict changes in animal and plant populations based on descriptions of climatic changes. • Explain how global warming could result in the extinction of a species. • Predict the impact of the melting of glaciers on coastal settlements.

Examples of Performance Expectations (cont.)

Using Scientific Inquiry. Students can:

- Given data on climate change due to global warming and change in the number of a particular species, critique possible explanations.
- Identify patterns in various climate indicators such as maximum and minimum temperatures, rainfall, and cloud cover from data spanning 10 years.
- Given data spanning several decades, propose relationships between ocean temperatures and frequency of hurricanes and between ocean temperatures and strength of hurricanes.
- Critique a scientific investigation that monitors changes in plants and animals and climate indicators such as temperature.

Using Technological Design. Students can:

- Describe measures that can be taken to reduce global warming.
- Discuss the scientific trade-offs of a measure taken to reduce global warming.
- Discuss two alternatives to the burning of fossil fuels that provide energy for human needs.

Items to Assess Identifying Science Principles

Illustrative Item

The burning of fossil fuels has increased the carbon dioxide content of the atmosphere. What is a possible effect that the increased amount of carbon dioxide is likely to have on our planet?

A. A warmer climate
B. A cooler climate
C. Lower relative humidity
D. More ozone in the atmosphere

Key: A
Source: TIMSS 2003, Grade 8

Item Suggestion

Melting of major ice shelves at the south pole will cause

A. changes in atmospheric temperature.
B. rising sea level.
C. ocean waves.
D. rising ocean temperatures.

Key: B

Item to Assess Using Science Principles

Item Suggestion

Scientists are concerned that the average temperature of Earth is increasing a little bit every year because of

A. the hole in the ozone layer.
B. fossil fuel burning that releases excess heat.
C. a gradual increase in the temperature of the sun.
D. fossil fuel use that releases carbon dioxide.

Key: D

Interpretation: A very common naïve conception is option A. Although the reduction of ozone in the stratosphere permits slightly more solar radiation to reach the surface, the effect is very small compared with the property of carbon dioxide to absorb heat (infrared) energy. Consequently, to successfully answer this question, students need to apply what they know about how the atmosphere maintains a nearly constant liveable temperature on Earth.

Item to Assess Using Technological Design

Item Suggestion

In order to reduce the use of fossil fuels and lessen the dangers associated with global warming, the National Park Service decided to build a dam and hydroelectric power station in a national forest. People who lived in the local area objected, noting that when the dam is built, it would flood a river valley and form a large human-made lake. In your opinion, what are the scientific trade-offs that the National Park Service should consider before it goes ahead?

Interpretation: In order to understand the advantages of the proposed dam, students need to understand how the burning of fossil fuels may bring about global warming. In order to understand the disadvantages of the proposed dam, students need to understand now natural ecosystems provide services for humans and other animal life. A successful answer will view the trade-off as being between reduction of global warming on the one hand, and the services of the natural ecosystem that will be submerged on the other.

Illustrative items to assess Using Scientific Inquiry are not provided.

Grade 12: Earth Systems – Biogeochemical Cycles

Content Statement

E12.13: Natural ecosystems provide an array of basic processes that affect humans. These processes include maintenance of the quality of the atmosphere, generation of soils, control of the hydrologic cycle, disposal of wastes, and recycling of nutrients.

Commentary

This content statement calls on high school students to integrate a wide range of information and ideas from the physical, life, and Earth sciences to recognize how natural systems make human life possible. Emphasis is on the positive ways that natural systems provide an environment that supports living systems including humans, although human activities can in turn affect the functioning of natural ecosystems. Since each of the examples given in the statement can be developed into an entire course of study, it will be important to limit the scope of the assessment to just the major ideas, such as the following:

Commentary (cont.)

- Maintain the quality of the atmosphere: Focus on how the short-term carbon cycle allows for the exchange of gases between plants and animals, how rain helps to remove pollutants from the atmosphere, and how the presence of carbon dioxide and water vapor provides a "natural greenhouse effect" to maintain the climate at a liveable temperature.
- Generation of soils: Include physical processes (e.g., weathering, erosion, deposition) that provide the inorganic constituents of soils, as well as decomposition processes that provide the organic constituents of soils.
- Control of the hydrologic cycle: Expand on the concept of the water cycle that was introduced in earlier grades by including the effects of mountain ranges, rainforests, and other natural environments on the water cycle.
- Disposal of wastes: Focus on how natural environments dispose of plant and animal wastes through processes such as decay and decomposition and how these systems may be overwhelmed where human populations expand rapidly (e.g., in cities).
- Recycling of nutrients: Emphasize that plants and animals in an ecosystem absorb nutrients from the air, water, and soil and by consuming other organisms. These nutrients are eventually returned to the environment with the assistance of decomposers so that other organisms can in turn absorb the nutrients.

It is important to distinguish this content statement from a similar statement found in the introduction to the Life Science section of the *Framework* and *Specifications*, "The species interaction in an ecosystem, the dynamics of population growth and decline, the use of resources by multiple species, their impact on their environment, and the complex interactions among all of these have enormous consequences to the survival of all species, including humans" (*Specifications*, p. 53).

Much more could be included in this content statement's list of natural processes. For example, "maintain the quality of the atmosphere" could be expanded to include the long-term carbon cycle and various feedback mechanisms. While important in the study of Earth sciences, these related topics are not recommended for the NAEP Science Assessment, which should focus on the key ideas outlined above and in the content boundaries found in Chapter Two.

Examples of Performance Expectations

Identifying Science Principles. Students can:

- Describe how the short-term carbon cycle recycles atmospheric gases needed by plants and animals to survive.

Using Science Principles. Students can:

- Provide three possible reasons that the soil layer may be thicker in some locations and thinner in others.
- Explain why people do not typically smell rotten garbage in natural woods or meadows, even though plant matter dies on a regular basis.
- Explain how natural systems purify water.
- Describe how allowing a farmland to lay fallow for a period of years restores productivity after intensive farming.

Using Scientific Inquiry. Students can:

- Design a series of systematic observations to determine how changes in a forest ecosystem result in changes in the local atmosphere over time.
- Describe a procedure for determining the impact of chemical pollution in a stream.
- Propose a testable hypothesis for why animals live longer (or shorter) in one environment compared to animals of the same species in another environment.

Examples of Performance Expectations (cont.)

Using Technological Design. Students can:

- Apply knowledge of how soils are produced to develop a process for changing the properties of soil to improve plant growth.
- Propose feasible solutions to prevent the depletion of oxygen in waters, given information about how such zones result from the use of fertilizers elsewhere. Use the concept of scientific trade-offs to select which of two possible solutions is most promising.

Item to Assess Using Science Principles

Item Suggestion

A student learned two things that seemed to contradict each other: (1) river water is sometimes polluted, especially downstream of cities and manufacturing regions; and (2) in the water cycle, all river water eventually evaporates and falls as rain, and rainwater is nearly pollution-free. What would you say to the student to show that both statements are true and not contradictory?

Interpretation: To reconcile these statements, students must understand how natural systems (the water cycle) purify water.

Item to Assess Using Scientific Inquiry

Item Suggestion

Propose a testable hypothesis to explain why squirrels that live in a city tend to live shorter lives than squirrels that live in a forest.

Interpretation: In order to come up with a reasonable hypothesis, students need to recognize ways in which natural systems provide the conditions that animals need to survive and that some of these conditions may not be present in cities. In addition, students should be able to state their hypothesis in testable terms, that is, in terms that can be objectively and quantitatively observed and recorded. The ability to develop a testable hypothesis is an important component of the Using Scientific Inquiry practice.

Items to Assess Using Technological Design

Item Suggestion 1

News item: Farmers in the Midwestern United States use nitrogen-based fertilizers to encourage growth of crops. Rain dissolves any fertilizers not taken up by plants and washes the excess nitrogen into the Mississippi River. This nitrogen is eventually discharged into the Gulf of Mexico. Algae thrive on the additional nitrogen and expand in numbers. The huge algae population absorbs nearly all of the water's oxygen, causing shrimp and other aquatic life that depend on dissolved oxygen to suffocate.

a. Propose at least three solutions for this problem.
b. List at least two criteria for choosing the best solution.
c. Choose the best solution and explain why. Use the concept of scientific "trade-off" in your answer.

Items to Assess Using Technological Design (cont.) Interpretation: In order to make sense of this item, students need to understand the short-term carbon cycle in an aquatic environment. A feasible solution could include stopping the nitrogen before it enters the Mississippi River, before the nitrogen-rich water is discharged into the Gulf of Mexico, or before the algae absorb it. Alternatively, efforts may be made to control the reproduction of algae or to oxygenate the water. Other solutions may be proposed as well. These ideas should be clearly stated and distinct from one other. Criteria for determining the most successful solution should include measurable ways to judge success (e.g., a decrease in the nitrogen content of the water, an increase in the dissolved oxygen content of the water, an increase in the fish population). Students should use a systematic approach when applying the criteria to the potential solutions. Overall, students' ideas should make sense in terms of their knowledge of how natural ecosystems provide an array of basic processes that affect humans. Item Suggestion 2 List at least five components that an enclosed city on the moon should have in order for it to rain occasionally.
Items to assess Identifying Science Principles are not provided.

APPENDIX I

EXAMPLES OF LEARNING PROGRESSIONS AND RELATED ITEMS

An example of a learning progression developed for a curriculum on floating and sinking is shown in Table 31 (Kennedy, Brown, Draney, & Wilson, 2005; see also Shavelson, 2006).[48] This simple example of a learning progression in a specific, well-researched area (expressed as a scoring guide) has been chosen to illustrate how items can be related to learning progressions and how they might be used in the NAEP Science Assessment.

Table 31. Example of a Learning Progression for Floating and Sinking
(adapted from Kennedy et al., 2005)

<table>
<tr><th colspan="2">Level</th><th colspan="2">What the Student Knows</th><th>Example Responses
(“why things sink and float” essay)</th></tr>
<tr><td colspan="2">RD</td><td colspan="2">Relative Density
Student knows that floating depends on having less density than the medium.</td><td>“An object floats when its density is less than the density of the medium.”</td></tr>
<tr><td colspan="2">D</td><td colspan="2">Density
Student knows that floating depends on having a small density.</td><td>“An object floats when its density is small.”</td></tr>
<tr><td colspan="2">MV</td><td colspan="2">Mass and Volume
Student knows that floating depends on having a small mass and a large volume.</td><td>“An object floats when its mass is small and its volume is large.”
“An object that is light for its size will float.”</td></tr>
<tr><td>M</td><td>V</td><td>Mass*
Student knows that floating depends on having a small measured mass.</td><td>Volume
Student knows that floating depends on having a large measured volume.</td><td>“An object floats when its mass is small.”
“An object floats when its volume is large.”</td></tr>
<tr><td colspan="2">PM</td><td colspan="2">Productive Misconception
Student thinks that floating depends on felt weight (heft), a small size or amount, or that it depends on being made out of a particular material.</td><td>“An object floats when it is small.”
“An object floats when it is a light material.”</td></tr>
<tr><td colspan="2">UF</td><td colspan="2">Unconventional Feature
Student thinks that floating depends on being flat, hollow, filled with air, or having holes.</td><td>“An object floats when it is hollow.”
“An object floats when it is flat.”</td></tr>
<tr><td colspan="2">OT</td><td colspan="2">Off Target
Student does not attend to any property or feature to explain floating.</td><td>“I have no idea.”</td></tr>
<tr><td colspan="2">NR</td><td colspan="2">No Response
Student left the response blank.</td><td>[blank]</td></tr>
</table>

*Weight is acceptable terminology for students at grade 4.

[48] See grade 8 “Properties of Matter” content boundaries on measurement of density and relative density.

Learning progressions can be used as a guide to identify points at which assessment would be useful to locate a student's position on a learning progression. Items may be developed that are related to a learning progression in two ways. In the first case, items are designed to be an indicator of whether or not a student's understanding has reached a certain level on the learning progression. For example, Item A below would be an indicator that a student had reached the relative density level of understanding (the RD level in Table 31).

Item A. Example of an Item that Targets the Relative Density Level of the Learning Progression

George has a block with a density of 1.1 g/cm^3. He puts the block in salt water with density of 1.2 g/cm^3. What will the block do?

A. Float
B. Subsurface float
C. Sink
D. Not enough information to tell

Key: A

P8.4, Using Science Principles

Item B below (Yin, 2005) is an example of an item that would be an indicator of the next level below, that is, if a student has an understanding of density, the D level on the learning progression. It would not, however, tell whether the student understands relative density (RD).

Item B. Example of an Item that Targets the Density Level of the Learning Progression

Which object listed in the table has the greatest density?

Object	Mass of Object	Volume of Object
W	11.0 grams	24 cubic centimeters
X	11.0 grams	12 cubic centimeters
Y	5.5 grams	4 cubic centimeters
Z	5.5 grams	11 cubic centimeters

A. W
B. X
C. Y
D. Z

Key: C

P8.4, Using Scientific Inquiry

One way to use items like these within a grade level tested on NAEP is as follows: A series of individual items such as Items A and B can be used to test which levels of understanding on the learning progression students have achieved. In addition, when sets of items are released, an illustration of the learning progression could be represented by the locations of the items on an item map (assuming that they come up in the expected order). The learning progression contributes to the quality of the item set since the items are founded upon a research-based learning progression.

The second way that items can be related to a learning progression is to create items that span multiple levels of the learning progression. The items may be selected-response or constructed-response items. For example, Item C below elicits a range of responses that can be scored using a rubric that is based upon the learning progression (Shavelson, 2006; Yin, 2005). In fact, Table 31 is a generalized form of the rubric for Item C.

Item C. Example of a Constructed Response Item That Spans Multiple Levels of the Learning Progression

Why do things sink or float?
Write as much information as you need to explain your answer. Use evidence, examples, and what you have learned to support your explanations.

P8.4, Using Science Principles

In a similar way, clusters of ordered multiple-choice items (Briggs et al., 2006) may be used to locate the level of understanding a student has reached regarding Earth and the solar system. An example of how this may be accomplished is shown in Table 32 and Item Set D. Table 32 shows a simple learning progression (called a construct map by Briggs et al., 2006) that represents how student understanding of the motions of the Earth, moon, and sun in the solar system develops over time. Note that the top level of the learning progression is designed to approximate the understanding typical of 8th graders in national standards documents. Similarly, the fourth level of the learning progression is designed to approximate the understanding typically expected at the 5th grade level. Because students' understanding of the Earth and the solar system develops throughout their schooling, the learning progression is designed to describe the understandings of both 5th and 8th grade students. However, the top level is not expected of 5th graders, and it is unlikely that many 8th grade students would fall into the lowest levels of the continuum.

Table 32: Learning Progression for Student Understanding of Earth in the Solar System (from Briggs et al., 2006)

Level	Description
5 (grade 8)	Student is able to put the motions of the Earth and Moon into a complete description of motion in the Solar System which explains: • the day/night cycle • the phases of the Moon (including the illumination of the Moon by the Sun) • the seasons
4 (grade 5)	Student is able to coordinate apparent and actual motion of objects in the sky. Student knows that: • the Earth is both orbiting the Sun and rotating on its axis • the Earth orbits the Sun once per year • the Earth rotates on its axis once per day, causing the day/night cycle and the appearance that the Sun moves across the sky • the Moon orbits the Earth once every 28 days, producing the phases of the Moon COMMON ERROR: Seasons are caused by the changing distance between the Earth and Sun. COMMON ERROR: The phases of the Moon are caused by a shadow of the planets, the Sun, or the Earth falling on the Moon.
3	Student knows that: • the Earth orbits the Sun • the Moon orbits the Earth • the Earth rotates on its axis However, student has not put this knowledge together with an understanding of apparent motion to form explanations and may not recognize that the Earth is both rotating and orbiting simultaneously. COMMON ERROR: It gets dark at night because the Earth goes around the Sun once a day.
2	Student recognizes that: • the Sun appears to move across the sky every day • the observable shape of the Moon changes every 28 days Student may believe that the Sun moves around the Earth. COMMON ERROR: All motion in the sky is due to the Earth spinning on its axis. COMMON ERROR: The Sun travels around the Earth. COMMON ERROR: It gets dark at night because the Sun goes around the Earth once a day. COMMON ERROR: The Earth is the center of the universe.
1	Student does not recognize the systematic nature of the appearance of objects in the sky. Students may not recognize that the Earth is spherical. COMMON ERROR: It gets dark at night because something (e.g., clouds, the atmosphere, "darkness") covers the Sun. COMMON ERROR: The phases of the Moon are caused by clouds covering the Moon. COMMON ERROR: The Sun goes below the Earth at night.
0	No evidence or off-track

Item Set D shows two ordered multiple-choice items that are carefully constructed so that the answer choices map to different levels of the learning progression shown in Table 32. Note that each item contains answer choices that map differently to the progress variable: the first item's answer choices map to levels 1, 2, 3 and 4, whereas the second item's answer choices map to levels 3, 4, and 5. The first item is focused at the 5th grade because it contains answer choices up to and including a 5th grade level response.[49] The second item is focused at the 8th grade level because the answer choices cover typical understandings and common errors at that level and below. Accordingly, between the two items, there is coverage of the whole learning progression, except for level 0, which is one that is more likely to be addressed in a constructed-response item rather than a selected-response item where answer choices should all sound plausible to an uninformed test taker.

Item Set D: Example of Two Ordered Multiple-Choice Items Based Upon the Earth in the Solar System Progress Variable (adapted from Briggs et al., 2006)

It is most likely colder at night because

A.	the Earth is at the farthest point in its orbit around the Sun.	Level 3
B.	the Sun has traveled to the other side of the Earth.	Level 2
C.	the Sun is below the Earth and the Moon does not emit as much heat as the Sun.	Level 1
D.	the place where it is night on Earth is rotated away from the Sun.	Level 4

E8.2, Using Science Principles
© WestEd, 2002

Which is the best explanation for why we experience different seasons (winter, summer, etc.) on Earth?

A.	The Earth's orbit around the Sun makes us closer to the Sun in summer and farther away in winter.	Level 4
B.	The Earth's orbit around the Sun makes us face the Sun in the summer and away from the Sun in the winter.	Level 3
C.	The Earth's tilt causes the Sun to shine more directly in summer than in winter.	Level 5
D.	The Earth's tilt makes us closer to the Sun in summer than in winter.	Level 4

E8.12, Using Science Principles
© WestEd, 2002

These types of items can be used both within a given grade level as well as across grade levels (e.g., at grades 8 and 12 for Item Set D). Administration of items across grade levels allows for comparisons of success rates.

[49] The first item in Item Set D goes beyond the grade 4 expectations described in the *Framework* and *Specifications*. NAEP expects an Earth-centric perspective through grade 4 (see Chapter Two's content boundaries for "Objects in the Universe").

Table 33 uses a different format to illustrate another learning progression. Using a set of related science content statements across grades 4, 8, and 12, it follows the format of the science practices defined in the *Framework* and *Specifications.* It illustrates how relevant research (e.g., Smith et al., 2004) can be used as an opportunity to assess content in greater depth—available research on student learning is used to inform the generation of related performance expectations across grades. The table includes examples of performance expectations for a possible learning progression for States of Matter. These *illustrative* performance expectations are not intended to denote a sense of content priority or importance, nor should they be interpreted as a complete representation of the research currently available.

Table 33. Examples of Performance Expectations for States of Matter

Grade 4 (See content statement P4.3.)	**Grade 8** (See content statement P8.1.)	**Grade 12** (See content statement P12.1.)
Identifying Science Principles Classify samples of material as solid, liquid, or gas.	**Identifying Science Principles** Given an animation of molecules in motion, identify the substance that is being illustrated as a solid, liquid, or gas.	**Identifying Science Principles** Explain why ice is harder than liquid water in terms of the strength of the force between the molecules.
Using Science Principles Infer that a change of state (e.g., freezing or melting) affects the identity of an object, but not the identity of the material of which it is made.	**Using Science Principles** Predict how the mass of a sample of iodine will change after sublimation. Justify the prediction based on what occurs during sublimation at a molecular level.	**Using Science Principles** Use the concept of molecular arrangements and bonds to explain why graphite is very soft and diamond is very hard, even though they are both made of pure carbon.
Using Scientific Inquiry Collect, display, and interpret data showing how the temperature of a substance changes over time as it cools and becomes a solid.	**Using Scientific Inquiry** Plan and conduct an investigation to determine the melting point and boiling point of an unknown substance.	**Using Scientific Inquiry** Using molecular theory, explain the results of experiments showing how the volume of three different liquids changes when they are heated.
Using Technological Design Propose a method for determining for certain whether holiday chocolates that have been shaped by different processes (melting, freezing, reshaping, or breaking into pieces) have the same amount of chocolate in them.	**Using Technological Design** Choose the best solution for increasing the altitude of a hot air balloon, based on an understanding of the macroscopic and microscopic changes that occur when the gas inside the balloon is heated.	**Using Technological Design** Design an instrument to measure temperature as accurately as possible, taking into account both the thermal properties of liquids and solids to be used in the device, and structural shape and dimensions of the device.

APPENDIX J

NAEP ITEM DEVELOPMENT AND REVIEW POLICY STATEMENT

Adopted: May 18, 2002

National Assessment Governing Board

NAEP Item Development and Review Policy Statement

It is the policy of the National Assessment Governing Board to require the highest standards of fairness, accuracy, and technical quality in the design, construction, and final approval of all test questions and assessments developed and administered under the National Assessment of Educational Progress (NAEP). All NAEP test questions or items must be designed and constructed to reflect carefully the assessment objectives approved by the National Assessment Governing Board. The final assessments shall adhere to the requirements outlined in the following Guiding Principles, Policies and Procedures for NAEP Item Development and Review.

The Governing Board's Assessment Development Committee, with assistance from other Board members as needed, shall be responsible for reviewing and approving NAEP test questions at several stages during the development cycle. In so doing, the Guiding Principles, Policies and Procedures must be adhered to rigorously.

Introduction

The No Child Left Behind Act of 2001 (P.L. 107-110) contains a number of important provisions regarding item development and review for the National Assessment of Educational Progress (NAEP). The legislation requires that:

- "the purpose [of NAEP] is to provide…a fair and accurate measurement of student academic achievement"
- "[NAEP shall]…use widely accepted professional testing standards, objectively measure academic achievement, knowledge, and skills, and ensure that any academic assessment authorized….be tests that do not evaluate or assess personal or family beliefs and attitudes or publicly disclose personally identifiable information;"
- "[NAEP shall]…only collect information that is directly related to the appraisal of academic achievement, and to the fair and accurate presentation of such information;"
- "the Board shall develop assessment objectives consistent with the requirements of this section and test specifications that produce an assessment that is valid and reliable, and are based on relevant widely accepted professional standards;"
- "the Board shall have final authority on the appropriateness of all assessment items;"

- "the Board shall take steps to ensure that all items selected for use in the National Assessment are free from racial, cultural, gender, or regional bias and are secular, neutral, and non-ideological;" and
- "the Board shall develop a process for review of the assessment which includes the active participation of teachers, curriculum specialists, local school administrators, parents, and concerned members of the public."

Given the importance of these mandates, it is incumbent upon the Board to ensure that the highest standards of test fairness and technical quality are employed in the design, construction, and final approval of all test questions for the National Assessment. The validity of educational inferences made using NAEP data could be seriously impaired without high standards and rigorous procedures for test item development, review, and selection.

Test questions used in the National Assessment must yield assessment data that are both valid and reliable in order to be appropriate. Consequently, technical acceptability is a necessary, but not a sufficient condition, for judging the appropriateness of items. In addition, the process for item development must be thorough and accurate, with sufficient reviews and checkpoints to ensure that accuracy. The Guiding Principles, Policies, and Procedures governing item development, if fully implemented throughout the development cycle, will result in items that are fair and of the highest technical quality, and which will yield valid and reliable assessment data.

Each of the following Guiding Principles is accompanied by Policies and Procedures. Full implementation of this policy will require supporting documentation from the National Center for Education Statistics (NCES) regarding all aspects of the Policies and Procedures for which they are responsible.

This policy complies with the documents listed below which express acceptable technical and professional standards for item development and use. These standards reflect the current agreement of recognized experts in the field, as well as the policy positions of major professional and technical associations concerned with educational testing.

Standards for educational and psychological testing. (1999). Washington, DC: American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME).

Code of fair testing practices in education. (1988). Washington, DC: Joint Committee on Testing Practices.

National Center for Education Statistics (NCES) Statistical Standards, DRAFT, February 2002.

Guiding Principles – Item Development and Review Policy

Principle 1
NAEP test questions selected for a given content area shall be representative of the content domain to which inferences will be made and shall match the NAEP assessment framework and specifications for a particular assessment.

Principle 2
The achievement level descriptions for basic, proficient, and advanced performance shall be an important consideration in all phases of NAEP development and review.

Principle 3
The Governing Board shall have final authority over all NAEP test questions. This authority includes, but is not limited to, the development of items, establishing the criteria for reviewing items, and the process for review.

Principle 4
The Governing Board shall review all NAEP test questions that are to be administered in conjunction with a pilot test, field test, operational assessment, or special study administered as part of NAEP.

Principle 5
NAEP test questions will be accurate in their presentation and free from error. Scoring criteria will be accurate, clear, and explicit.

Principle 6
All NAEP test questions will be free from racial, cultural, gender, or regional bias, and must be secular, neutral, and non-ideological. NAEP will not evaluate or assess personal or family beliefs, feelings, and attitudes, or publicly disclose personally identifiable information.

Policies and Procedures for Guiding Principles

Principle 1

NAEP test questions selected for a given content area shall be representative of the content domain to which inferences will be made and shall match the NAEP assessment framework and specifications for a particular assessment.

Policies and Procedures

1. Under the direction of the Board, the framework for each assessment will be developed in a manner that defines the content to be assessed, consistent with NAEP's purpose and the context of a large-scale assessment. The framework development process shall result in a rationale for each NAEP assessment, which delineates the scope of the assessment relative to the content domain. The framework will consist of a statement of purpose, assessment objectives, format requirements, and other guidelines for developing the assessment and items.
2. In addition to the framework, the Board shall develop assessment and item specifications to define the: a) content and process dimensions for the assessment; b) distribution of items across content and process dimensions at each grade level; c) stimulus and response attributes (or what the test question provides to students and the format for answering the item); d) types of scoring procedures; e) test administration conditions; and f) other specifications pertaining to the particular subject area assessment.
3. The Board will forward the framework and specifications to NCES, in accordance with an appropriate timeline, so that NCES may carry out its responsibilities for assessment development and administration.
4. In order to ensure that valid inferences can be made from the assessment, it is critical that the pool of test questions measures the construct as defined in the framework. Demonstrating that the items selected for the assessment are representative of the subject matter to which inferences will be made is a major type of validity evidence needed to establish the appropriateness of items.
5. A second type of validity evidence is needed to ensure that NAEP test items match the specific objectives of a given assessment. The items must reflect the objectives, and the item pool must match the percentage distribution for the content and cognitive dimensions at each grade level, as stated in the framework. Minor deviations, if any, from the content domain as defined by the framework will be explained in supporting materials.
6. Supporting material submitted with the NAEP items will provide a description of procedures followed by item writers during development of NAEP test questions. This description will include the expertise, training, and demographic characteristics of the groups. This supporting material must show that all item writing and review groups have the required expertise and training in the subject matter, bias, fairness, and assessment development.

7. In submitting items for review by the Board, NCES will provide information on the relationship of the specifications and the content/process elements of the pool of NAEP items. This will include procedures used in classifying each item.
8. The item types used in an assessment must match the content requirements as stated in the framework and specifications, to the extent possible. The match between an objective and the item format must be informed by specifications pertaining to the content, knowledge or skill to be measured, cognitive complexity, overall appropriateness, and efficiency of the item type. NAEP assessments shall use a variety of item types as best fit the requirements stated in the framework and specifications.
9. In order to ensure consistency between the framework and specifications documents and the item pools, NCES will ensure that the development contractor engages a minimum of 20% of the membership of the framework project committees in each subject area to serve on the item writing and review groups as the NAEP test questions are being developed. This overlap between the framework development committees and the item developers will provide stability throughout the NAEP development process, and ensure that the framework and specifications approved by the Board have been faithfully executed in developing NAEP test questions.

Principle 2

The achievement level descriptions for basic, proficient, and advanced performance shall be an important consideration in all phases of NAEP development and review.

Policies and Procedures

1. During the framework development process, the project committees shall draft preliminary descriptions of the achievement levels for each grade to be assessed. These preliminary descriptions will define what students should know and be able to do at each grade, in terms of the content and process dimensions of the framework at the basic, proficient, and advanced levels. Subsequent to Board adoption, the final achievement level descriptions shall be an important consideration in all future test item development for a given subject area framework.
2. The achievement level descriptions will be used to ensure a match between the descriptions and the resulting NAEP items. The achievement level descriptions will be examined, and appropriate instruction provided to item writers to ensure that the items represent the stated descriptions, while adhering to the content and process requirements of the framework and specifications. The descriptions will be used to evaluate the test questions to make certain that the pool of questions encompasses the range of content and process demands specified in the achievement level descriptions, including items within each achievement level interval, and items that scale below basic.

3. As the NAEP item pool is being constructed, additional questions may need to be written for certain content/skill areas if there appear to be any gaps in the pool, relative to the achievement level descriptions.
4. Supporting materials will show the relationship between the achievement levels descriptions and the pool of NAEP test questions.

Principle 3

The Governing Board shall have final authority over all NAEP test questions. This authority includes, but is not limited to, the development of items, establishing the criteria for reviewing items, and the process for review.

Policies and Procedures

1. Under the No Child Left Behind Act, a primary duty of the Governing Board pertains to “All Cognitive and Noncognitive Assessment Items.” Specifically, the statute states that, “The Board shall have final authority on the appropriateness of all assessment items.” Under the law, the Board is therefore responsible for all NAEP test questions as well as all NAEP background questions administered as part of the assessment.
2. To meet this statutory requirement, the Board’s Policy on NAEP Item Development and Review shall be adhered to during all phases of NAEP item writing, reviewing, editing, and assessment construction. The National Center for Education Statistic (NCES), which oversees the operational aspects of NAEP, shall ensure that all internal and external groups involved in NAEP item development activities follow the Guiding Principles, Policies and Procedures as set forth in this Board policy.
3. Final review of all NAEP test questions for bias and appropriateness shall be performed by the Board, after all other review procedures have been completed, and prior to administration of the items to students.

Principle 4

The Governing Board shall review all NAEP test questions that are to be administered in conjunction with a pilot test, field test, operational assessment, or special study administered as part of NAEP.

Policies and Procedures

1. To fulfill its statutory responsibility for NAEP item review, the Board shall receive, in a timely manner and with appropriate documentation, all test questions that will be administered to students under the auspices of a NAEP assessment. These items include those slated for pilot testing, field testing, and operational administration.

2. The Board shall review all test items developed for special studies, where the purpose of the special study is to investigate alternate item formats or new technologies for possible future inclusion as part of main NAEP, or as part of a special study to augment main NAEP data collection.
3. The Board shall <u>not</u> review items being administered as part of test development activities, such as small-scale, informal try-outs with limited groups of students designed to refine items prior to large-scale pilot, field, or operational assessment.
4. NCES shall submit NAEP items to the Board for review in accordance with a mutually agreeable timeline. Items will be accompanied by appropriate documentation as required in this policy. Such information shall consist of procedures and personnel involved in item development and review, the match between the item pool and the framework content and process dimensions, and other related information.
5. For its first review, the Board will examine all items prior to the pilot test or field test stage. In the case of the NAEP reading assessment, all reading passages will be reviewed by the Board prior to item development. For each reading passage, NCES will provide the source, author, publication date, passage length, rationale for minor editing to the passage (if any), and notation of such editing applied to the original passage. NCES will provide information and explanatory material on passages deleted in its fairness review procedures.
6. For its second review, the Board will examine items following pilot or field testing. The items will be accompanied by statistics obtained during the pilot test or field test stage. These statistics shall be provided in a clear format, with definitions for each item analysis statistic collected. Such statistics shall include, but shall not be limited to: p-values for multiple-choice items, number and percentage of students selecting each option for a multiple-choice item, number and percentage not reaching or omitting the item (for multiple-choice and open-ended), number and percentage of students receiving various score points for open-ended questions, mean score point value for open-ended items, appropriate biserial statistics, and other relevant data.
7. At a <u>third stage</u>, for some assessments, the Board will receive a report from the calibration field test stage, which occurs prior to the operational administration. This "exceptions report" will contain information pertaining to any items that were dropped due to differential item functioning (DIF) analysis for bias, other items to be deleted from the operational assessment and the rationale for this decision, and the final match between the framework distribution and the item pool. If the technology becomes available to perform statistically sound item-level substitutions at this point in the cycle (from the initial field test pool), the Board shall be informed of this process as well.
8. All NAEP test items will be reviewed by the Board in a secure manner via in-person meetings, teleconference or videoconference settings, or on-line via a password-protected Internet site. The Board's Assessment Development Committee shall have primary responsibility for item review and approval. However, the Assessment Development Committee, in consultation with the Board Chair, may involve other NAGB members in the item review process on an *ad hoc* basis. The Board may also submit items to external experts, identified by

the Board for their subject area expertise, to assist in various duties related to item review. Such experts will follow strict procedures to maintain item security, including signing a Nondisclosure Agreement.

9. Items that are edited between assessments by NCES and/or its item review committees, for potential use in a subsequent assessment, shall be re-examined by the Board prior to a second round of pilot or field testing.
10. Documentation of the Board's final written decision on editing and deleting NAEP items shall be provided to NCES within 10 business days following completion of Board review at each stage in the process.

Principle 5

NAEP test questions will be accurate in their presentation, and free from error. Scoring criteria will be accurate, clear, and explicit.

Policies and Procedures

1. NCES, through its subject area content experts, trained item writers, and item review panels, will examine each item carefully to ensure its accuracy. All materials taken from published sources must be carefully documented by the item writer. Graphics that accompany test items must be clear, correctly labeled, and include the data source where appropriate. Items will be clear, grammatically correct, succinct, and unambiguous, using language appropriate to the grade level being assessed. Item writers will adhere to the specifications document regarding appropriate and inappropriate stimulus materials, terminology, answer choices or distractors, and other requirements for a given subject area. Items will not contain extraneous or irrelevant information that may differentially distract or disadvantage various subgroups of students from the main task of the item.
2. Scoring criteria will accompany each constructed-response item. Such criteria will be clear, accurate, and explicit. Carefully constructed scoring criteria will ensure valid and reliable use of those criteria to evaluate student responses to maximize the accuracy and efficiency of scoring.
3. Constructed-response scoring criteria will be developed initially by the item writers, refined during item review, and finalized during pilot or field test scoring. During pilot or field test scoring, the scoring guides will be expanded to include examples of actual student responses to illustrate each score point. Actual student responses will be used as well, to inform scorers of unacceptable answers.
4. Procedures used to train scorers and to conduct scoring of constructed-response items must be provided to the Board, along with information regarding the reliability and validity of such scoring. If the technology becomes available to score student responses electronically, the Board must be informed of the reliability and validity of such scoring protocol, as compared to human scoring.

Principle 6

All NAEP test questions will be free from racial, cultural, gender, or regional bias, and must be secular, neutral, and non-ideological. NAEP will not evaluate or assess personal or family beliefs, feelings, and attitudes, or publicly disclose personally identifiable information.

Policies and Procedures

1. An item is considered biased if it unfairly disadvantages a particular subgroup of students by requiring knowledge of obscure information unrelated to the construct being assessed. A test question or passage is biased if it contains material derisive or derogatory toward a particular group. For example, a geometry item requiring prior knowledge of the specific dimensions of a basketball court would result in lower scores for students unfamiliar with that sport, even if those students know the geometric concept being measured. Use of a regional term for a soft drink in an item context may provide an unfair advantage to students from that area of the country. Also, an item that refers to a low-achieving student as "slow" would be unacceptable.
2. In conducting bias reviews, steps should be taken to rid the item pool of questions that, because of their content or format, either appear biased on their face, or yield biased estimates of performance for certain subpopulations based on gender, race, ethnicity, or regional culture. A statistical finding of differential item functioning (DIF) will result in a review aimed at identifying possible explanations for the finding. However, such an item will not automatically be deleted if it is deemed valid for measuring what was intended, based on the NAEP assessment framework. Items in which clear bias is found will be eliminated. This policy acknowledges that there may be real and substantial differences in performance among subgroups of students. Learning about such differences, so that performance may be improved, is part of the value of the National Assessment.
3. Items shall be secular, neutral, and non-ideological. Neither NAEP nor its questions shall advocate a particular religious belief or political stance. Where appropriate, NAEP questions may deal with religious and political issues in a fair and objective way. The following definitions shall apply to the review of all NAEP test questions, reading passages, and supplementary materials used in the assessment of various subject areas:
    - Secular – NAEP questions will not contain language that advocates or opposes any particular religious views or beliefs, nor will items compare one religion unfavorably to another. However, items may contain references to religions, religious symbolism, or members of religious groups where appropriate. Examples: The following phrases would be acceptable: "shaped like a Christmas tree", "religious tolerance is one of the key aspects of a free society," "Dr. Martin Luther King, Jr. was a Baptist minister," or "Hinduism is the predominant religion in India."

- Neutral and Non-ideological - Items will not advocate for a particular political party or partisan issue, for any specific legislative or electoral result, or for a single perspective on a controversial issue. An item may ask students to explain both sides of a debate, or it may ask them to analyze an issue, or to explain the arguments of proponents or opponents, without requiring students to endorse personally the position they are describing. Item writers should have the flexibility to develop questions that measure important knowledge and skills without requiring both pro and con responses to every item. Examples: Students may be asked to compare and contrast positions on states rights, based on excerpts from speeches by X and Y; to analyze the themes of Franklin D. Roosevelt's first and second inaugural addresses; to identify the purpose of the Monroe Doctrine; or to select a position on the issue of suburban growth and cite evidence to support this position. Or, students may be asked to provide arguments either for or against Woodrow Wilson's decision to enter World War I. A NAEP question could ask students to summarize the dissenting opinion in a landmark Supreme Court case. The criteria of neutral and non-ideological also pertain to decisions about the pool of test questions in a subject area, taken as a whole. The Board shall review the entire item pool for a subject area to ensure that it is balanced in terms of the perspectives and issues presented.

4. The Board shall review both stimulus materials and test items to ensure adherence to the NAEP statute and the polices in this statement. Stimulus materials include reading passages, articles, documents, graphs, maps, photographs, quotations, and all other information provided to students in a NAEP test question.
5. NAEP questions will not ask a student to reveal personal or family beliefs, feelings, or attitudes, or publicly disclose personally identifiable information.

BIBLIOGRAPHY

Abedi, J. (2003). *Impact of student language background on content-based performance: Analyses of extant data* (CSE Technical Report 603). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing, Center for the Study of Evaluation, Graduate School of Education and Information Studies.

Abedi, J. (2004). *Inclusion of students with limited English proficiency in NAEP: Classification and measurement issues* (CSE Report 629). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing, Center for the Study of Evaluation, Graduate School of Education and Information Studies.

Almond R. G., Steinberg, L., & Mislevy, R. J. (2003). *A four-process architecture for assessment delivery, with connections to assessment design* (CSE Report 616). Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing, Center for the Study of Evaluation, Graduate School of Education and Information Studies.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy.* New York: Oxford University Press.

American Association for the Advancement of Science. (1994). *Science for all Americans* (rev. ed.). New York: Oxford University Press.

American Association for the Advancement of Science. (1997). Comparisons of benchmarks to national standards. In *Resources for science literacy: Professional development* [CD-ROM]. New York: Oxford University Press.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Ayala, C. A., Shavelson, R. J., Yin, Y., & Schultz, S. (2002). Reasoning dimensions underlying science achievement: The case of performance assessment. *Educational Assessment, 8*(2), 101-122.

Bar, V., Sneider, C., & Martimbeau, N. (1997, April). What research says: Is there gravity in space? *Science and Children*, 38-43.

Bar, V., & Zinn, B. (1998). Similar frameworks of action-at-a-distance in early scientists' and pupils' ideas. *Science and Education, 7*(5).

Bar, V., Zinn, B., Goldmuntz, R., & Sneider, C. (1994). Children's concepts about weight and free fall. *Science Education, 78*(2), 149-169.

Barnett, M., & Morran, J. (2002). Addressing children's alternative frameworks of the moon's phases and eclipses. *International Journal of Science Education, 24,* 859-879.

Bialystok, E. (Ed.). (1991). *Language processing in bilingual children*. Cambridge, UK: Cambridge University Press.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment, 11*(1), 33-63.

Catley, K., Lehrer, R., & Reiser, B. (2005). *Tracing a prospective learning progression for developing understanding of evolution.* Paper commissioned by the Committee on Test Design for K-12 Science Achievement, Board on Testing and Assessment, Center for Education, National Research Council, Washington, DC.

Champagne, A., Bergin, K., Bybee, R., Duschl, R., & Gallagher, J. (2004). *NAEP 2009 science framework development: Issues and recommendations*. Paper commissioned by the National Assessment Governing Board, Washington, DC.

Chandler, D. (1991). Weightlessness and microgravity. *The Physics Teacher, 29*(5), 312.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*(1), 66-71.

Council of Chief State School Officers. (2005). *A summary of national feedback provided on preliminary drafts gathered from surveys and regional and national feedback meetings.* Paper prepared for the National Assessment Governing Board. [Available from the National Assessment Governing Board, Washington, DC.]

Cummins, J. (1999). Alternative paradigms in bilingual education research: Does theory have a place? *Educational Researcher 28*, 26-32.

Dalton, B., Morocco, C. C., Tivnan, T., & Rawson, P. (1994). Effect of format on learning disabled and non-learning disabled students' performance on a hands-on science assessment. *International Journal of Educational Research*, *21*(3), 299-316.

De Boeck, P., & Wilson, M. (Eds.). (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York: Springer-Verlag.

Delgado, J. (2005, April). *Vocabulary assessment: Issues and ideas from science education.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). *Children's ideas in science.* Philadelphia: Open University Press.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: Research into children's ideas.* London: Routledge.

Duit, R. (2004). *Bibliography: Students' and teachers' conceptions and science education*. Retrieved February 15, 2006, from http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html

Ericsson, K. A. (2002). Attaining excellence through deliberate practice: Insights from the study of expert performance. In M. Ferrari (Ed.), *The pursuit of excellence in education* (pp. 21-55). Hillsdale, NJ: Erlbaum.

Finegold, M., & Gorsky, P. (1991). Students' concepts of force as applied to related physical systems: A search for consistency. *International Journal of Science Education, 13*(1), 97-113.

Fleming, S. (2002). *When and how should keyboarding be taught in elementary school?* Unpublished Masters Thesis, University of Northern Iowa, Cedar Falls, IA. Retrieved December 27, 2005, from http://ci.coe.uni.edu/facstaff/zeitz/web/general/keyboardingresearch.html

Genesee, F. (1994). Introduction. In F. Genesee (Ed.), *Educating second language children: The whole child, the whole curriculum, the whole community* (pp. 1-11). Cambridge, UK: Cambridge University Press.

Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gonzales, P., Guzmán, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., & Williams, T. (2004). *Highlights from the Trends in International Mathematics and Science Study (TIMSS) 2003* (NCES 2005-005). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Gunstone, R., & Watts, M. (1985). Force and Motion. In R. Driver, E. Guesne, & A. Tiberghien (Eds.), *Children's ideas in science.* Philadelphia: Open University Press.

Hamilton, L. S., Nussbaum, E. M., Kupermintz, H., Kerkhoven, J. I. M., & Snow, R. E. (1995). Enhancing the validity and usefulness of large-scale educational assessments: II. NELS:88 science achievement. *American Educational Research Journal, 32*(3), 555-581.

Herl, H. E., O'Neil, H. F., Chung, G. K. W. K., & Schacter, J. (1999). Reliability and validity of a computer-based knowledge mapping system to measure content understanding. *Computers in Human Behavior, 15*, 315-334.

Hewson, P. W., Beeth, M. E., & Thorley, N. R. (1998). Teaching for conceptual change. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education.* Great Britain: Kluwer, pp. 199-218.

Hourcade, J. P., Bederson, B. B., Druin, A., & Guimbretière, F. (2004). Differences in pointing task performance between preschool children and adults using mice. *Association for Computing Machines (ACM) Transactions on Computer-Human Interaction, 11*(4), 357-386.

International Technology Education Association. (2000). *Standards for technological literacy: Content for the study of technology.* Reston, VA: Author.

Kavanagh, C., Agan, L., & Sneider, C. (2005). Learning about phases of the moon and eclipses: A guide for teachers and curriculum developers. *Astronomy Education Review, 4*(1). Retrieved from http://aer.noao.edu/

Kendall, J. S., & Marzano, R. J. (2004). *Content knowledge: A compendium of standards and benchmarks for K-12 education* (4th ed.). Retrieved December 20, 2004, from http://www.mcrel.org/standards-benchmarks/

Kennedy, C. A., Brown, N. J. S., Draney, K., & Wilson, M. (2005, April). *Using progress variables and embedded assessment to improve teaching and learning.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Ketterlin-Geller, L. R. (2005). Knowing what all students know: Procedures for developing universal design for assessment. *The Journal of Technology, Learning, and Assessment, 4*(2). Retrieved from http://www.jtla.org

Kopriva, R. (2000). *Ensuring accuracy in testing for English language learners.* Washington, DC: Council of Chief State School Officers.

Kopriva, R., & Saez, S. (1997). *Guide to scoring LEP student responses to open-ended mathematics items.* Washington, DC: Council of Chief State School Officers.

Koretz, D., & Barton, K. (2003-2004). Assessing students with disabilities: Issues and evidence. *Educational Assessment, 9*(1 & 2), 29-60.

Lee, O. (1999). Equity implications based on the conceptions of science achievement in major reform documents. *Review of Educational Research, 69*, 83-115.

Leighton, J. P. (2004). Avoiding misconception, misuse, and missed opportunities: The collection of verbal reports in educational achievement testing. *Educational Measurement: Issues and Practice, 23*(4), 6-15.

Li, M., Ruiz-Primo, M. A., & Shavelson, R. J. (In press). Towards a science achievement framework: The case of TIMSS 1999. In S. J. Howie & T. Plomp (Eds.), *Contexts of learning mathematics and science: Lessons learned from TIMSS.*

Li, M., & Shavelson, R. J. (2001, April). *Validating the links between knowledge and test items from a protocol analysis.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Mackey, W. F. (1962). The description of bilingualism. *Canadian Journal of Linguistics, 7*, 51-85.

Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist, 34*, 207-218.

Masters, G. N., Adams, R. A., & Wilson, M. (1990). Charting of student progress. In T. Husen & T. N. Postlethwaite (Eds.), *International encyclopedia of education: Research and studies* (supplementary Vol. 2, pp. 628-634). Oxford: Pergamon Press.

Mazzeo, J., Carlson, J. E., Voelkl, K. E., & Lutkus, A. D. (1999). *Increasing the participation of special needs students in NAEP. A report on 1996 research activities* (NCES Publication No. 2000-473). Washington, DC: National Center for Education Statistics.

Meade, S. D., & Dugger, W. E., Jr. (2004, October). Reporting on the status of technology education in the U.S. *The Technology Teacher*, 29-35.

Metz, K. E. (1995). Reassessment of developmental constraints on children's science instruction. *Review of Educational Research, 65*(2), 93-127.

Minstrell, J. (1998). *Student thinking and related instruction: Creating a facet-based learning environment.* Paper presented at the meeting of the Committee on Foundations of Assessment, Woods Hole, MA.

Morrison, R. C. (1999). Weight and gravity—the need for consistent definitions. *The Physics Teacher, 37*(1), 51.

Mullis, I. V. S., Martin, M. O., Smith, T. A., Garden, R. A., Gregory, K. D., Gonzalez, E. J., Chrostowski, S. J., & O'Connor, K. M. (2001). *TIMSS assessment frameworks and specifications 2003.* Boston: Boston College, International Study Center, Lynch School of Education.

National Center for Education Statistics. (2005a). NAEP inclusion policy. *The Nation's Report Card.* Retrieved September 7, 2005, from http://nces.ed.gov/nationsreportcard/about/inclusion.asp

National Center for Education Statistics. (2005b). What are the differences between long-term trend NAEP and main NAEP? *The Nation's Report Card.* Retrieved November 4, 2005, from http://nces.ed.gov/nationsreportcard/about/ltt_main_diff.asp

National Research Council. (1996). *National science education standards.* National Committee on Science Education Standards and Assessment. Coordinating Council for Education. Washington, DC: National Academy Press.

National Research Council. (1997a). *Educating one and all: Students with disabilities and standards-based reform.* Committee on Goals 2000 and the Inclusion of Students with Disabilities. L. McDonnell, M. McLaughlin, & P. Morison (Eds). Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (1997b). *Improving schooling for language-minority children: A research agenda*. Committee on Developing a Research Agenda on the Education of Limited-English-Proficient and Bilingual Students. D. August & K. Hakuta (Eds.). Board on Children, Youth, and Families, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (1999a). *Evaluation of the voluntary national tests: Phase 1.* L. L. Wise, R. M. Hauser, K. J. Mitchell, & M. J. Feuer (Eds.). Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (1999b). *High stakes: Testing for tracking, promotion, and graduation.* Committee on Appropriate Test Use. J. P. Heubert & R. M. Hauser (Eds.). Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (1999c). *How people learn: Brain, mind, experience, and school.* Committee on Developments in the Science of Learning. J. D. Bransford, A. L. Brown, & R. R. Cocking (Eds.). Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment.* Committee on the Foundations of Assessment. J. Pellegrino, N. Chudowsky, & R. Glaser (Eds.). Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (2005a). *America's lab report: Investigations in high school science.* Committee on High School Science Laboratories: Role and Vision. S. R. Singer, M. L. Hilton, & H. A. Schweingruber (Eds.). Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (2005b). *How students learn: History, mathematics, and science in the classroom.* Committee on *How People Learn*, a Targeted Report for Teachers. M. S. Donovan & J. D. Bransford (Eds.). Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council. (2005c). *Systems for state science assessment.* Committee on Test Design for K-12 Science Achievement. M. Wilson & M. Bertenthal (Eds.). Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Neidorf, T. S., Binkley, M., & Stephens, M. (2006). *Comparing science content in the National Assessment of Educational Progress (NAEP) 2000 and Trends in International Mathematics and Science Study (TIMSS) 2003 Assessments* (NCES 2006-026). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved April 10, 2006, from http://nces.ed.gov/pubsearch

Nielsen, J. (1993). *Usability engineering*. San Diego, CA: Academic Press.

Nielsen, J. (2000). *Designing web usability*. Indianapolis, IN: New Riders Publishing.

Nielsen, J., & Mack, R. L. (1994). *Usability inspection methods*. New York: John Wiley.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn.* New York and Cambridge, UK: Cambridge University Press.

Oakland T., & Lane, H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing, 4*(3), 239-252.

Olson, J. F., & Goldstein, A. A. (1997). *The inclusion of students with disabilities and limited English proficient students in large-scale assessments: A summary of recent progress* (NCES Publication No. 97-482). Washington, DC: National Center for Education Statistics.

Organisation for Economic Cooperation and Development. (2005). *PISA 2006 scientific literacy framework.* Paris: OECD/PISA.

Osborne, R. J., & Freyberg, P. (1985). *Learning in science: The implications of children's science.* Auckland, New Zealand: Heinemann.

Partnership for 21st Century Skills. (2004). *ICT literacy map for science.* Retrieved from http://www.21stcenturyskills.org/images/stories/matrices/ictmap_science.pdf

Perie, M., Grigg, W., & Dion, G. (2005). *The nation's report card: Mathematics 2005* (NCES 2006-453). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office.

Perie, M., Moran, R., & Lutkus, A. D. (2005). *NAEP 2004 trends in academic progress: Three decades of student performance in reading and mathematics* (NCES 2005–464). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office.

Perkins, K., & Wieman, C. (2005). Web watch: Free online resource connects real-life phenomena with science. *Physics Education, 40*(1), 93-95.

Persky, H., Bennett, R. E., Weiss, A. R., & Jenkins, F. (2005). *Problem solving in technology-rich environments: A report from the NAEP technology-based assessment project* (NCES 2005-XXX). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office.

Pine, J., Baxter, G., & Shavelson, R. (1993). Assessments for hands-on elementary science curricula. *MSTA Journal, 39*(2), 3, 5-19.

Quellmalz, E. S., Griffin, M., Hurst, K., Kreikemeier, P., Rosenquist, A., & Zalles, D. (2004, April). *Integrated performance assessments with technology (IPAT): Design model and prototypes*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA. Funded by the National Science Foundation, REC 0126350. Retrieved from http://ipat.sri.com

Quellmalz, E. S., Haertel, G. D., DeBarger, A., & Kreikemeier, P. (2005). *A study of evidence of the validities of assessments of science inquiry in the National Assessment of Educational Progress (NAEP), Trends in Mathematics and Science Survey (TIMSS), and the New Standards Science Reference Exam (NSSRE) in Science* (Validities Technical Report #1). Menlo Park, CA: SRI International. Funded by the National Science Foundation, REC 9903287.

Resnick, L. B. & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments* (pp. 37-76). Boston: Kluwer.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford, UK: Blackwell.

Rosenquist, A., Shavelson, R. J., & Ruiz-Primo, M. A. (2000). *On the "exchangeability" of hands-on and computer simulation science performance assessments.* Los Angeles: University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing, Center for the Study of Evaluation, Graduate School of Education and Information Studies.

Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching, 38*(2), 260-278.

Ruiz-Primo, M. A., & Shavelson, R. J. (1996a). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33*(6), 569-600.

Ruiz-Primo, M. A., & Shavelson, R. J. (1996b). Rhetoric and reality in science performance assessments: An update. *Journal of Research in Science Teaching, 33*(10), 1045-1063.

Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching, 39*(5), 369-393.

Ruiz-Primo, M. A., Shavelson, R. J., Li, M., & Schultz, S. E. (2001). On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. *Educational Assessment, 7*(2), 99-141.

Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching, 35*(3), 265-296.

Scalise, K. (2004). *BEAR CAT: Toward a theoretical basis for dynamically driven content in computer-mediated environments*. Unpublished doctoral dissertation, University of California, Berkeley.

Schatz, D., & Cooper, D. (1994). *Astro adventures*. Seattle, WA: Pacific Science Center.

Schmidt, W. H., Raizen, S. A., Britton, E. D., Bianchi, L. J., & Wolfe, R. G. (1997). *Many visions, many aims, volume 2: A cross-national investigation of curricular intentions in school science*. Dordrecht, The Netherlands: Kluwer Academic.

Sequeira, M., & Leite, L. (1991). Alternative conceptions and history of science in physics teacher education. *Science Education, 75*(1), 45-56.

Shavelson, R. J. (2006). On the integration of formative assessment in teaching and learning: Implications for new pathways in teacher education. In F. Oser, F. Achtenhagen, & U. Renold (Eds.), *Competence-oriented teacher training: Old research demands and new pathways.* Utrecht, The Netherlands: Sense Publishers.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1991). Performance assessment in science. R. Stiggins & B. Plake (Guest Eds.), *Applied Measurement in Education* [Special issue], *4*(4), 347-362.

Shavelson, R. J., Gao, X., & Baxter, G. P. (1995). On the content validity of performance assessments: Centrality of domain specification. In M. Birenbaum & F. Dochy (Eds.), *Alternatives in assessment of achievements, learning focuses and prior knowledge.* Boston: Kluwer.

Shavelson, R. J., & Ruiz-Primo, M. A. (1999). On the psychometrics of assessing science understanding. In J. J. Mintzes, J. H. Wamhersee, & J. D. Novak (Eds.), *Assessing science understanding: A human constructivist view.* New York: Academic Press.

Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. W. (2005). Windows into the mind. *Higher Education, 49,* 413-430.

Smith, C., Wiser, M., Anderson, C. W., Krajcik, J., & Coppola, B. (2004). *Implications of research on children's learning for assessment: Matter and atomic molecular theory.* Paper commissioned by the Committee on Test Design for K-12 Science Achievement, Board on Testing and Assessment, Center for Education, National Research Council, Washington, DC.

Solano-Flores, G., Lara., J., Sexton, U., & Navarrete, C. (2001). *Testing English language learners: A sampler of student responses to science and mathematics test items*. Washington, DC: Council of Chief State School Officers.

Solano-Flores, G., & Li, M. (2006). The use of generalizability (G) theory in the testing of linguistic minorities. *Educational Measurement: Issues and Practice 25*(1), 13-22.

Solano-Flores, G., & Shavelson, R. J. (1997). Development of performance assessments in science: Conceptual, practical, and logistical issues. *Educational Measurement: Issues and Practice, 16*(3), 16-24.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32*(2), 3-13.

Taylor, K. L., & Dionne, J. P. (2000). Accessing problem-solving strategy knowledge: The complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educational Psychology, 92*(3), 413-425.

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved September 13, 2005, from http://education.umn.edu/NCEO/OnlinePubs/Synthesis44.html

Thompson, S., Thurlow, M., & Moore, M. (2003). *Using computer-based tests with students with disabilities* (Policy Directions No. 15). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved December 27, 2005, from http://education.umn.edu/NCEO/OnlinePubs/Policy15.html

Valdés, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex.

Waner, K., Behymer, J., & McCrary, S. (1992). Two points of view on elementary school keyboarding. *Business Education Forum*, 27-35.

Wellington, J., & Osborne, J. (2001). Language and literacy in science education. Philadelphia: Open University Press.

White, R. T., & Gunstone, R. (1992). *Probing understanding.* New York: Falmer Press.

Wilson, E. B. (1928). *The cell in development and heredity*. New York: Macmillan.

Yin, Y. (2005). *The influence of formative assessments on student motivation, achievement, and conceptual change.* Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching, 42*(2), 166-184.