This policy research document is intended for policymakers to use when examining possible changes to the state assessment's alignment with the NAEP.

The 2009 NAEP test is not yet in existence, so the purpose of this report is to give policymakers a headstart in determining where they might, if they so decide, begin to make changes in their assessment standards and specifications to develop an assessment system more closely aligned to that used for the NAEP.

BACKGROUND TO THE STUDY

This report presents the findings of an alignment study comparing the new science framework for the 2009 National Assessment of Educational Progress (NAEP) and the accompanying science assessment and item specifications with the Texas state science assessment. More details about the documents compared are in appendix A. The study was conducted for the Regional Education Laboratory Southwest, funded by the Institute of Education Sciences to provide research and support to Arkansas, Louisiana, New Mexico, Oklahoma, and Texas. The study was undertaken in anticipation of a growing need in the region to be better informed about how state assessment standards in science compare with those tested in the NAEP.

Five factors make this study timely. First, the importance of state science assessments has been increased by the No Child Left Behind Act. Beginning in the 2007/08 school year, states are required to administer science assessments to all students in each of the elementary, middle, and high school levels, holding states and local school districts accountable for student academic achievement in science (NCLB, 2001).

Second, the NAEP is increasingly being used as a benchmark against which student achievement across the nation can be compared (Linn, 2005; Linn, Baker & Herman, 2005). The NAEP has been dubbed the "nation's report card," and when fresh NAEP results are released—as they were for science in 2006, following an administration of the test in 2005—the media report the results (Cavanagh, 2006a, 2006b). Although states are not sanctioned for failing to demonstrate NAEP student performance improvement, NAEP data do provide an external accountability benchmark and serve to verify student achievement on state assessments. In fact, the National Center for Education Statistics has a website (http://nces.ed.gov/nationsreportcard/nde/statecomp/) that allows anyone to create customized comparative reports based on the latest NAEP data. So anyone can create tables that compare states and jurisdictions based on the average scale scores for selected groups of public school students within a single assessment year, or compare the change in performance between two assessment years.

Third, NAEP data are being used more in educational research to investigate how the No Child Left Behind Act provisions have played out in different states. For example, Olson (2005) compared the percentages of students at or above

the proficient level on the 2005 state grade 8 mathematics assessments in 33 states. The study showed that, on average, 33 percent more students scored at or above the proficient level according to the state assessment than did so according to the NAEP. As yet, no similar study has been done of science, but with the release of the 2005 NAEP results it is now possible to do so.

Fourth, political attention is beginning to focus on using the NAEP as a yardstick for measuring state standards (Olson, 2007). In January 2007 two bills were introduced in Congress, one seeking to encourage states to benchmark their own standards and tests to NAEP and the other calling for states to adopt voluntary "American education content standards" in mathematics and science that would be developed by the National Assessment Governing Board, the body responsible for the NAEP. These issues will doubtless be a topic of debate in the upcoming reauthorization of the No Child Left Behind Act.

Fifth, the standards and test specifications that form the blueprint for the content the NAEP science assessment covers and the types of items it uses were revised in 2006. The 2009 NAEP framework takes account of the latest knowledge on science learning and assessment, which suggests that measuring student understanding involves much more than assessing factual knowledge. It defines the science knowledge and skills that science-literate students should possess at grades 4, 8, and 12. The assessment itself, while retaining some familiar paper-and-pencil assessment formats, will also include student performance assessments in both classroom settings and computer simulations. The 2009 NAEP framework will determine the shape of NAEP science assessments through 2017, setting the direction of science assessment across the nation.

States want to know how well their standards align with the NAEP so they can make informed decisions about possible changes to their own standards and assessment systems

These factors are working together to gradually raise the status of the NAEP to a de facto national benchmark, and states naturally want to know how well their state standards align with the NAEP so they can make informed decisions about possible changes to their own standards and assessment systems. This report describes the results of a systematic alignment study conducted for that purpose. Details of the study are in appendix B.

The intent of this report is to inform those in the Texas Education Agency responsible for shaping the state assessment in science how the current assessment standards and test specifications compare with those of the national NAEP 2009 assessment. It is hoped that this study will be of use to policymakers and others in the state who are interested in the Texas state assessments in science.

Similar reports have been completed for Arkansas, Louisiana, New Mexico, and Oklahoma, but there is no intent to compare Texas with these other states. This report shows where there is good content alignment with NAEP standards, identifies where there is partial alignment, pinpoints NAEP standards where there are no corresponding state standards, and highlights where the Texas standards go beyond the NAEP. It also deals with the assessment specifications, showing what percentages of the NAEP assessment at each grade level are devoted to different science topics and comparing that to the coverage of the topics in the Texas assessment. And it compares the proportions of types of items used to test students' science knowledge and skills. Through comprehensive comparative analysis, the report provides a way for the Texas Education Agency to gauge how well its tests are doing in covering the depth of science understanding expected on the NAEP test.

The results are presented in the summary tables and narratives in the sections that follow. Those sections provide an analysis that highlights the differences found between the NAEP assessment and the Texas state assessment. For more detail about the alignment of the Texas Assessment of Knowledge and Skills (TAKS) to the individual content statements of the NAEP, turn to the tables in appendixes C–E. They show exactly which Texas

standards align with a particular NAEP statement and, in cases of partial alignment, explain why. For a discussion of methodology, see box 1 and appendix B.

CONTENT ALIGNMENT AT GRADE 4

The NAEP grade 4 science standards were compared with the Texas Essential Knowledge and Skills statements in the TAKS Information Booklet for Science Grade 5 (Texas Education Agency, 2004a).

For grade 4 the NAEP provides 33 distinct content statements (displayed in parentheses in table 1). Four of these content statements (12 percent) are fully addressed by Texas standards in TAKS, 25 (76 percent) partially addressed, and 4 (12 percent) unaddressed (figure 1).

The average alignment rating for grade 4 is 2 (table 1). The majority of content statements were given ratings of 2, which means that state standards partially address the NAEP content statements (figure 1 and appendix C).

Areas of full alignment

Four NAEP grade 4 content statements are fully addressed by Texas grade 5 assessment standards. One of 15 physical science NAEP statements has full alignment with Texas, as do 2 of 7 life science statements and 1 of 11 Earth and space science statements.

BOX 1

Methodology

The chief research questions driving this study were these: "To what extent do current state assessment standards cover the content on which NAEP 2009 assessments will be based?" and "To what extent do current state assessment specifications align with the NAEP 2009 assessment specifications?"

The methodology used to answer the questions followed the successful pattern of a similar study conducted by WestEd in New England, which examined the alignment of math and reading standards with the NAEP. The methodology developed by WestEd for the New England study was designed to include all the most prominent alignment methodologies, discussed in appendix B. Thus far, alignment studies and methods have focused on aligning standards and tests, whereas the objective of this study was to compare one set of assessment standards and specifications with another. In this study, however, the methodology is based upon methodologies for aligning standards to tests, because similar principles are used in both types of alignments.

In this study reviewers followed the methodology of the portion of the previous study examining alignment between two sets of standards. Following the methodology of Achieve, test blueprints were examined to find correspondence between the two documents (see appendix B). Reviewers performed gap analyses to identify content included in one set of standards but not the other, identified issues of order so they could reveal differences in the grade levels at which standards appear, and examined the degree to which the standards and assessments cover content to the same depth and have similar cognitive demands (depth-of-knowledge consistency) and the degree to which assessments cover the same range of content as the corresponding standards (range-of-knowledge correspondence) to determine whether there was a match between Texas and NAEP in the level of detail, the cognitive demands, and the range of content covered. A coding scheme was used to indicate alignment issues and reviewer ratings, and a matrix-like format was created to facilitate alignment.

Reviewers attended several training sessions, conducted individual reviews, and then met in teams of two to reach consensus on ratings. This consensus method was designed to create one consensus rating per NAEP standard with the help of a moderator and was not intended to allow for disagreements. This methodology was determined to be best suited to the scope and timing of this study. The consensus methodology is designed to highlight areas for states to examine, not to gather large amounts of data, record multiple ratings, or measure inter-rater reliability (see appendix B for more on methodology).

TABLE 1

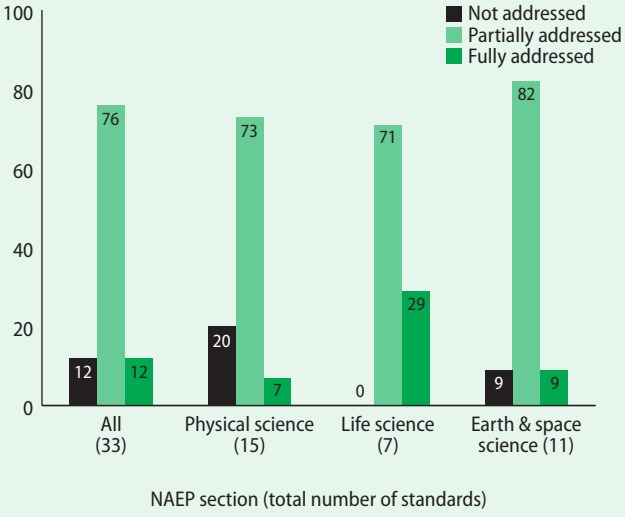
Average ratings of alignment of Texas grade 5 standards and National Assessment of Educational Progress grade 4 science content statements

NAEP content area (number of NAEP standards)	Average rating
Overall physical science (15)	1.9
Matter (6)	2.0
Energy (5)	2.2
Motion (4)	1.3
Overall life science (7)	2.3
Structures and functions of living systems (4)	2.3
Changes in living systems (3)	2.3
Overall Earth and space science (11)	2.0
Earth and space in time (3)	2.0
Earth structures (3)	2.0
Earth systems (5)	2.0
All content (33)	2.0

Note: Rating is based on a scale of 1 to 3, where 1 indicates that state standards do not address NAEP content statement, 2 that state standards partially address NAEP content statement, and 3 that state standards fully address or exceed NAEP content statement by targeted grade level.

FIGURE 1

The majority of Texas grade 5 standards partially address National Assessment of Educational Progress content statements



The four NAEP grade 4 content statements fully addressed by the Texas Essential Knowledge and Skills for Science curriculum are P4.7—forms of energy (heat, electricity, light, and sound), L4.1—needs of organisms, L4.6—plants and animals closely resemble their parents, and E4.7—the sun warms the land, air and water and helps plants grow.

Areas of partial alignment

Twenty-five NAEP grade 4 content statements (76 percent) have partial alignment, in large part because many Texas benchmarks imply content explicitly stated by the NAEP and because the NAEP content statements are often more detailed than the Texas standards.

Raters found that many Texas standards imply content that the NAEP addresses in depth. For example, Texas 5.7 (D) mentions observing and measuring characteristic properties such as boiling points and melting points, which most likely implies the content regarding measuring weight (mass) and volume in the corresponding NAEP standard (P4.1). In life science, NAEP's L4.4 states that some plants and animals survive and reproduce, die, or move to new locations when the environment changes, but the corresponding Texas statements include only comparing adaptive characteristics (5.9 A) and thriving, becoming ill, or perishing (3.8 C) but exclude moving. In Earth Science, E4.3 states that some changes on the surface of Earth are due to slow processes and others are due to rapid processes. Texas has four corresponding standards (3.6B, 5.11A, 5.5A, and 5.12A), but none delineate the slow versus the fast processes of Earth changes.

NAEP items are also addressed in higher Texas grade levels. For example, NAEP P4.10 addresses vibrating objects and producing varying pitches of sound, and while Texas 5.8 (D) states that vibrating an object can produce sound, Texas does not include how to vary the pitch of sound until grade 8. The NAEP addresses, in L4.2, the basic needs of animals and plants, including "a source of energy and building material for growth and repair." The corresponding Texas statements (3.8 and 3.8A) do not address the need for a source of energy or cellular use, but a higher grade statement in Texas

(6.10 B) covers these topics. In Earth and space science E4.2 deals with the observable shape of the moon. And although Texas 5.6 (A) includes events and changes of the lunar cycle, a higher grade statement, 7.13 B, more closely addresses the NAEP statement by including "the observed cyclical phases of the moon."

Areas of nonalignment

Three NAEP statements in physical science and one in Earth and space science are unaddressed by Texas content statements. P4.12 and P4.13 cover motion at the macroscopic level, including descriptions of position and motion, relative observation and the definition of speed. P4.15 covers the concept of gravity. E4.9 covers the use of tools for observing, recording, and predicting weather changes over days and over seasons.

Areas where Texas standards go beyond the NAEP content statements

Texas has 54 benchmarks in the TAKS Information Booklet for Science Grade 5. The NAEP does not address, in its content statements, the 10 Texas statements in nature of science, 3 of the 15 in life science, or 6 of the 17 in Earth and space sciences.

The NAEP does not address the nature of science statements because it discusses inquiry in a section separate from the content statements, called "science practices," intended to crosscut all NAEP content.

In life science the NAEP does not cover habitat or niche (5.9 B), predicting changes from adaptation (5.9 C), or learned characteristics resulting from environmental influence (5.10 B).

In Earth and space science the NAEP does not address tree rings and sedimentary rock (5.11 B), past events that led to the formation of Earth's resources (5.11 C), identifying the planets and their positions in the solar system (3.11 C), describing the sun's characteristics (3.11 D), the effects of oceans on land (4.11 B), or comparing the physical characteristics of Earth with those of the moon (5.12 C).

The NAEP addresses all Texas physical science statements. But for parts of some physical science NAEP statements (P4.2, P4.4, and P4.7), Texas contains more detailed content than the NAEP (denoted by the code "MD-TX").

Summary of grade 4 alignment

Most grade 4 NAEP content statements are, to some degree, addressed by the grade 5 Texas content statements, but the Texas statements typically are only partially aligned to the NAEP statements. The Texas Assessment of Knowledge and Skills also contains several content items not listed in NAEP content statements. Most of the NAEP content is implied in the Texas content, not explicit. In a few cases Texas addresses a topic at a higher grade and in more detail than the NAEP did. The overall alignment rating is 2, which indicates partial alignment.

Most grade 4 NAEP content statements are, to some degree, addressed by the grade 5 Texas content statements, but the Texas statements typically are only partially aligned to the NAEP statements

CONTENT ALIGNMENT AT GRADE 8

The NAEP grade 8 science standards were compared with the Texas Essential Knowledge and Skills statements found in the TAKS Information Booklet for Science Grade 8 (Texas Education Agency, 2005).

For grade 8, the NAEP provides 43 distinct content statements (displayed in parentheses in table 2). Five (12 percent) are fully addressed by Texas standards in TAKS, 17 (40 percent) partially addressed, and 21 (49 percent) unaddressed.

The average alignment rating for grade 8 is 1.6. The majority of content statements were given a rating of 1, which means that most NAEP content statements are unaddressed by grade 8 TAKS (figure 2 and appendix D).

TABLE 2

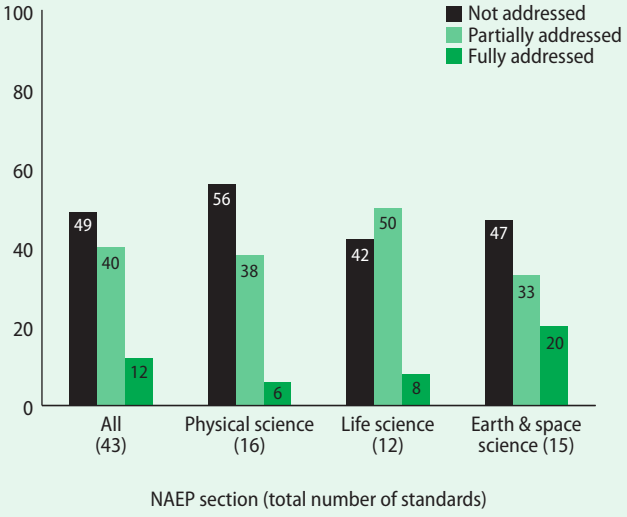
Average ratings of alignment of Texas grade 8 standards and National Assessment of Educational Progres grade 8 science content statements

NAEP content area (number of NAEP standards)	Average rating
Overall physical science (16)	1.5
Matter (7)	1.7
Energy (6)	1.2
Motion (3)	1.7
Overall life science (12)	1.7
Structures and functions of living systems (8)	1.8
Changes in living systems (4)	1.5
Overall Earth and space science (15)	1.7
Earth and space in time (4)	1.3
Earth structures (6)	1.3
Earth systems (5)	2.6
All content (43)	1.6

Note: Rating is based on a scale of 1 to 3, where 1 indicates that state standards do not address NAEP content statement, 2 that state standards partially address NAEP content statement, and 3 that state standards fully address or exceed NAEP content statement by targeted grade level.

FIGURE 2

Many Texas grade 8 standards do not address National Assessment of Educational Progress content statements



Areas of full alignment

Five NAEP grade 8 content statements are fully addressed by Texas grade 8 assessment standards. One of 16 physical science NAEP statements has full alignment with Texas, as do 1 of 12 life science statements and 3 of 15 Earth and space science statements.

The five NAEP grade 8 content statements fully addressed by the TEKS are P8.3—all substances are composed of elements in the periodic table, which organizes elements with similar properties, L8.5—consumers and decomposers have different ways of meeting energy needs, E8.11 and E8.12—the sun's observable effects, and E8.15—human-induced changes in Earth materials and systems.

Areas of partial alignment

Seventeen NAEP grade 8 content statements (40 percent) have partial alignment.

Raters found that many Texas content statements do not have as much content or detail as NAEP's statements. For example, Texas 6.7 (B) says, "classify substances by their physical and chemical properties," while NAEP gives a similar statement in P8.5 but continues to give examples and descriptions of metals and acids to further illustrate the classification of substances. In life science NAEP's L8.3 describes the functioning of cells, including growth, division, and the use of food. The corresponding Texas statement (6.10 (B)) contains the more general statement, "determine that all organisms are composed of cells that carry on functions to sustain life." In Earth and space science E8.9 in the NAEP contains content on the constant rate of movement of lithospheric plates and the resulting geological events. Texas 8.14 (A) does not contain a similar amount of detail, as it more generally asks for the prediction of land features resulting from gradual geologic changes.

Several instances of implied content were also found for Texas content statements. Texas statement 7.12 (C) states, "describe how different environments support different varieties of organisms," while NAEP L8.7 goes into more detail on the biotic and abiotic factors that support organisms.

Texas 6.8 (B) contains a statement about explaining and illustrating the water cycle, while NAEP E8.14 describes the processes in the water cycle.

Areas of nonalignment

Nine NAEP statements in physical science, five in life science, and seven in Earth and space science are unaddressed by Texas content statements.

In NAEP physical science the unaddressed content statements are P8.1—the particulate model of matter that explains properties of matter, P8.2—the arrangement of atoms and molecules that explain chemical properties, P8.6—changes of state, P8.8—kinetic energy, P8.9—potential energy, P8.11—light energy from the sun, P8.12– energy transfer and conservation of energy, P8.13—nuclear reactions in the sun and plants' usage of sunlight, and P8.15—forces acting at a distance.

In life science the following NAEP statements are unaddressed by Texas: L8.2—cell division and differentiation, L8.6—interactions and relationships of organisms, L8.8—causation of changes to organisms' environments, L8.9—reproduction, and L8.12—anatomical features of organisms and classification.

In Earth and space science the following are unaddressed: E8.1—a model of the solar system, E8.3—fossils, E8.4—Earth processes and the measurement of geologic time, E8.6—the composition of soils, E8.7—the composition of the atmosphere, and E8.10—Earth's magnetic field.

Areas where Texas standards go beyond the NAEP content statements

Texas has 53 "knowledge and skills statements" listed in the TAKS Information Booklet for Science Grade 8. The NAEP does not address, in its content statements, the 11 Texas statements in nature of science, 6 of the 13 in living systems and the environment, 3 of the 8 in structures and properties of matter, 5 of the 7 in motion, forces, and energy, or 6 of the 14 in Earth and space systems.

The NAEP does not address nature of science statements because it discusses inquiry in a section separate from the content statements, called "science practices," intended to crosscut all NAEP content.

In living systems and the environment the NAEP does not cover 6.5 (B)—describing differences between properties of a system and properties of its parts, 7.12 (D)—observing and describing ecological successions, 8.6 (A)—interactions among human systems, 8.6 (B)—feedback mechanisms for maintaining equilibrium, 8.6 (C)—interactions within ecosystems, or 8.11 (C)—predictions about outcomes of genetic combinations.

In structures and properties of matter the NAEP does not address 8.8 (A)—describing the structure and parts of an atom, 8.8 (B)—identifying the properties of an atom, or 8.10 (A)—illustrating interactions between matter and energy.

For motion, forces, and energy the NAEP does not address 6.9 (A)—identifying energy transformations, 7.8 (A)—illustrating examples of potential and kinetic energy, 7.6 (A)—demonstrating basic relationship between force and motion, 7.6 (C)—relating forces to basic processes in organisms, and 8.7 (B)—recognizing that waves are generated and can travel through different media.

In Earth and space systems the NAEP does not address 6.14 (B)—identifying relationship between groundwater and surface water in a watershed, 8.12 (A)—analyzing and predicting the sequence of events in lunar and rock cycles, 8.13 (A)—describing characteristics of the universe such as stars and galaxies, 7.14 (A)—describing and predicting the impact of different catastrophic events on Earth, 7.14 (B)—analyzing effects of regional erosional deposition and weathering, or 7.14 (C)—making inferences and drawing conclusions about effects of human activity on Earth's resources.

Almost half the NAEP grade 8 content statements are unaddressed by the grade 8 Texas statements

Summary of grade 8 alignment

Almost half the NAEP grade 8 content statements are unaddressed by the grade 8 Texas statements. Most other Texas content statements are only partially aligned to the NAEP. But Texas also contains many content items not listed in NAEP's content statements. Most of the partially aligned NAEP statements contain more content or more detail than the corresponding Texas statements, and Texas often contains statements that imply content explicitly stated by the NAEP. The overall alignment rating is 1.6, between nonalignment and partial alignment.

CONTENT ALIGNMENT AT GRADE 12

The NAEP grade 12 science standards were compared with the Texas Essential Knowledge and Skills statements in the TAKS information booklets for grades 10 and 11/exit level (Texas Education Agency, 2004b).

For grade 12 the NAEP provides 49 distinct content statements (displayed in parentheses in table 3). Four (8 percent) were fully addressed by Texas standards in TAKS, 22 (45 percent) were partially addressed, and 23 (47 percent) were unaddressed.

The average alignment rating for grade 12 is 1.6. The rating used most was 1, which means that many NAEP statements are unaddressed (figure 3 and appendix E).

Areas of full alignment

Four NAEP grade 12 content statements are fully addressed by Texas' grades 10 and 11 assessment standards. Two of 23 physical science NAEP statements have full alignment with Texas, as do 2 of 13 life science statements.

The four NAEP grade 12 content statements fully addressed by the TEKS are P12.16—total energy is conserved in a closed system, P12.20—acceleration and its relationship to force and mass, L12.9—DNA and genes, and

TABLE 3

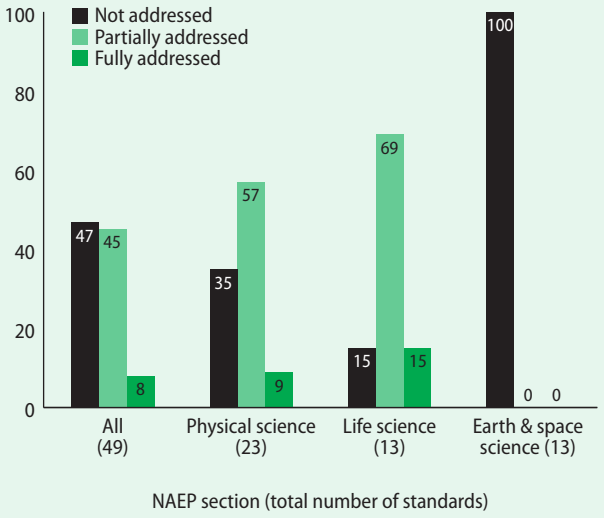
Average ratings of alignment of Texas grade 10 and 11/exit level standards and National Assessment of Educational Progress grade 12 science content statements

NAEP content area (number of NAEP standards)	Average rating
Overall physical science (23)	1.7
Matter (7)	1.7
Energy (9)	1.7
Motion (7)	1.9
Overall life science (13)	2.0
Structures and functions of living systems (7)	1.9
Changes in living systems (6)	2.2
Overall Earth and space science (13)	1.0
Earth and space in time (7)	1.0
Earth structures (1)	1.0
Earth systems (5)	1.0
All content (49)	1.6

Note: Rating is based on a scale of 1 to 3, where 1 indicates that state standards do not address NAEP content statement, 2 that state standards partially address NAEP content statement, and 3 that state standards fully address or exceed NAEP content statement by targeted grade level.

FIGURE 3

Many Texas grades 10 and 11/exit level standards do not fully address National Assessment of Educational Progress content statements



L12.12—molecular and anatomical evidence for evolution.

Areas of partial alignment

Twenty-two NAEP grade 12 content statements (45 percent) have partial alignment.

Raters found that many Texas content statements do not have as much content or detail as NAEP's statements. For example, although Texas has many standards that match to NAEP's P12.1, the standards do not specify solid, liquid, and gas—and they imply content on molecular attraction that NAEP states explicitly. In P12.12, the NAEP includes content on the increase of translational, rotational, and vibrational energy of atoms during heating, while the corresponding Texas statement, IPC (6)(B), addresses the movement of heat by convection, conduction, and radiation. In L12.6 the NAEP gives many details about chemical recombination during cycles and flows of matter and energy, while Texas bio (9)(D) contains a general statement about analyzing the flow of matter and energy.

Some Texas content statements imply content stated in the NAEP. For example, P12.9 states, "Energy may be transferred from one object to another during collisions," while Texas IPC (6)(A) implies this content with its statement describing the law of conservation of energy. In addition, L12.11 has content on modern ideas about evolution that provide a scientific explanation for the history of life on Earth, implied by Texas bio (7)(B), which states, "illustrate the results of natural selection in speciation, diversity, phylogeny, adaptation, behavior, and extinction."

Areas of nonalignment

Twenty-three NAEP grade 12 statements (47 percent) are unaddressed by Texas. This is mostly due to the fact that Texas grade 10 and 11 TAKS tests contain only content from biology, chemistry, and integrated physics and do not cover Earth and space science, while NAEP's grade 12 content includes physical science, life science, and Earth and space science. So there is no alignment between Texas and the entire NAEP section for Earth and space science.

In NAEP physical science, the unaddressed content statements are P12.2—the components of an atom, P12.4—neutral atoms and isotopes, P12.11—fission and fusion, P12.13—potential energy of an object on Earth's surface, P12.14—exothermic and endothermic reactions, P12.15—the conversion of very small amounts of matter into appreciable amounts of energy through nuclear reactions, P12.22—gravitational force, and P12.23—electric force.

In life science two NAEP statements are unaddressed by Texas: L12.4—plants' transformation of energy from light to sugar molecules to amino acids and organic molecules to larger molecules with biological activity, and L12.10—large variety of possible gene combinations resulting from sorting and recombination of genes in sexual reproduction.

Many NAEP content statements contain more content and are more detailed than Texas' corresponding content statements for grade 12

In Earth and space science all NAEP statements are unaddressed by Texas. NAEP's Earth and space science section includes E12.1—the origin of the universe and the "big bang" theory, E12.2—the formation of stars and galaxies, E12.3—nuclear reactions in stars and the formation of all elements, E12.4—relative and absolute dating, E12.5—the conclusion that the solar system formed from a nebular cloud 4.6 billion years ago, E12.6—early Earth, including bacteria and the composition of the atmosphere, E12.7—the influence of sporadic and gradual events on Earth's current structure, E12.8—theory of plate tectonics, E12.9—internal and external sources of energy in Earth systems, E12.10—systems that influence climate, E12.11—the movement and forms of elements, E12.12—the movement of matter through Earth's systems, and E12.13—the processes of natural ecosystems.

Areas where Texas standards go beyond the NAEP content statements

Texas has 40 "knowledge and skills statements" in the grade 10 and 11 TAKS information booklets. The NAEP does not address 18 (45 percent) of them: the 7 Texas statements in nature of science, 3 of the 8 in organization of living systems, 2 of the 8 in interdependence of organisms, 2 of the 8 in structures and properties of matter, and 4 of the 9 in motion, forces, and energy.

The NAEP does not address the nature of science (Objective 1) statements because it discusses inquiry in a section separate from the content statements, called "science practices," intended to crosscut all NAEP content.

In organization of living systems the NAEP does not address bio (8)(C)—characteristics of kingdoms including monerans, protests, fungi, plants, and animals, bio (10)(A)—the functions of systems in organisms, or bio (10)(B)—the interrelationships of organ systems.

In interdependence of organisms the NAEP does not address bio (4)(C)—structures and functions of viruses and cells and the role of viruses in causing various diseases and conditions or bio (4)(D)—the role of bacteria in maintaining health and in causing diseases.

In structures and properties of matter the NAEP does not address IPC (7)(A)—properties of fluids including density, viscosity, and buoyancy or IPC (8)(C)—investigating and identifying the law of conservation of mass.

To improve the assessment of the range of science knowledge and skills, the last two NAEP science frameworks have expanded the range of item types on the test

In motion, forces, and energy the NAEP does not address IPC (4)(D)—the mechanical advantage and efficiency of various machines, IPC (5)(A)—demonstrating wave types and characteristics through activities, and interpreting seismic wave data, IPC (6)(D)—economic and environmental impacts of various energy sources, or IPC (6)(F)—series and parallel circuits.

Summary of NAEP grade 12 alignment

No Earth and space science content statements in the NAEP are addressed by Texas because TAKS includes only biology, chemistry, and integrated physics. So the overall alignment between NAEP grade 12 and Texas high school TAKS tests is fairly low. But in the NAEP's physical and life science sections, Texas was most often partially aligned: Many NAEP content statements contain more content and are more detailed than Texas' corresponding content statements. In addition, Texas content statements often imply content explicitly stated by the NAEP. The overall alignment rating between only the physical and life science sections of the NAEP and the Texas content statements is 1.8. The overall alignment rating including Earth and space science statements—all with ratings of 1—is 1.6.

TEST SPECIFICATIONS ALIGNMENT

The assessment specifications alignment involved two parts: examining the types of items found in NAEP and in TAKS, and comparing NAEP's and TAKS' distribution of items between the different science strands.

Science is a discipline with a strong tradition of investigation, experimentation, and application of knowledge and skills. Before the 2005 assessment, NAEP science assessments consisted primarily of short-answer, paper-and-pencil questions that were mostly multiple-choice, which can only go so far in assessing skills. To improve the assessment of the range of science knowledge and skills, the last two NAEP science frameworks have expanded the range of item types on the test. In particular, the 2009 NAEP framework takes advantage of advances in educational measurement and the development of computer-based assessments. Due to the varying ways in which differing item types assess and reveal what students know and can do, the NAEP 2009 assessment specifications require

future NAEP tests to incorporate a range of item types, allowing students to reveal their understanding in ways beyond traditional selected-response methods. Multiple-choice items, short constructed-response items, extended constructed-response items, hands-on performance tasks, and interactive computer tasks will all be used to more accurately assess student knowledge, thinking, and skills.

Each type of assessment item demands a unique response from students (selecting a response from a set of alternatives, writing an explanation or justification, performing a virtual lab experiment). Individual items may draw on different types of stimuli (verbal, graphic, manipulative) to access the knowledge and skills required and may be scored in a variety of ways (right/wrong, partial credit, human scorers, computer software). By using several types of items the 2009 NAEP science assessment will require students to draw on multiple types of knowledge and a variety of skills for using and expressing that knowledge, thereby giving a more accurate picture of the breadth and depth of their learning. In this study, the following item types from NAEP were compared with the types in use from the states.

In multiple-choice items, students reflect on the material and then select an answer from a limited number of alternatives. Well constructed multiple-choice items can probe important facts, broad concepts, and themes of science, as well as deductive reasoning skills.

Constructed-response items, in which students answer without reference to a provided list of alternatives, include short constructed-response items and extended constructed-response items. Constructed-response items can provide insights into students' levels of conceptual understanding and assess their abilities to communicate about science. They can also be used to probe student abilities to generate information related to science content statements and their interconnections (how two or more cyclic events are related). Constructed-response items may be particularly useful for probing the practices of using scientific inquiry or using technological design (interpret given data or provide a solution to a real-world problem).

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 student abilities to combine science knowledge with the investigative skills reflective of the nature of science and inquiry.

By using several types of items the 2009 NAEP science assessment will require students to draw on multiple types of knowledge and a variety of skills for using and expressing that knowledge, thereby giving a more accurate picture of the breadth and depth of their learning

Interactive computer tasks in the 2009 NAEP science assessment may involve information search and analysis, empirical investigation, simulation, or concept mapping. The broad purpose of interactive computer tasks in this context is to tap performance expectations that are more advantageously assessed in a virtual format, such as scientific modeling of microscopic or temporal phenomena, repeated experiments, or simulations of hazardous or messy lab situations. Interactive computer tasks are intended as a complement to the hands-on performance tasks, not as a replacement.

The NAEP specifications also include two other types of items, item clusters and predict-observe-explain item sets. Item clusters are groups of related items that provide more in-depth analysis of student performance than would a collection of discrete, unrelated items. They can be particularly useful in exploring student conceptions, predictions, or explanations of the natural world. The predict-observe-explain item sets (White & Gunstone, 1992) describe a situation and ask the student to predict, observe, and/or explain the outcome, sometimes with additional supporting detail. Predict-observe-explain items may involve using science principles or the cognitive demand

of "knowing why (schematic knowledge)." Because these are really ways of clustering items and are not usually included in state test specifications, they were not used for comparison in this study.

The NAEP stipulates that 50 percent of student response time should be spent on multiple-choice items and the other 50 percent on constructed-response items (including short constructed-response, extended constructed-response, and concept-mapping tasks). Within these two categories are item clusters, predict-observe-explain item sets, hands-on performance tasks, and interactive computer tasks. There will be at least one item cluster, one predict-observe-explain item set, one hands-on performance task, and one interactive computer task at each grade level, and the total number of interactive computer tasks plus hands-on performance tasks will be at least four at each grade level.

The 2009 NAEP will have 50 percent of student response time allocated to multiple-choice items and 50 percent to constructed-response items. The Texas tests contain 100 percent multiple-choice items

Table 4 shows the percentages of various item types found in the NAEP and in Texas, with the number of Texas items in parentheses. The 2009 NAEP will have 50 percent of student response time allocated to multiple-choice items and 50 percent of student response time allocated to constructed-response items (short and extended). The current Texas tests contain 100 percent multiple-choice items. The number of items used to test students in science in Texas differs across grades, increasing by ten items in grade 8 and five more items in grades 10 and 11. In grade 5 there are 40 multiple-choice items with no short constructed-response or extended constructed-response items and no hands-on performance tasks. In grade 8 there are 50 multiple-choice items with no short constructed-response or extended constructed-response items and no hands-on performance tasks. And in grades 10 and 11 there are 55 multiple-choice items with no short constructed-response or extended constructed-response items and no hands-on performance tasks.

To consider how the state test coverage of the NAEP science topics matched, table 5 shows the proportions of testing time devoted to each of the three content areas for NAEP and for the Texas test. The first column of the table lists all the science topic areas that are included on the Texas test. The first three topic areas (physical, life, and Earth and space science) are those that are covered in NAEP, and the two topics below those (science as inquiry and science and the environment) are not separately assessed on the NAEP test.

Under the column heading for elementary school, three subcolumns are shown. The first shows the proportion of testing time devoted to each topic for the three NAEP topic areas. The second shows the proportion of testing time devoted to each of the four Texas topics at grade 5. The third shows the comparison of NAEP and Texas testing times for each of the three NAEP topics, a positive number if

TABLE 4
Percentages of different item types on the Texas science assessment

	NAEP	Texas		
NAEP item types	All grades	Grade 5	Grade 8	Grades 10/11
Multiple-choice items	50 percent	100 percent (40)	100 percent (50)	100 percent (55)
Short constructed-response items	50 percent			
Extended constructed-response items				
Hands-on performance tasks[a]	(≥1)			
Interactive computer tasks[a]	(≥1)			

a. Hands-on performance tasks and interactive computer tasks are combination items and can be categorized as multiple-choice or constructed-response.

TABLE 5
Approximate testing time allocated to different science topics on the Texas science assessment (percent of time)

	Elementary school			Middle school			High school		
	NAEP Grade 4	Texas Grade 5	Difference	NAEP Grade 8	Texas Grade 8	Difference	NAEP Grade 12	Texas Grade 11	Difference
Physical science	33.3	22.5	–10.83	30.0	24.0	–6.0	37.5	40.0	2.5
Life science	33.3	22.5	–10.83	30.0	24.0	–6.0	37.5	29.0	–8.5
Earth and space science	33.3	22.5	–10.83	40.0	24.0	–16.0	25.0	0.0	–25.0
Nature of science	0.0	32.5		0.0	28.0		0.0	31.0	

TABLE 6
Comparison of the proportions of testing time allocated to the NAEP science topics (percent of time)

	Elementary school		Middle school		High school	
	NAEP Grade 4	Texas Grade 5	NAEP Grade 8	Texas Grade 8	NAEP Grade 12	Texas Grade 11
Physical science	33.3	33.3	30.0	33.3	37.5	58.0
Life science	33.3	33.3	30.0	33.3	37.5	42.0
Earth and space science	33.3	33.3	40.0	33.3	25.0	0.0

the Texas test devotes more and a negative number if the NAEP devotes more. This pattern of columns is repeated for middle and high school. For this comparison Texas grade 11 (or exit level) was used instead of grade 10, because the grade 11 TAKS is the test given at the grade closest to the NAEP's grade 12 high school assessment.

At the elementary school level the proportion of Texas testing time is approximately 11 percentage points less in all three NAEP strands. However, the Texas test devotes 32.5 percent of testing time to the nature of science, which in the NAEP is tested as part of each of the three topic areas, not separately. For Texas at grade 8 testing time is 16 percentage points less for Earth and space science and 6 percentage points less in both physical and life sciences than in the NAEP. Nature of science accounts for 28 percent of Texas grade 8 testing time. At grade 11 the distribution of testing time in Texas exceeds that in the NAEP by 2.5 percentage points in physical science but is 8.5 percentage points lower in life science, and Earth and space science is not tested at all. Nature of science accounts for 31 percent of testing time at grade 11.

Table 6 ignores the testing time devoted to nature of science, which is not separately tested in NAEP, and shows how the testing time for the three NAEP topics compares with the time in the Texas state test. At the elementary school level there is no difference in the proportions of time. At grade 8 Texas devotes 33 percent of time equally to all three topics, whereas NAEP emphasizes Earth and space science by allocating it 40 percent of testing time compared to 30 percent for the other two topics. At grade 11 the Texas test does not address Earth and space science, which is covered in the NAEP with 25 percent of testing time. Physical science receives 58 percent of the time in Texas, and life science 42 percent.