

# ALIGNMENT AND THE STATES

## Three Approaches to Aligning the National Assessment of Educational Progress with State Assessments, Other Assessments, and Standards

Martha Vockley  
Vockley Lang



Prepared for the Council of Chief State School Officers  
and the U.S. Department of Education, National Center for Education Statistics  
August 2009

*This report was produced under U.S. Department of Education Contract No. ED-01-CO-0040/0009 with the Council of Chief State School Officers. The views expressed herein do not necessarily represent the positions or policies of the Department of Education. No official endorsement by the U.S. Department of Education of any product, commodity, service or enterprise mentioned in this publication is intended or should be inferred.*

## THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

The NAEP Informational Network at CCSSO is funded by the U.S. Department of Education to facilitate communication between the National Center for Education Statistics (NCES) and state assessment and policy leaders. States provide input to NCES on NAEP administration; items, frameworks, and trend lines; data reporting and use; templates and tools; as well as the design and state implications of NAEP studies commissioned by NCES.

Author  
Martha Vockley  
Principal and Founder, Vockley Lang, Reston, VA

## COUNCIL OF CHIEF STATE SCHOOL OFFICERS

Susan Gendron (Maine), President

Gene Wilhoit, Executive Director

Council of Chief State School Officers  
One Massachusetts Avenue, NW, Suite 700  
Washington, DC 20001-1431  
Phone (202) 336-7000  
Fax (202) 408-8072  
[www.ccsso.org](http://www.ccsso.org)

Copyright © 2009 by the Council of Chief State School Officers, Washington, DC  
*All rights reserved.*

## Contents

Introduction	page 4
Understanding Alignment	page 8
Descriptions of Three Approaches to Alignment	page 10
The NAEP ESSI Procedural Manual	page 10
The HumRRO Alignment Model	page 23
The Surveys of Enacted Curriculum Alignment Model	page 33
Conclusion	page 42
Features of the Three Alignment Approaches	page 43
References	page 46
Glossary	page 47
Appendix A: Webb’s “Depth-of-Knowledge” Model for Alignment	page 49

# Introduction

Since 1969, the National Assessment of Educational Progress (NAEP) has been the common yardstick for measuring the progress of students' education over time across the country. Teachers, principals, parents, policymakers, and researchers all use NAEP results to assess progress and develop ways to improve education in America.

Administered by the National Center for Education Statistics (NCES) at the U.S. Department of Education, NAEP is the only nationally representative and continuing assessment of what students in grades 4, 8, and 12 know and can do in various subjects. Assessments are conducted periodically in mathematics, reading, science, writing, the arts, civics, economics, geography, and U.S. history.

NAEP has taken on greater significance over the years. In the early 1990s, NAEP expanded its reporting of results to include descriptions of the comparative standings of state performance. The Elementary and Secondary Education Act, reauthorized in 2001 as No Child Left Behind (NCLB), took NAEP's role even further. Every state is now required to participate in NAEP mathematics and reading assessments for grades 4 and 8 to receive Title I funds for disadvantaged students. The law also holds states accountable for closing performance gaps, as measured on state assessments, between groups or subgroups of students.

State assessment results often are compared to NAEP. Since NAEP assessments are administered uniformly, the results serve as a common metric for all states and for selected large urban districts. NAEP can be used as a comparison measure for state-reported progress. NAEP is a benchmark of student achievement across the states in reading and mathematics.

## A Surge of Interest in Alignment

To make the relationships between NAEP and state education systems more explicit, governors, policymakers, and staff members of state education departments and large urban districts increasingly are interested in comparing their assessments, standards, and more (including content coverage, test items, and cognitive demand) to NAEP. State officials are major consumers of NAEP results and NCES reports that examine these results, trends over time, and aspects of performance. They want answers to key questions, such as:

- How does NAEP compare to state assessments?
- Does NAEP measure the same general knowledge and skills that state assessments measure?

- Does a level of “proficient” on state assessments mean the same thing as “proficient” on NAEP?
- Are students learning what is being tested on state assessments and on NAEP?
- Does the relative standing of a group or subgroup (or groups or subgroups) of students remain the same regardless of the assessment?
- How do NAEP results relate to and inform school improvement policy initiatives?

The theme of *alignment* runs through these questions. That is, how well do states’ assessments and standards match to NAEP?

In addition, states might be interested in finding out how well their state assessments and standards align with other tests, including:

- The Programme for International Student Assessment (PISA)
- The Progress in International Reading Literacy Study (PIRLS)
- The Trends in International Mathematics and Science Study (TIMSS)
- Commercially available standardized tests, such as the ACT or SAT
- Other measures they use or are evaluating for use

### Three Approaches to Alignment

To assist states in answering their questions, NCES supported the development of three approaches that could be useful for state alignment initiatives:

- **The NAEP ESSI Procedural Manual**—This manual is designed to help state or district officials think through their alignment questions, understand what they want to achieve, and choose an appropriate methodology (or alignment model). The manual also can be used to compare NAEP frameworks, specifications, and assessment items to state frameworks and assessments in any subject area.

Developed by the NAEP Education Statistics Services Institute (NAEP ESSI) at the American Institutes for Research, the NAEP ESSI procedural manual includes a sequence of three comparison approaches that could be used to evaluate the alignment between NAEP and state standards and

assessments. Or the initial planning protocol included with this manual could precede an alignment study using another model.

- **The HumRRO Model**—This model is designed to evaluate the alignment between NAEP and state assessments, frameworks, and items in mathematics and reading. It includes a series of eight tasks, with protocols and facilitation materials, for completing an alignment study.

The Human Resources Research Organization (HumRRO), an independent research and consulting organization, developed this model.

- **The Surveys of Enacted Curriculum (SEC) Model**—This model is designed to evaluate the alignment of standards, assessments, and classroom instruction in English language arts, mathematics, and science, with a focus on content. It includes Web-based tools for collecting, comparing, and analyzing data.

The Council of Chief State School Officers (CCSSO), in conjunction with several partner districts and states, developed this model.

## About This Report

This report will help states understand the three alignment approaches in terms of the questions they are designed to answer; the methodologies, tools, resources, analyses, and products they offer; and the amount of time and human resources they require. Additionally, this report includes:

- **A comparison table** that highlights the key features, resources, and procedures for conducting an alignment study with each of the three alignment approaches (see page 41)
- **A glossary** that will help states understand the different terminology used to describe assessments and alignment (see page 46)

State NAEP coordinators, state and district assessment and curriculum directors, and other state and district officials can use this information to compare and select an alignment approach or model that best matches their needs, expectations, and resources to achieve their alignment objectives. These alignment approaches also could prove useful to states interested in comparing their assessment systems to international assessments.

This is a descriptive report about three alignment models that NCES has supported. Other alignment models are available to states. The U.S. Department of Education does not endorse any model. In addition, while this report responds

to states' interest in alignment, it is not intended to imply that states need to make any changes to their curriculum.

### **Accessing and Using the NAEP Assessment Items**

State and districts interested in conducting an alignment study will need access to the NAEP assessment items, or questions. To maintain assessment security, the operational pool of items—the assessment questions still in use—is not normally made available for use in alignment studies.

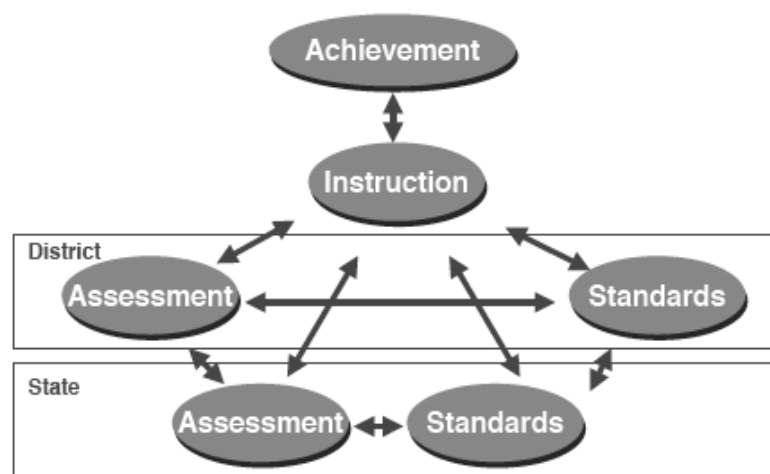
However, about one-third of the pool of items for any given subject and grade are made public after each administration of the assessment. These released items, along with their framework classifications and performance data, are available via the NAEP Questions Tool (<http://nces.ed.gov/nationsreportcard/itmrls/>). More than 2,000 questions are currently available.

States and districts can select from these released items for alignment studies. In making this selection, it is important to understand that items from any given assessment administration address a *sample* of objectives, not all of the objectives, in the corresponding NAEP framework for the given subject and grade. It is therefore recommended that states and districts use blocks of released items from multiple years of assessment administrations whenever possible to get the best representation of NAEP items for alignment studies.

# Understanding Alignment

Alignment is the “core idea” in systemic, standards-based education reform (Porter, 2002; Smith & O’Day, 1991). Within their states and districts, policymakers and educators strive for alignment between standards and curriculum, standards and instruction, and standards and assessments. Figure 1 shows these relationships in terms of vertical and horizontal alignment:

**Figure 1. Vertical and Horizontal Alignment**



Source: Porter, A. (2002)

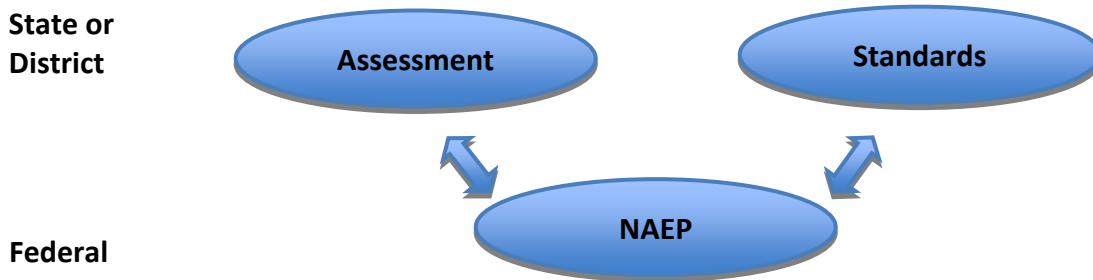
The point of alignment is to create a coherent educational system that conveys a clear and unified message about expectations and goals. A coherent educational system should inform efficient, effective instruction that is focused on “what matters,” as defined by the standards, and motivate student achievement.

Curriculum standards articulate broadly what students should know and be able to do in a given subject area. Benchmarks and indicators add specificity to standards by spelling out grade-level or age-span expectations of how well students should know or demonstrate the content. The curriculum encapsulates the content of the standards, benchmarks, and indicators. The standards and curriculum make up the content of instruction. Assessments measure how well students have learned the content.

NAEP alignment studies can reveal similar relationships to state standards and assessments, as shown in figure 2:



**Figure 2. Relationships for NAEP Alignment Studies**



The purpose of NAEP alignment studies is to determine how well state content standards align with NAEP test frameworks, test items, and/or type of cognitive demand or level of difficulty, as measured by student performance. NAEP alignment studies allow states or districts to compare their expectations to the content of the national assessment.

Measurement research and alignment methodologies have become more sophisticated over time, which makes possible more fine-grained—and thus more useful—comparisons and analysis. The three alignment approaches described in this section share some common research foundations and methodological approaches. They are, however, distinct, in that they are designed to answer different alignment questions, use different methodologies and procedures, and focus on particular aspects of alignment.

The “depth-of-knowledge” alignment model, developed by Norman Webb of the Wisconsin Center for Education Research, is groundbreaking work that has been influential in developing and refining many other approaches and models for aligning standards with assessments (Webb, 1997; Webb 1999; Webb 2002; Webb and Smithson, 1999).

Depth-of-knowledge models describe the cognitive demand, or types of thinking or reasoning, required to perform tasks. Appendix A describes Webb’s depth-of-knowledge levels—recall and reproduction, skills and concepts, strategic thinking, and extended thinking.

# Descriptions of Three Alignment Approaches

## The NAEP ESSI Procedural Manual

### Focus of NAEP-Supported Alignment Studies

- Grades 8 and 12 mathematics

### Applicability for State and District Alignment Studies

- Any subject
- Any grade level

The NAEP ESSI procedural manual (*Comparing National Assessment of Educational Progress Frameworks, Specifications, and Assessment Items to State Frameworks and Assessments: A Procedural Manual*) was developed by the NAEP Education Statistics Services Institute at the American Institutes for Research. The manual is a sequence of procedures for comparing NAEP frameworks, specifications, and assessment items to state frameworks and assessments. It features a decision tree with a series of questions that, together, create a decision-making tool for planning and conducting an alignment study.

In a pilot study, NAEP ESSI worked with three states beginning in 2003 to develop the procedures, which lead state or district officials through an in-depth process for comparing subject-area content measured by NAEP and state assessments. The study focused on grade 8 mathematics, but the procedures could be used for any subject or grade level.

In 2009, the manual served as the basis for a second study, which compared the content of the 2005 and 2009 NAEP grade 12 mathematics frameworks and item pools. As a result of the second study, refinements were made to the manual and some procedures were added.

## Methodology of the NAEP ESSI Procedural Manual

The NAEP ESSI procedural manual begins with a **Plan for the Comparison**. Time and resources available to conduct an alignment study often are limited. Decisions about how to allocate resources need to be made for each step of the process. The steps will depend on the kind of comparison to be conducted.

Good planning, therefore, is a critical first step—and one that will yield answers to questions about assessments, instruction, and student performance that typically prompted the decision to conduct a comparison in the first place.

The Plan for Comparison begins with these **Key Questions**:

- What is the intent of the comparison? (Why are we doing this? What do we want to know?)
- What comparisons between NAEP and state frameworks and assessments will be conducted?
- Which state documents and state assessment items will be used in the comparison?
- Who will participate in the comparison?
- How will participants in the comparison procedure become familiar with the documents and item pools to be compared?

Each Key Question includes **Considerations**, a series of probing questions and guidance that help state or district officials clarify their purposes, select an appropriate alignment methodology (or methodologies), and understand and plan the scope of work. The Key Questions will help state or district officials decide what kind of comparison(s) they need and whether the NAEP ESSI manual—or some other alignment model—is appropriate. A sample of a completed Key Questions documents is shown in exhibit 1.

## **Exhibit 1: Sample of Completed Key Questions**

### **What is the intent of the comparison? (Why are we doing this? What do we want to know?)**

The comparison is being done to determine if there are key concepts assessed by NAEP that our current state assessments do not include by grades 4 or 8. The results of the comparison will be used along with other pieces of information for two purposes:

1. to provide contextual information for the state superintendent in preparation for the next release of NAEP results, and;
2. as a discussion starter for the state's assessment advisory panel as it considers whether revisions are needed to our state's current grades 3–8 mathematics assessment blueprints and the content expectations for algebra 1 end-of-course exam.

### **What comparisons between NAEP and state frameworks and assessments will be conducted?**

We will be comparing the 2009 NAEP specifications documents for grades 4 and 8 with our state's mathematics content expectations for grades 3–8 and our algebra end-of-course exam. In addition, we will compare how NAEP incorporates cognitive skills, such as reasoning and problem solving, as compared to our state's approach. We will *not* be comparing assessment item pools at this time, but may revisit this decision at a later date depending upon results of the specifications/content expectation comparisons.

### **Which state documents and state assessment items will be used in the comparison?**

- *2009 NAEP Mathematics specifications document*
- *State assessment blueprints for grades 3–8*
- *Content expectations for state algebra 1 end-of-course exam*

### **Who will participate in the comparison?**

Since the resulting document will be used to help initiate a discussion of potential revisions to content expectations, our state has decided not to commit a tremendous amount of outside resources to the alignment effort; thus, the comparison will be done by our state's elementary and high school mathematics state supervisors, our Title I mathematics coordinator, and the NAEP state coordinator.

### **How will participants in the comparison procedure become familiar with the documents and item pools to be compared?**

The NAEP state coordinator will train staff on the NAEP mathematics specifications document. The mathematics state supervisors will do a quick refresher on the various state blueprints. The Title I mathematics coordinator will be responsible for training the participants on the comparison procedure that will be used, identifying special cases that might arise, and facilitating the discussion of decision rules that will be needed before the process starts.

Source: NAEP ESS1

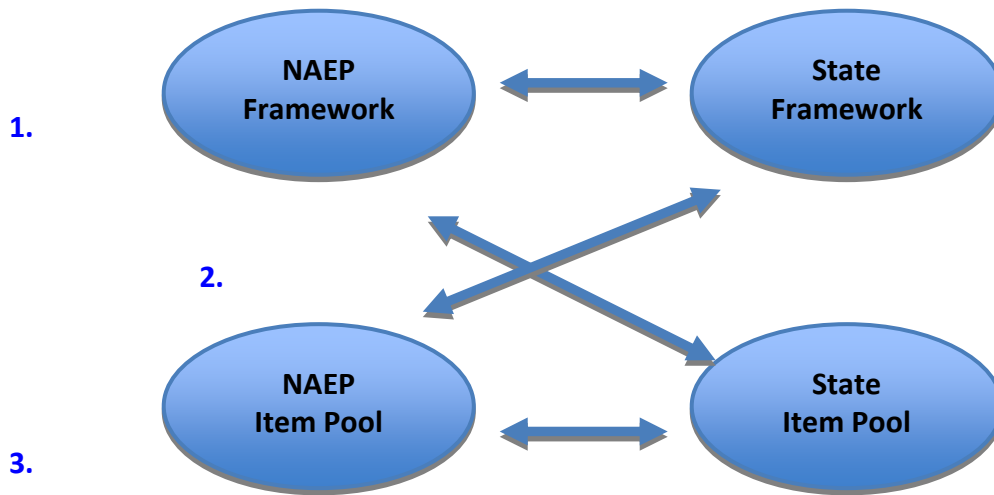
Next, the NAEP ESSI manual provides **Procedures** for investigating content similarities and differences between NAEP and state assessments. The procedures build on existing procedures that were used to compare NAEP with international assessments, such as PISA, PIRLS, and TIMSS.

There are three components to the process in the NAEP ESSI manual, which are described in more detail below:

1. **Comparing frameworks** to understand the descriptions of content to be assessed on NAEP and on state assessments.
2. **Cross-classifying assessment items to frameworks** to investigate similarities and differences between NAEP and state assessments by comparing *items* from one assessment to the framework *objectives* of the other.
3. **Comparing attributes of NAEP and state assessment items** to discover similarities and differences between NAEP and state assessments items.

The sequence starts with a comparison of frameworks, moves to an item-to-framework comparison, and ends with an item-to-item comparison. Thus, it supports increasingly more fine-grained analysis. Depending on the questions they want to answer, state or district officials can go through the entire three-component process or select only the components that meet their needs. Figure 3 illustrates the three components of comparison. For example, the NAEP framework could be compared to the state framework or to the state item pool.

**Figure 3. Three Components of Comparison in the NAEP ESSI Procedural Manual**



Source: NAEP ESSI

### **Component 1. Compare Frameworks**

The purpose of assessment development documents, known as *frameworks* or *specifications*, is to define the boundaries of the subject-area domain to be assessed. Frameworks are based on beliefs—intentions that may or may not be explicitly stated—about content expectations and assessment.

A first step in aligning NAEP with other assessments often is a comparison of frameworks to discern the subject-area expectations of each assessment program: Are NAEP and the state assessment intended to measure the same content in the same way?

This component of the NAEP ESSI procedural manual examines NAEP and state frameworks by mapping one framework onto the other. This comparison is designed to:

- Pinpoint similarities and differences in the descriptions of the subject-area domain
- Inform decisions about whether to conduct further investigation of the similarities and differences
- Provide background information for the next two components of the comparison process

This framework-to-framework comparison includes key questions, considerations and guidance, and step-by-step procedures. Completing this process will show the points at which NAEP and state subject-area domains intersect, as well as points of divergence or incomplete agreement. Potential products of this analysis include:

- **A map or grid** of subject-area domains that illustrates schematically the similarities and differences between NAEP and state frameworks, as shown in table 1
- **A quantitative summary** of the similarities and differences of the subject-area domains assessed by NAEP and state frameworks, as shown in exhibit 2
- **A qualitative summary** of the similarities and differences between how the subject-area domain is assessed by NAEP and the state assessment, as shown in exhibit 3

**Table 1. Excerpts from a Sample Framework-to-Framework Mapping Matrix**

State X Framework	NAEP 2005 Mathematics Framework—Grade 8	Comparison Comments
Generate a different representation of data such as a table, graph, equation, or verbal description.	Translate between different representations of linear expressions using symbols, graphs, tables, diagrams, or written descriptions.	Objectives are the <b>same</b> ; NAEP limits the objective to linear expressions.
Write, simplify, and evaluate expressions	Write algebraic expressions, equations, or inequalities to represent a situation.  Perform basic operations using appropriate tools on linear algebraic expressions (including grouping and order of multiple operations involving basic operations, exponents, roots, simplifying, and expanding).	Objectives are <b>similar</b> ; NAEP expands the objective to include writing equations and inequalities. NAEP has a separate objective for evaluating expressions but limits the operations to linear algebraic expressions.
	Interpret the meaning of slope or intercepts in linear functions.	<b>No match</b> ; there is no corresponding state objective.

Source: NAEP ESS1

**Exhibit 2. Excerpt from a Sample Framework-to-Framework Quantitative Summary**

“Mapping of the geometry content area revealed **five** objectives that were similar, **two** objectives that are assessed at different grade levels by the state, and **three** objectives in NAEP that were not assessed by the state at any grade level.”

Source: NAEP ESSI

**Exhibit 3. Excerpt from a Sample Framework-to-Framework Qualitative Summary**

Two significant differences were found in the algebra content area for grade 8:

- NAEP focuses on linear expressions, equations, and inequalities. Our state extends this focus to quadratic equations with positive integer roots.
- NAEP assesses ability to interpret the meaning of slope or intercepts of linear equations, a topic our state does not assess until the end-of-course assessment in algebra [at the end of grade 8, another grade, or whenever a student takes algebra?]

Source: NAEP ESSI

**Component 2. Cross-Classify Items to Frameworks**

The framework comparisons described above provide insight into how NAEP and states broadly define the domain of an assessed subject area. Comparing items to frameworks provides more in-depth information, including whether the frameworks (intentionally or unintentionally) accommodate different item content and characteristics.

Cross-classifying state assessment items to the NAEP framework—and/or NAEP items to the state framework—illustrates potential differences in interpretation of framework objectives by different audiences, including teachers, curriculum specialists, and assessment developers. This comparison further delineates each assessment’s domain. Cross-classifying items to frameworks can:

- Provide information about the *constructs* and *dimensions* of the assessments (constructs refer to the underlying trait—knowledge, skills—



an assessment is intended to measure; *dimensions* refer to the content of a subject and the *cognitive demand*, or types of thinking or reasoning processes, to be assessed)

- Demonstrate whether framework objectives clearly describe the content to be measured
- Reveal areas where items from one assessment's item pool fulfill a particular objective or group of objectives in the other assessment framework
- Identify NAEP items that assess content not included in the state's framework and state items that test content not included in the NAEP framework

This component of the NAEP ESSI procedural manual also provides key questions, considerations and guidance, and step-by-step procedures for conducting two types of cross-classifications:

- **Option 1**, to determine whether NAEP items would appear on a state assessment
- **Option 2**, to determine whether state items would appear on NAEP

States or districts interested in understanding the subject-area domains assessed by NAEP and the state, respectively, and in improving their ability to synthesize and distinguish the two domains for curriculum development purposes should complete both options.

Potential products of this component of the NAEP ESSI procedural manual include:

- **Cross-classifications of released NAEP items to state frameworks**, as shown in table 2
- **Cross-classifications of state items to the NAEP framework**, as shown in table 3
- **Quantitative summary** of items that assess NAEP and state content objectives, along with identification of content objectives that NAEP and the state do *not* have in common, as shown in table 4.

### Degree of Match Coding Scheme

This component of the NAEP ESSI procedural manual includes a new coding scheme, known as **Degree of Match** (Mueller and Gattis, 2004), piloted in the three-state study. Degree of Match builds on Norman Webb's depth-of-knowledge model (see Appendix A) to rate how closely an assessment item matches the content of one or more of the objectives it is intended to measure.

### Degree of Match Categories

- **Full Match:** The item fully assesses a single objective (or part of a single objective).
- **Cumulative Match:** The item fully assesses multiple objectives (or parts of multiple objectives).
- **Minimal Match:** The item only partially assesses the intent of a single objective.
- **Limited Match:** The item only partially assesses the intent of multiple objectives.
- **No Match:** There is no objective at any grade level that describes the item.

**Table 2. Excerpts from a Sample Cross-Classification:  
NAEP Items to State Framework**

NAEP Item Number	Content Classification in State Framework				Other Dimensions	Summary Comments
	Standard	Topic	Indicator	Grade Level	Degree of Match	
A-1	1	A	1	8	Full	Item asks students to find a remote term
A-2	1	A	1	8	Cumulative	Item combines two objectives: recognizing patterns and different representations
A-3	1	A	1	8	Full	Non-numerical pattern

Source: NAEP ESSI

**Table 3. Excerpts from a Sample Cross-Classification:  
State Items to NAEP Framework**

	Content Classification in NAEP Framework				Other Dimensions		
State Item Number	Content Area	Subtopic	Objective	Grade Level	Degree of Match	Levels of Complexity	Summary Comments
A-1	A	1	a	8	Full	Moderate	Finding an expression for the “nth” term might be a way to shift the item to high complexity
A-2	A	1	a	8	Full	Moderate	Pattern is not uniquely defined

Source: NAEP ESSI

**Table 4. Excerpts from a Sample Quantitative Summary**

Content Area	NAEP Grade 4 Items Classified to State Grade 3 Objectives	NAEP Grade 4 Items Classified to State Grade 4 Objectives	State Grades 3 Items Classified to NAEP Grade 4 Objectives
Number	10	42	33
Measurement	5	25	14
Geometry	6	21	12
Data and Probability	2	26	15
Algebra	3	32	21
No Grade-level Match	—	8	7

Source: NAEP ESSI

### **Component 3. Compare Attributes of NAEP and State Assessment Items**

The third and final component of the NAEP ESSI procedural manual is the analysis of the similarities and differences in the assessment items themselves. Examining the *attributes*, or characteristics, of individual items or groups of items is more thorough than an investigation limited to frameworks and assessment instruments. This examination yields a deeper understanding of whether NAEP and the state assess similar content objectives in the same way and at the same level of difficulty.

Item-to-item comparisons can:

- Pinpoint similarities and differences in the content and cognitive demands of individual items or groups of items
- Contribute to a fuller description of similarities and differences between NAEP and state assessments
- Be used to clarify whether NAEP and state assessments measure a particular section of a subject-area domain in the same way and at the same level of cognitive demand

Finally, combining performance information with item attributes may uncover patterns of student performance clustered across objectives or groups of objectives that illustrate differences in the content and cognitive demands of NAEP and state assessments.

This component of the NAEP ESSI procedural manual provides key questions, guidance, and step-by-step procedures for conducting two types of item-to-item comparisons:

- **Option 1**, to determine if NAEP and state assessments measure the same construct in the same way
- **Option 2**, to examine content and construct similarities and differences within items on which groups of students perform at the same level of difficulty

Potential products of this component of the NAEP ESSI procedural manual include:

- **Summary statements** comparing attributes of NAEP and state items aligned to the same objective or group of objectives, as shown in exhibit 4
- **Summary statements** comparing subject-area content and cognitive demands on NAEP and state assessments when students are performing on the assessments at similar performance levels, as shown in exhibit 5

#### Exhibit 4. Sample Summary Statements Comparing NAEP and State Item Attributes

- State and NAEP items aligned to our state’s grade 4 computation objectives used different kinds of numbers. State items used only whole numbers while NAEP items used both whole numbers and simple fractions.
- NAEP items aligned to our state’s grade 5–8 probability objectives use fractions, decimals, and percents to represent probability. State items use fractions only at grades 5 and 6, and fractions, decimals, and percents for grade 7 and 8 items.
- Our state classifies all grade 4 items involving number sentences under the “solve equations” objective in algebra, while NAEP tends to classify number sentence items in “number.”
- In geometry, all of the grade 8 NAEP items with  $p$ -values\* that are higher than the highest grade 8 state item  $p$ -value test concepts that the state would test at lower grade levels.

Source: NAEP ESS1

\* Note: A  $p$ -value is the proportion of sampled students who responded correctly to a test question. The  $p$ -value actually represents item “easiness,” but traditionally psychometricians refer to it as item difficulty. The NAEP Question Tool states: “Difficulty is a measure of student performance on a question. Multiple-choice or constructed-response questions scored either right or wrong are rated ‘easy’ if answered correctly by 60 percent or more of students, ‘medium’ if answered correctly by 40 to 59 percent of students, or ‘hard’ if answered correctly by fewer than 40 percent of students.”

**Exhibit 5. Sample Summary Statements Comparing NAEP and State Content and Cognitive Demands**

- The majority of the state's grade 8 number items have a  $p$ -value greater than 70%. This is not true for NAEP items.
- More than 75% of the 8th-grade algebra items with  $p$ -values of 60% or more involve solving only single-step equations. This is not true for NAEP items.
- All state grade 4 items with  $p$ -values of 10% or less are in the data and probability content area. That does not appear to be true for NAEP.

Source: NAEP ESSI

**Learn More**

To learn more about the NAEP ESSI model, visit <http://www.air.org>.

## The HumRRO Alignment Model

### **Focus of NAEP-Supported Alignment Studies**

- Grades 4 and 8 mathematics
- Grades 4 and 8 reading

### **Applicability for State and District Alignment Studies**

- Grades 4 and 8 mathematics
- Grades 4 and 8 reading

The Human Resources Research Organization, an independent research and consulting organization, developed the HumRRO model, an alignment method for examining similarities and differences between NAEP and state frameworks or standards and assessments in reading and mathematics.

The HumRRO model is the result of two alignment studies. In the first study, conducted in 2003, HumRRO supported one state in comparing its assessments, items, and standards to NAEP assessments, items, and “content expectations” (comparable to state standards). For this study, HumRRO:

- Created a series of alignment tasks (discussed below)
- Facilitated workshops in which teachers and other educators completed alignment tasks to compare NAEP to state standards and assessments
- Completed alignment tasks involving test administration and scoring
- Prepared the final report of the alignment findings for the state

In the second study, conducted in 2006 with a second state, HumRRO refined the alignment tasks and developed resources that states can use to conduct alignment studies independently. These resources include:

- Background information and a facilitator’s guide for workshop participants
- An alignment guide with descriptions of the alignment tasks
- Directions and worksheets for completing and documenting the tasks

- Reference material on the NAEP reading and mathematics content expectations, frameworks, and released items, plus a link to the NAEP Web site to retrieve updated information
- An electronic “report shell,” or template, for preparing the final alignment document

HumRRO piloted these resources in the second study and, based on evaluations from participants, refined its methodology and resources to support their use in other states.

## Methodology for the HumRRO Model

The HumRRO model entails eight tasks:

- **Tasks 1–5** focus on comparing state standards and released items to NAEP.

For these five tasks, the HumRRO model employs a three-day workshop format, with two groups of participants (one for reading and one for mathematics). Each group has five members: four teachers, coaches, or other school or district educators with expertise in their state’s reading or mathematics standards and assessments, plus one state department of education official who serves as the group’s facilitator. Another state official serves as administrator to both groups and may participate in the alignment tasks as well.

HumRRO recommends that participants represent a range of expertise and grade spans that reflect the scope of the alignment. For example, to compare state assessments to NAEP at grades 4 and 8, elementary and middle school educators should be included to provide comprehensive information about the content of their state standards and assessments.

- **Tasks 6–8** focus on comparing state *operational* (currently used) items, and assessment administration and scoring procedures, to NAEP.

State department of education officials complete Tasks 6–8. These tasks require more specialized knowledge about assessments that teachers might not have, such as comparisons of test administration and scoring between state assessments and NAEP. Task 6, in particular, may require access to secure testing documents; limiting access to department officials is advised to protect assessment security.



## **Eight Tasks for Comparing State Assessments to NAEP**

### **Task 1. Matching State Standards and NAEP Content Expectations**

Participants match their own state standards to NAEP content expectations and create four categories:

- Exact match
- Partial match
- Unique state standard
- Unique NAEP content expectation (no match)

In addition, participants rate each partial match on a scale of 1 to 4 points, with 1 representing a strong partial match; 2, a slightly less strong partial match; 3, a weak partial match; and 4, a very weak partial match. Exhibit 6 shows a worksheet that participants use for this task.

The purpose of this task is to document the degree of match between state standards and NAEP content expectations. Participants tabulate the results to determine the percentage of match, and the degree of match, between the state standards and NAEP.

Task 1 results also form the foundation of Task 6.

## Exhibit 7. Worksheet for Matching the NAEP Framework to State Standards

Match: E=Exact, P=Partial, #=Strength of match

NAEP ID No.	NAEP Description	State Standard and Description	Match			Comment
			E	P	#	
8.L.01.A	What is the moral in the story? Use evidence from the story in your response.					
8.L.01.B	How does the setting help to illustrate the theme of the story?					
8.L.01.C	Do you think the lesson in this story is true today? Why or why not?					
8.L.01.D	Explain what makes this story a fable.					
8.L.02.A	What was the major character's opinion of ____?					
8.L.02.B	What causes the main character to do ____? Use evidence from the story in your response.					
8.L.02.C	How do you think the character's actions might be different today? Support your response with evidence from the story.					
8.L.02.D	How does the author's description of ____ help explain the character's actions?					

Source: HumRRO

### Task 2. Matching State and NAEP Released Items onto the NAEP Matrix

Participants match state and NAEP released items to one another in terms of their content strand and cognitive complexity, for mathematics, and in terms of their context and aspect, for reading.

The purpose of this task is to show how NAEP items are structured, as well as how state items “fit” onto the NAEP matrix (or framework) structure. Participants place matching items on a matrix and document which items are placed on which cell of the matrix. Exhibit 7 is a photograph of how participants in one group sorted excerpts of frameworks and content standards into NAEP matrix groupings.

**Exhibit 7. Excerpts of State and NAEP Release Items Sorted into NAEP Matrix Groupings**



Source: HumRRO

**Task 3. Sorting State and NAEP Items onto a State Taxonomy**

Participants sort the state and NAEP released items onto the cognitive taxonomy—the classification of the categories of thinking or reasoning skills required by items—that the state uses. If the state has no preferred taxonomy, participants use Webb’s depth-of-knowledge model, described in Appendix A. A training manual for using this model (Webb, Alt, Ely, and Vesperman, 2005) is included in the HumRRO alignment guide.

The purpose of this task is to compare the cognitive complexity of items, using a taxonomy with which the participants are familiar, if possible. They document the taxonomy levels and the NAEP and state items they assign to these levels.

Exhibit 8 shows a worksheet used to document the results of this task.

## Exhibit 8. Worksheet of Cognitive Complexity

### Task 3 Worksheet and Results Documentation—Reading

You will need two copies of this worksheet, one for Grade 4 and one for Grade 8.

As you place state released items in the State taxonomy matrix, ask the following two questions:

1. What State reading context does this item best represent?
2. What State taxonomy level of reading does this item best represent?

Document results by filling in the number of **state released items** that were placed in each cell (use this matrix, or modify as needed to best fit your state's taxonomy):

	Cognitive Complexity Level			
State reading context	Level 1	Level 2	Level 3	Level 4

How many state released items were unable to be placed in any of the cells? \_\_\_\_\_ items

What were the reasons these state items could not be matched?

Document results by filling in the number of **NAEP released items** that were placed in each cell:

	Cognitive Complexity Level			
State reading context	Level 1	Level 2	Level 3	Level 4

How many NAEP released items were unable to be placed in any of the cells? \_\_\_\_\_ items

What were the reasons these NAEP items could not be matched?

Source: HumRRO

### Task 4. Comparing State and NAEP Item Format

Participants use a series of guiding questions to compare test items according to:

- Question type (multiple choice or constructed response)
- Formatting

- Graphic elements

This task allows participants to consider differences in item structure—and how these differences might affect assessment scores. For example, how might students perform on NAEP, which has a combination of multiple-choice and constructed-response items, if their state assessment has only multiple-choice items? Participants document differences on a worksheet.

### **Task 5. (Reading Only.) Comparing State and NAEP Reading Passages**

Participants examine state and NAEP reading passages that accompany reading items. They compare reading passages according to length, difficulty, topics, bias, and stereotyping, and document differences on a worksheet.

### **Task 6. Linking Operational State Assessment Items to NAEP Content Expectations**

State department of education officials link current state assessment items to NAEP content expectations. To complete this task, they use the state’s item specification document and the state standards-to-NAEP matching results from Task 1.

Item specifications typically report specific state standards that have been assigned to specific assessment items. First, officials count the number of items that represent a particular state standard. Next, when all of the items have been accounted for, officials use the results of Task 1 to determine the relationship between state standards and NAEP content expectations on items. For example, officials may determine that 16 items are assigned to state standard 1. By using Task 1 results, they know that state standard 1 is an exact match to NAEP content expectation X.

An automated feature in the report shell converts this information into a series of bar graphs showing the number of items for each state standard. Some states allow state test items to be coded with more than one standard (as primary, secondary, or tertiary standards). In these instances, the left bar of the double bar indicates how many state items are coded with the standard used as the primary standard. The right bar indicates how many state items are coded with the standard used as primary, secondary, or tertiary standards. The color of the bar graph indicates the degree of match to a NAEP content expectation, with black indicating an exact match; gray, a partial match; and white, no match.

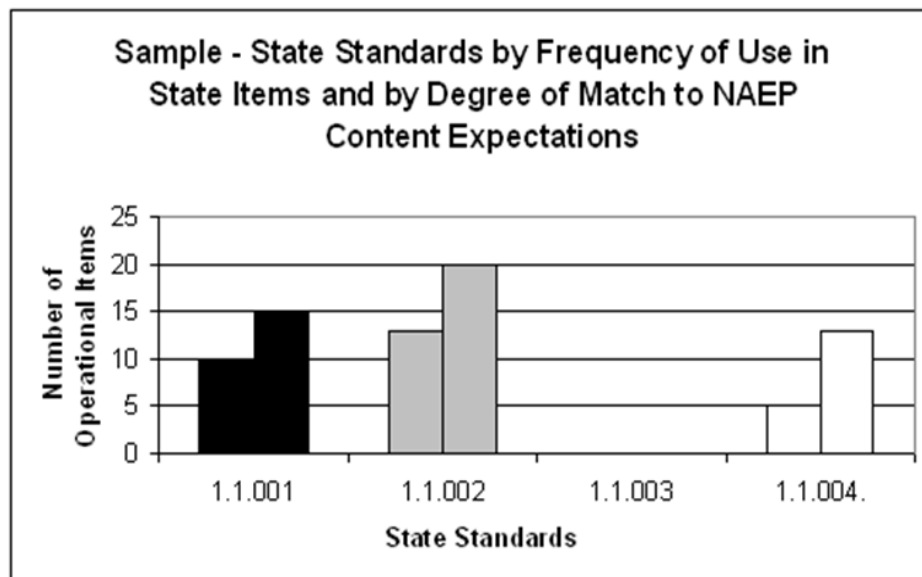
The results of this task will show how well state items match to NAEP content expectations. Exhibit 9 shows a sample item content table, which indicates how well several state standards match to NAEP. Exhibit 10 shows the same information as a bar graph.

## Exhibit 9. Item Content Table

Item Content Table (example)								
State standard number	Frequency of use as:				Links to NAEP content expectation(s):	Quality of match		
	Primary	Secondary	Tertiary	Total		Exact	Partial (1-4)	Unique
1.1.001	10	2	3	15	4.N.1.a	yes		
1.1.002	13	7	0	20	4.N.1.b		3	
1.1.003	0	0	0	0	4.N.2.a	yes		
1.1.004	5	2	6	13	Unique			yes

Source: HumRRO

## Exhibit 10. Item Content Graph



Source: HumRRO

## Task 7. Comparing Test Administration Procedures

Officials answer a series of guiding questions that will help them compare administration procedures for state assessments and NAEP. For example, NAEP tests only a sample of students in a state, but all students may be required to take the state assessment.

This task provides officials with a more complete understanding of differences between the administration procedures of state assessments and NAEP. Exhibit 11 shows a worksheet officials use to summarize the results of this task.



## Exhibit 11. Worksheet for Comparing Test Administration Procedures

Test Administration Worksheet Summary				
Use this table to summarize the results of the test administration analysis.				
+	Questions	No difference (A)	Minor difference (B)	Major difference (C)
	1. Who administers?			
	2. Who is accountable?			
	3. Test consequences?			
	4. Test format?			
	5. Number of tests taken?			
	6. Who takes—sample/all?			
	7. Alternate assessment?			
	8. Special education inclusion?			
	9. Test preparation?			
	10. Accommodations?			
	11. Time limits?			
	12. Manipulatives permitted?			
	13. Calculators permitted?			
	14. When tested?			
	15. Testing in other languages?			

Source: HumRRO

### Task 8. Comparing Test Scoring Procedures

Officials answer a series of guiding questions that will help them compare scoring procedures for state assessments and NAEP. For example, how are state assessment scores reported—at the student level, or by grade, school, or district? How do these results compare to the way NAEP scores are reported?

State officials input the documented results from the eight tasks into the report shell that is included with the HumRRO materials. The report shell uses a standard report format (Background, Methodology, Results, Discussion, and Recommendations). Much of the background information is included in the report shell, while the remaining sections contain suggested subheadings for the state's analysis and discussion.

A separate Excel® spreadsheet file for creating bar graphs to display Task 6 data is included in the report shell, as shown in exhibit 12. The graphs can then be copied and pasted into the final report.

## Exhibit 12. Worksheet for Comparing Test Scoring Procedures

Test Scoring Worksheet Summary			
Use this table to summarize the results of the test scoring analysis.			
Questions	No difference (A)	Minor difference (B)	Major difference (C)
1. Scoring—hand/machine?			
2. Multiple test forms?			
3. Item types?			
4. Weighting of items?			
5. Score computation?			
6. Methodology used?			
7. Frequency of equating?			
8. Developmental scaling?			
9. Level of reporting?			
10. Trends?			
11. Equating across forms?			
12. Performance categories?			
13. Tracking student performance?			
14. Who gets scores?			
15. Use of scores?			
16. Special ed scores—how used?			
17. Use of rewards?			
18. Other factors in scoring?			

Source: HumRRO

### Learn More

To learn more about the HumRRO model, visit  
<http://apps.humrro.org/NAEPAMG/>.



## The Surveys of Enacted Curriculum Alignment Model

### Focus of NAEP-Supported Alignment Studies

- Grades 4 and 8 mathematics

### Applicability for State and District Alignment Studies

- For NAEP alignment studies, grades 4 and 8 in English language arts, mathematics, and science
- For curriculum alignment studies, any grade level for English language arts, mathematics, and science

The Council of Chief State School Officers (CCSSO), a nationwide, nonprofit nonpartisan organization, developed the Surveys of Enacted Curriculum (SEC) in conjunction with several partner districts and states.

The SEC model is intended primarily to help teachers, administrators, and policymakers answer questions like these:

- Is what we teach truly aligned with state standards—the content students should learn?
- Do teachers devote the right amount of instructional time to the right content?
- Is there a connection between current instructional practices and low performance relative to certain standards?
- Are instructional practices consistent with prevailing research on effective practices?
- What types of professional development do teachers need?

The SEC model has been used to analyze the content and alignment of standards, curriculum, and assessments in mathematics, English language arts, and science in more than 30 states. CCSSO has developed the SEC model into a Web-based tool that provides consistent data and graphic presentations of instructional practices and content (the “how” and the “what”) actually being taught in classrooms.

In 2007, CCSSO and its research contractor, the Wisconsin Center for Education Research (WCER), used the SEC model to conduct an alignment study between

the 2007 NAEP mathematics assessments for grades 4 and 8 and state standards and assessments.

## The SEC Methodology

To develop SEC, CCSSO worked with WCER in a collaborative project involving educators, researchers, and subject-area specialists. The project incorporated research insights, survey instruments, and data reporting methods developed by Andrew Porter, now dean of the Graduate School of Education at the University of Pennsylvania and former director of WCER, and his colleague at the center, John Smithson (Blank, Porter, and Smithson, 2001; Porter, 2002).

Porter's alignment methodology focuses on content, which comprises:

- The **"content of instruction,"** defined as the decisions that teachers make about what to teach (and how), how much time to spend on a particular subject, what topics to cover, when and in what order, to what standards of achievement, and to which students. These decisions and their implementation make up the content of instruction, which plays a primary role in determining student achievement.
- The **"content of instructional materials,"** such as content standards, textbooks, and achievement tests that influence teachers' content decisions

Porter's methodology refines three types of research-based instruments for measuring content and alignment:

- **Surveys** of teachers on the content of their instruction
- **Content analyses** of instructional materials
- **Alignment indices** describing the degree of overlap in content between, for example, standards and assessments

The "central idea" behind these instruments is the development of a **uniform language** for describing the content of instruction, which makes it possible to build useful indices of alignment.

The uniform language consists of descriptors of topics covered in a particular subject and categories of cognitive demand that distinguish what students are expected to know or be able to do (such as Webb's depth-of-knowledge levels, described in Appendix A). Porter uses five categories of cognitive demand (memorize, perform procedures, communicate understanding, solve nonroutine problems, and conjecture/generalize/prove).

The uniform language describes the degree of overlap in content between, for example, standards and assessments, or standards and instructional materials. Table 5 shows a two-dimensional mathematics **content matrix**, akin to the instruments used in the NAEP alignment study, which uses a uniform language to describe mathematics content. (SEC subsequently modified “communicate understanding” to “demonstrate understanding,” and switched the order of “solve nonroutine problems” and conjecture/generalize/prove,” among other revisions.)

The SEC model is distinguished by the use of an external content matrix and descriptive language to which standards and assessments are coded (a third reference point) instead of directly comparing standards to assessments. The same content matrix and language can be used to collect data and analyze instructional content as reported by teachers.

**Table 5. Mathematics Content Matrix**

Topic	Category of cognitive demand				
	Memorize	Perform procedures	Communicate understanding	Solve nonroutine problems	Conjecture/generalize/prove
Multiple-step equations					
Inequalities					
Linear equations					
Lines/slope and intercept					
Operations on polynomials					
Quadratic equations					

Source: Porter, A. (2002)

### Collecting, Analyzing, and Reporting Content Data

This content matrix is designed for collecting, analyzing, and reporting data on curriculum that has been taught and for comparing curriculum content in relation to standards—which the SEC model equates to the “intended curriculum”—as well as assessments.

The SEC model first collects the data for the content matrix by surveying teachers on the amount of time they devote to each topic (level of coverage) and then, for each topic, the relative emphasis they give to each student expectation (category of cognitive demand). These data are then incorporated into the content matrix as proportions of total instructional time spent on the content of instruction—the intersection of each cell by topic and cognitive demand. Across the cells, the proportions sum to 1 (Porter and Smithson, 2001).

## **A Four-Step Procedure for Conducting NAEP Alignment Studies**

For the NAEP alignment study, CCSSO developed a four-step procedure for collecting, analyzing, reporting, and comparing data:

### **Step 1: Training Educators**

A team of educators in the content area of the assessment (or content standards) to be analyzed is trained using standard training and procedures set by CCSSO and WCER. The teams learn to code assessment items (or content standards) according to the topic and category of cognitive demand.

For the NAEP alignment study, four mathematics specialists participated in the content analysis of 2007 NAEP grade 4 assessment items, and three mathematic specialists (the minimum number of analysts recommended) participated in the analysis of NAEP grade 8 assessment items.

### **Step 2: Collecting Data**

Working alone, each analyst examines each assessment item (or content standard) and matches it to a topic and category of cognitive demand in a corresponding content matrix. In other words, for comparison studies, the analysts work with separate content matrices—one for each assessment (or set of content standards) under review.

For the NAEP alignment study, the mathematics specialists analyzed the complete bank of 2007 NAEP mathematics assessment items for grades 4 and 8. The teams also analyzed the state standards and assessments.

### **Step 3: Synthesizing Data**

The data collection results in a set of data codes from each analyst, which are then synthesized and averaged across each team of analysts. This step produces a completed content matrix that represents the team's analysis and coding of the entire set of assessment items (or content standards).

For the NAEP alignment study, the data codes of the individual analysts were reviewed and analyzed to measure reliability, then averaged for each assessment item (or standard) across analysts.

### **Step 4: Analyzing, Comparing, and Reporting Content Data**

The completed content matrices (e.g., assessment to assessment or assessment to standard) are mapped to one another, using SEC alignment indices (Porter, 2002). This mapping results in a set of alignment statistics that report the degree of consistency or similarity in topics and categories of cognitive demand between the assessments (or content standards) that have been analyzed.

An alignment statistic ranges between 0 and 1. In a hypothetical world, 0 would

represent no alignment and 1 would represent perfect alignment. In reality, alignment statistics fall somewhere in between. Determining low vs. high alignment is a judgment based on the purpose of the alignment study and the targets of comparison. For example, assessment-to-standards alignment results generally are lower than assessment-to-assessment results.

For the NAEP alignment study, the completed NAEP content matrix was mapped to the completed content matrices of state standards and assessments. Alignment was examined by comparing both NAEP and state content to the SEC content matrices. State assessments vary from about .25 to .60 in alignment to NAEP. In addition, results of previous analyses of the 2005 NAEP frameworks are included for comparison purposes.

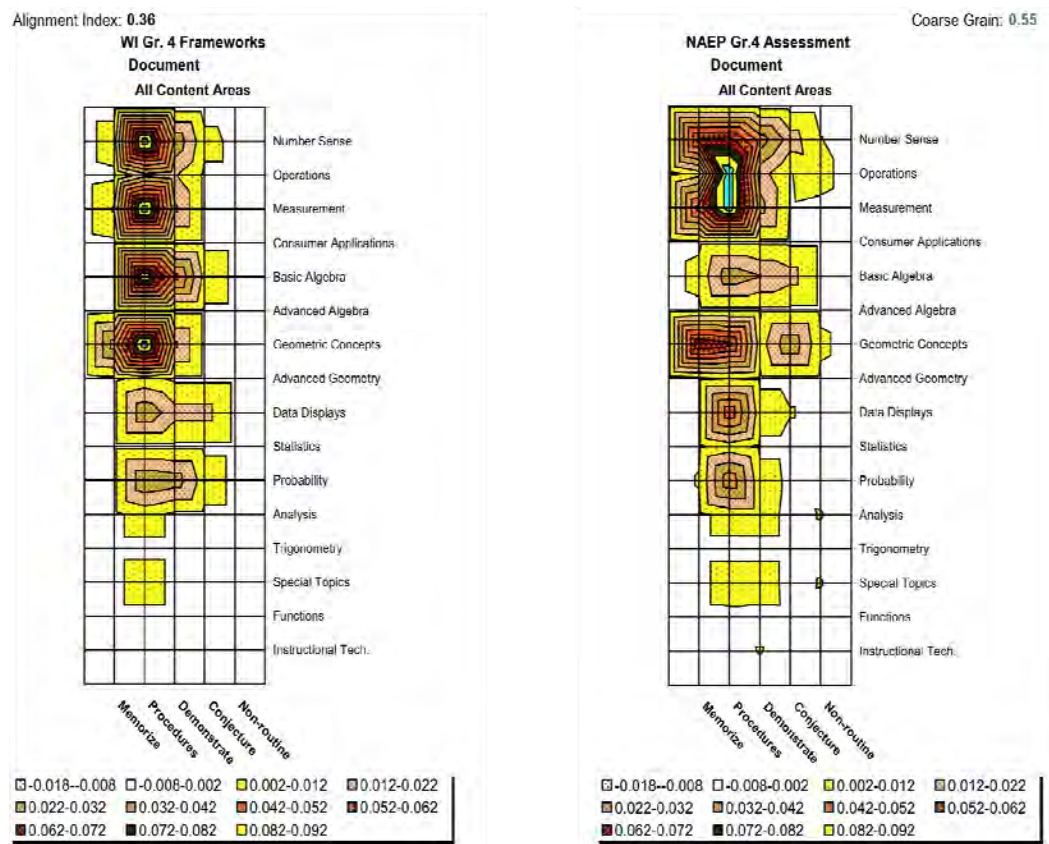
The result of this comparison was an online, interactive viewer (using Excel®) that allows users to examine the findings in several ways, including graphical content display maps, tile charts, and alignment tables.

Exhibit 13 shows a content map that provides a visual comparison of the degree of alignment between Wisconsin’s grade 4 mathematics framework (curriculum standards) and the NAEP grade 4 mathematics assessment. On this content map, “alignment index” refers to the degree of consistency or match between the content (by topic and category of cognitive demand) of the state framework and the NAEP assessment.

The alignment of .36 is considered high. The Wisconsin framework has strong consistency with NAEP topics, but the state heavily emphasizes procedural learning in all topics. NAEP has items coded for all five categories of expectations (memorize, perform procedures, demonstrate understanding, conjecture/generalize/prove and solve nonroutine problems).

The content map in this exhibit includes a “coarse-grain” statistic, which refers to the alignment or consistency of the main topics and expectations between the state and NAEP.

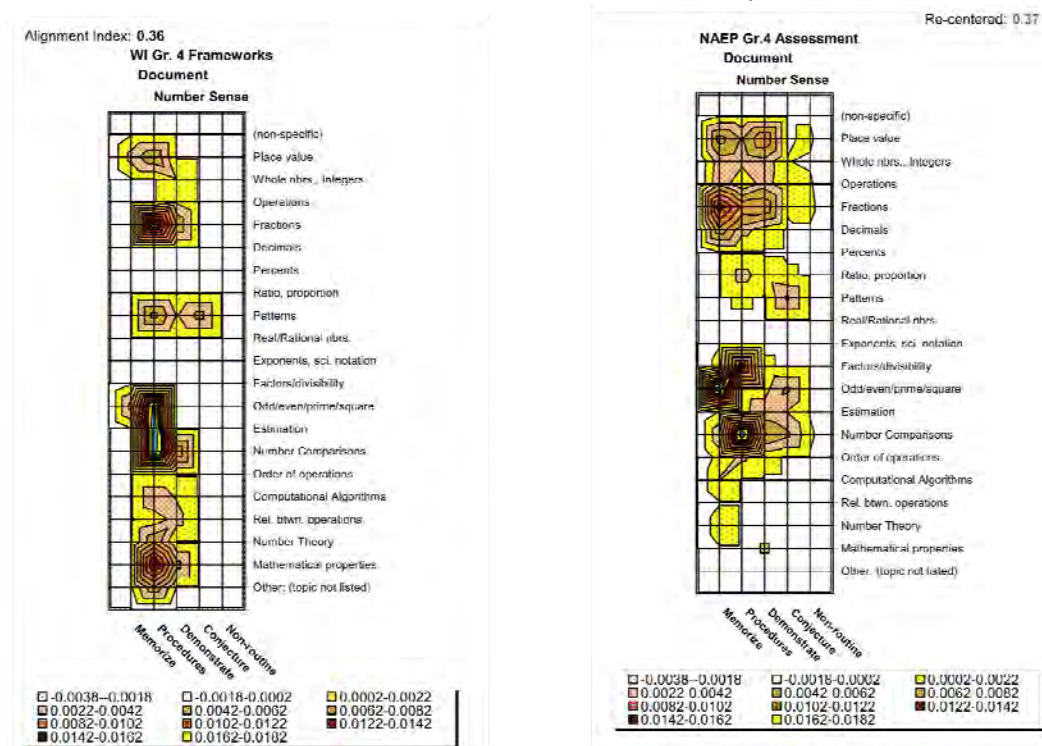
**Exhibit 13. A “Coarse-Grain” Content Map Illustrating Degree of Alignment Between One State’s Standards and NAEP: Grade 4 Mathematics**



Source: CCSSO

Exhibit 14 shows a “fine-grain” comparison of the alignment of one topic, Number Sense, on the state framework and NAEP. The “re-centered” statistics (under the content map) refer to the alignment. For this topic, NAEP includes items in the “memorize” category, while the state framework does not. NAEP also includes items covering operations, decimals, ratios/proportion, and factors/divisibility; the state framework does not cover these topics. The state framework places stronger emphasis than NAEP on mathematical properties.

**Exhibit 14. A “Fine-Grain” Content Map Illustrating Degree of Alignment on One State’s Standards and NAEP: Grade 4, Number Sense**

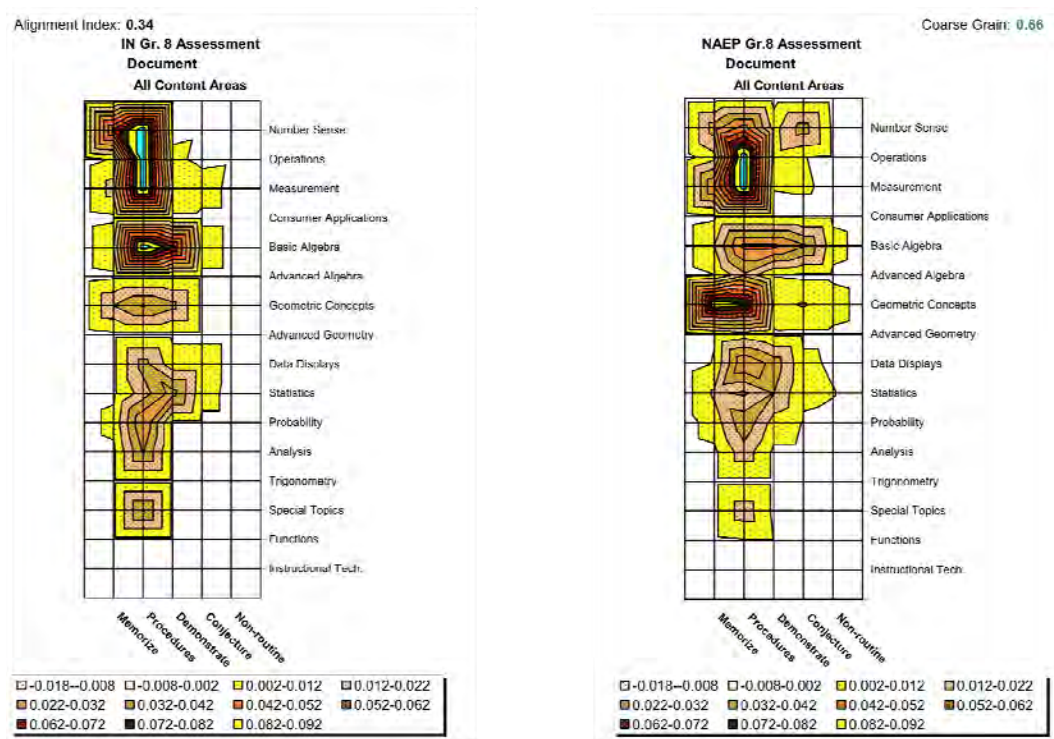


Source: CCSSO

Exhibit 15 illustrates the alignment between Indiana’s grade 8 mathematics assessment and the NAEP grade 8 mathematics assessment. The alignment is high at .34, and there is high consistency between the main math topics covered on both assessments. The NAEP assessment includes items in all five categories of cognitive demand, while the Indiana assessment has no items at the nonroutine level and a small number of conjecture/generalize/prove items. The Indiana assessment places greater emphasis than NAEP on Number Sense.



**Exhibit 15. A “Coarse-Grain” Content Map Illustrating Degree of Alignment Between One State’s Assessment and NAEP: Grade 8 Mathematics**

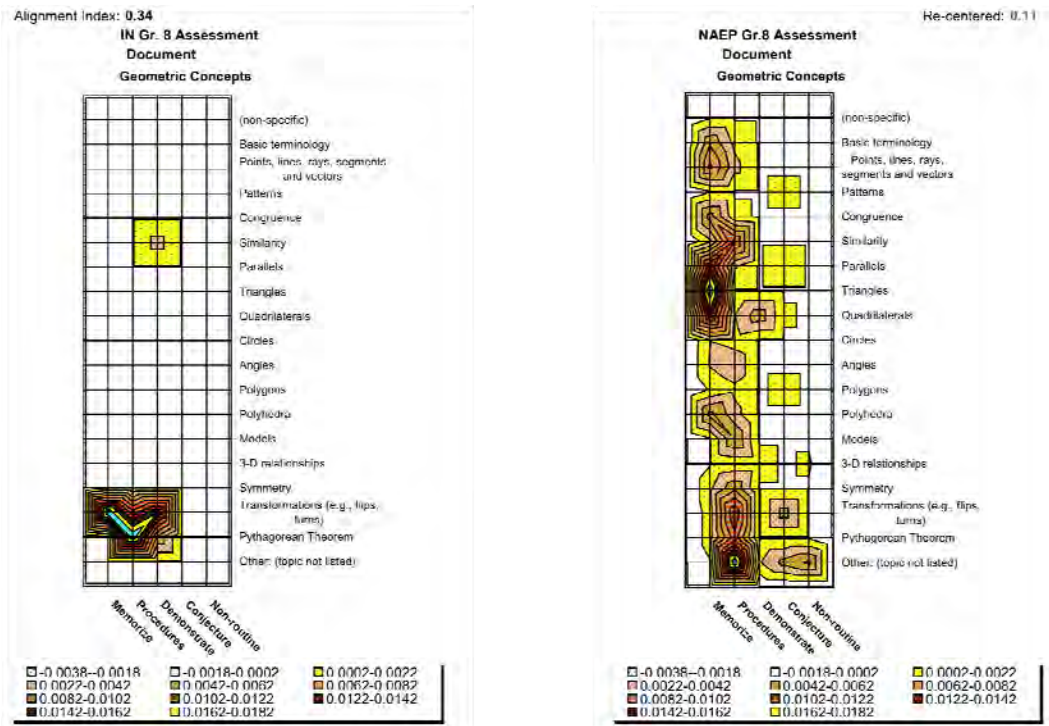


Source: CCSSO

Exhibit 16 shows a “fine-grain” comparison of the alignment of one topic, Geometric Concepts, on the state assessment and NAEP. The Indiana assessment emphasizes a small number of topics in geometry, which the NAEP assessment (which includes more total math items) emphasizes a broader range of topics. The “re-centered” alignment statistic for this topic is only .11, indicating that Indiana does not include much of the geometry content that NAEP assesses at grade 8.



**Exhibit 16. A “Fine-Grain” Content Map Illustrating Degree of Alignment on One State’s Assessment and NAEP: Grade 8, Geometric Concepts**



Source: CCSSO

All of the SEC content analyses for state, NAEP, and international standards and assessments are available at the SEC Web site.

**Learn More**

To learn more about the SEC model, visit <http://www.seconline.org/>

## Conclusion

This report provides states and districts with descriptions and features of three approaches supported by NCEs for planning and conducting alignment studies. The report also provides a synopsis of the surging interest in aligning state assessments and standards with NAEP, an overview of the rationale and purposes for alignment studies, and a glossary for understanding terminology (which often differs from NAEP to states and other organizations).

For state or district officials who have been asked whether their assessments or standards align with NAEP, the NAEP ESSI Procedural Manual could be used to answer the question, “What do I do next?” The NAEP Procedural Manual, the HumRRO Model, or the SEC Model could prove useful in moving forward with an alignment study, depending on the focus of inquiry. Each offers specific tools and resources, subject-area and grade-level coverage, and results. Other models are available as well.

However, alignment initiatives require more than a set of procedures. Invariably, alignment studies have political and educational contexts and consequences, which should be anticipated. State and district officials, educators, the media, and the public might expect a simple “yes” or “no” answer to the question, “Are we aligned with NAEP?” A “yes” or “no” answer does not suffice. Instead, the answer to this question depends on the subject and grade level; the region, district, or school; the students or groups of students; and so on.

Likewise, alignment studies might not produce simple answers. Rather, results might express degrees of alignment, or degrees of overlap or gaps in topic coverage, cognitive demand, or grade-level expectations. This could be extraordinarily valuable information—if states or districts have the political will to use it, for example, to examine and improve their standards, assessments, curriculum, instruction, or professional development.

In the end, the most compelling reason to conduct an alignment study might not be to answer simple questions, but to check the rigor of educational systems. Whether the comparison is to NAEP, to international assessments, or to other state assessments or standards, alignment initiatives carried out with this purpose in mind could yield the deepest insights.

## Features of the Three Alignment Approaches

	NAEP ESSI	HumRRO	Surveys of Enacted Curriculum
<b>What is the focus of comparison?</b>	NAEP frameworks, specifications, and assessment items to state frameworks, assessments, and items	NAEP content expectations, assessment administration, and items to state standards, assessment administration, and items	NAEP assessment items to state standards and assessment items  (This model also can be used to analyze the content and alignment of curriculum and instruction to standards and assessments)
<b>What subjects are covered?</b>	Any subject	Mathematics Reading	English language arts Mathematics Science
<b>Which grades are covered?</b>	Any grade level	Grades 4 and 8	Grades 4 and 8 for NAEP alignment studies; any grade level for curriculum alignment studies
<b>What tools or resources are provided?</b>	<p>Plan for Comparison:</p> <ul style="list-style-type: none"> <li>• Key Questions</li> <li>• Considerations</li> </ul> <p>Procedures:</p> <ul style="list-style-type: none"> <li>• Comparing NAEP and state frameworks</li> <li>• Cross-classifying assessment items to frameworks</li> <li>• Comparing attributes of NAEP and state assessment items</li> </ul>	<p>Background information and workshop facilitator's guide</p> <p>Alignment guide with procedures for eight alignment tasks</p> <p>Directions and worksheets for completing and documenting tasks</p> <p>Reference material on NAEP content expectations and released items</p> <p>Report shell for preparing text and graphics for final report</p>	<p>For NAEP alignment studies:</p> <ul style="list-style-type: none"> <li>• Content matrices for collecting, analyzing, and reporting data on topic coverage and categories of cognitive demand</li> <li>• Coded 2005 and 2007 NAEP mathematics assessment items</li> </ul> <p>For curriculum studies:</p> <ul style="list-style-type: none"> <li>• Surveys for teachers to describe the content of instruction</li> <li>• Content analysis tools and procedures for coding instructional materials</li> <li>• Alignment indices describing the degree of overlap in content between, for example, standards and assessments</li> </ul>
	NAEP ESSI	HumRRO	Surveys of Enacted Curriculum

<b>What products would result from the model?</b>	<p>Comparison maps or grids</p> <p>Quantitative and qualitative summaries of comparisons, cross-classifications, and attributes</p>	Final report with background methodology, results, discussion, recommendations, and bar graphs showing comparisons	<p>A Web-based interactive viewer for examining results in several ways, including content maps, marginal charts, and alignment tables</p> <p>Full study report (optional)</p>
<b>What is the format for conducting an alignment study?</b>	Flexible format determined by state or district	Workshop facilitated by state officials	Team collaboration and online data entry
<b>How much time would an alignment study take?</b>	<p>Preparation time for planning, assigning participants, gathering materials, and logistics, plus:</p> <p>Depends on depth and breadth of study:</p> <p>1 to 2 days for a high-level (coarse-grain) study, more for an in-depth (fine-grain) study of frameworks, test specifications, and cross-matching items to frameworks</p>	<p>Preparation time for planning, assigning participants, gathering materials, and logistics, plus:</p> <p>3.5 days per subject for a workshop to complete Tasks 1–5</p> <p>2 days for one state official to complete Tasks 6–8</p> <p>1 to 3 weeks to prepare a formal report, if required</p>	<p>Preparation time for planning, assigning participants, gathering materials, and logistics, plus:</p> <p>About 2 hours per subject per grade for content analysis, then about half an hour for online data entry</p> <p>About 10 minutes to access NAEP-to-state alignment statistics and graphic comparisons on the SEC Web site</p> <p>Production time for a full study report varies, depending on needs</p>
<b>How many participants are required?</b>	<p>1 person for a high-level study</p> <p>5 to 10 participants per grade level for an in-depth study to inform policy, curriculum decisions, assessment development, or explanation of performance</p>	<p>1 state administrator</p> <p>1 facilitator per workshop</p> <p>5 reviewers per workshop</p>	A team of 3 to 5 content specialists or educators per subject and grade

	<b>NAEP ESSI</b>	<b>HumRRO</b>	<b>Surveys of Enacted Curriculum</b>
<b>What qualifications do participants need?</b>	<p>Coders should have some familiarity with curriculum or assessment standards and assessment development</p> <p>It is good practice to include a mix of classroom teachers, curriculum specialists, and assessment item writers or developers.</p> <p>All participants should become familiar with the alignment materials and item pools</p>	<p>State administrator should be very familiar with NAEP and state tests and have access to state operational test items</p> <p>Facilitator(s) can be a state education official or a teacher and should be familiar with NAEP and the state testing system</p> <p>1 reviewer should be a state-level, K–12 curriculum specialist. Other reviewers should be current or recently retired teachers, with a mix of grade 4 and grade 8 experience</p>	<p>Degree in the subject and experience in teaching or supervision</p>
<b>Are there target levels of agreement among content specialists for coding frameworks, standards, or items?</b>	<p>Target agreement levels depend on whether consensus agreement or diversity of opinion is the goal, which is tied to the results required. For example, if the goal of the study is to determine if there is alignment to a subscale such as algebra, but it does not matter if raters agree on specific objectives an item is coded to, then the degree of match should be very high. The more fine-grained the coding, the tendency is to lower the standard of agreement. One study used the standard “more than half of the coders have to agree” while another study used “at least 6 of the 8 coders had to agree.” The key is to set the target of agreement during the planning stage and be sure the target is reasonable for the purpose.</p>	<p>Agreement is reached by consensus</p>	<p>No. Team members do not need to agree. The aim is to get an honest, objective analysis from each member without negotiation. The coded values are then averaged.</p>

## References

- Blank, R.K., with assistance from the Wisconsin Center for Education Research, University of Wisconsin–Madison. (November 2007). *Alignment Content Analysis of NAEP 2007 Mathematics Assessment Using the Surveys of Enacted Curriculum Methodology*. Washington, DC: Council of Chief State School Officers.
- Blank, R.K., Porter, A., and Smithson, S. (2001). *New Tools for Analyzing Teaching, Curriculum, and Standards in Mathematics and Science*. Report form Survey of Enacted Curriculum Project. (National Science Foundation REC98-03080). Washington, DC: Council of Chief State School Officers.
- Gattis, K. (in press). *Comparing National Assessment of Educational Progress Frameworks, Specifications, and Assessment Items to State Frameworks and Assessments: A Procedural Manual*. Washington, DC: American Institutes for Research.
- Koger, L.E., Deatz, R.C., Lozzi, D.E., and Furr, D. (2006) *Method Template for Describing the Alignment of State and NAEP Assessments*. Alexandria, VA: Human Resources Research Organization.
- Mueller, C., and Gattis, K. (2004). *Degree of Match*. Paper presented at the Three State Comparison Work Group Meeting #2, Washington, DC.
- Porter, A. (October 2002). “Measuring the Content of Instruction: Uses in Research and Practice.” *Educational Researcher*, Vol. 31, No. 7, pp 3–14.
- Porter, A., and Smithson, J. (2001). “Are Content Standards Being Implemented in the Classroom? A Methodology and Some Tentative Answers.” In S.H. Fuhrman (Ed.), *From the Capitol to the Classroom: Standards-Based Reform in the States* (100th Yearbook of the National Society for the Study of Education, part II, pp 60–80). Chicago: National Society for the Study of Education; distributed by University of Chicago Press.
- Smith, M.S., and O’Day, J. (1991). “Systemic school reform.” In S.H. Fuhrman & B. Malen (Eds.), *The Politics of Curriculum and Testing: The 1990 Yearbook of the Politics of Education Association* (pp. 233–267). Bristol, PA: Taylor & Francis.
- Webb, N.L. (1997). *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education*. Council of Chief State School Officers and National Institute for Science Education Research Monograph 6. Madison, WI: University of Wisconsin–Madison, Wisconsin Center for Education Research.
- Webb, N.L. (1999). *Alignment of Science and Mathematics Standards and Assessments in Four States* (Research Monograph No. 18). Madison, WI: National Institute for Science Education.
- Webb, N.L. (2002). *Alignment Study in Language Arts, Mathematics, Science, and Social Studies of State Standards and Assessments for Four States*. Washington, DC: Council of Chief State School Officers.
- Webb, N.L., Alt., M., Ely, R., and Vesperman, B. (2005). *Web Alignment Tool (WAT) Training Manual*. Washington, DC: Council of Chief State School Officers.
- Webb, N.L., and Smithson, J. (April 1999). *Alignment Between Standards and Assessments in Mathematics for Grade 8 in One State*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

## Glossary

**Alignment.** The connections between curriculum, standards and assessments

**Alignment index.** The degree of consistency or match between domains, such as standards and assessments

**Cognitive demand.** Types of thinking or reasoning processes

**Constructs.** The underlying traits (knowledge, skills) an assessment is intended to measure

**Content of instruction.** Decisions that teachers make about what to teach (and how), how much time to spend on a particular subject, what topics to cover, when and in what order, to what standards of achievement, and to which students

**Content of instructional materials.** Content standards, textbooks, and achievement tests that influence teachers' content

**Content map.** A visual presentation of content coverage

**Content matrix.** An instrument for alignment studies that uses a uniform language to describe content

**Coarse-grain comparison.** The alignment or consistency of the main topics and expectations between domains

**Degree of match.** How closely different standards, assessments, or items relate to one another

**Dimensions.** The content of a subject

**Fine-grain comparison.** The alignment or consistency of specific topics and expectations between domains

**Frameworks.** Standards of a subject-area domain or assessment development documents that specify the content to be assessed

**Objective.** The educational goal of an assessment question

**Operational items.** Assessment questions still in use and, typically, secured from the public

***p*-value.** A *p*-value is the proportion of sampled students who responded correctly to a test question. The *p*-value actually represents item "easiness," but traditionally psychometricians refer to it as item difficulty. The NAEP Question Tool states: "Difficulty is a measure of student performance on a question. Multiple-choice or constructed-response questions scored either right or wrong are rated 'easy' if answered correctly by 60 percent or more of students, 'medium' if answered correctly by 40 to 59 percent of students, or 'hard' if answered correctly by fewer than 40 percent of students."

**Released items.** Assessment questions no longer in use and that are available to the public

**Specifications.** Assessment development documents

**Standards.** Curriculum or content expectations

**Taxonomy.** The classification of the categories of thinking or reasoning skills required by items

**Uniform language.** Standard terms, definitions, and understandings for describing content



## **Appendix A**

### **Webb’s “Depth-of-Knowledge” Model for Alignment**

The “depth-of-knowledge” alignment model, developed by Norman Webb of the Wisconsin Center for Education Research, has been influential in aligning standards with assessments (Webb, 1997; Webb, 1999; Webb, 2002; Webb and Smithson, 1999). Webb’s depth-of-knowledge levels describe and show the progression of the rigor, or complexity, of content and expectations.

#### **Level 1: Recall and Reproduction**

Requires recall of information, such as a fact, term, or performance of a simple process or procedure

#### **Level 2: Skills and Concepts**

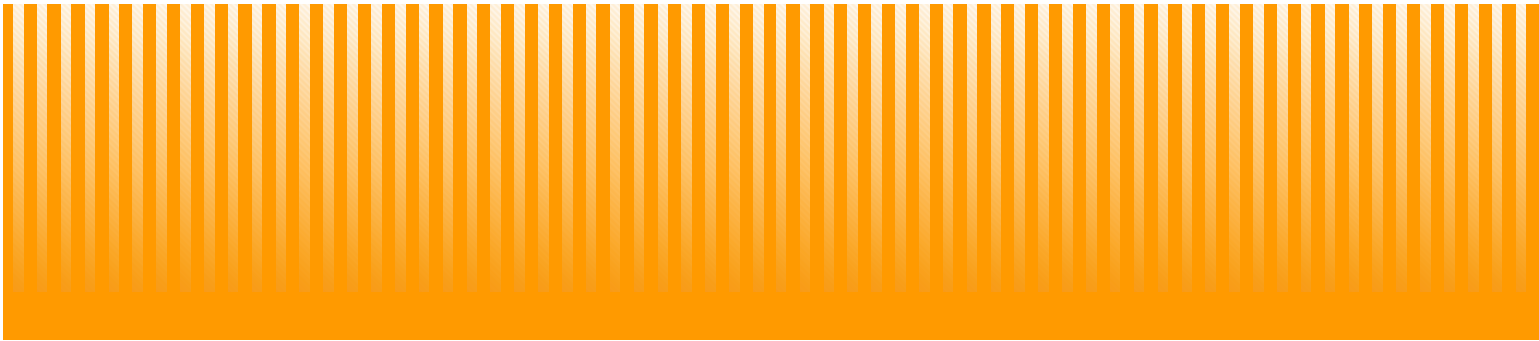
Requires engaging some mental process beyond recalling or reproducing a response, to make some decisions about how to approach a question or problem

#### **Level 3: Strategic Thinking**

Requires deep understanding as exhibited through planning, using evidence, and more demanding, complex, and abstract cognitive reasoning

#### **Level 4: Extended Thinking**

Requires high cognitive demand and is very complex. Students must make connections—related ideas within the content or among content areas—and select or devise one approach among many alternatives on how the situation can be solved. Most assessments do not include Level 4 items, because they probably require an extended period of time to carry out, but they are appropriate for standards, goals, and objectives and for instructional activities.



Council of Chief State School Officers | One Massachusetts Avenue, NW, Suite 700 | Washington, DC 20001-1431  
[www.ccsso.org](http://www.ccsso.org)

