

Assessing the CSAP

How Can We Assess the Quality of a Large-Scale, Standards-Based Assessment?

Vonda L. Kiplinger, Ph.D.
Colorado Department of Education
Student Assessment Unit

Presented to the Seventh Annual Standards and Assessments
Conference
Breckenridge, Colorado

June 18, 1998



Checklist: Properties of A Good Large-Scale Standards-Based Assessment

- ✓ Aligned with content standards
- ✓ Challenging
- ✓ Developmentally appropriate
- ✓ Unbiased
- ✓ Provide access & equal opportunities
- ✓ Administered under standardized conditions
- ✓ Accurate measure of what student **should** know & be able to do
- ✓ Provide accurate estimates of students' abilities
- ✓ Minimize effects of factors unrelated to knowledge/skill tested
- ✓ Reliable
- ✓ Valid

Statistical Properties of the Test

- ✓ Scaling method & properties of the items
- ✓ Consistency of measurements and reliability
 - Comparability
 - Relationship among standards
 - Relationship of standards to the whole test
- ✓ Validity
 - Does the test measure what it is intended to measure?
 - Is the test of appropriate difficulty?
- ✓ Checklist: Properties of a good assessment

Test Scaling

- ✓ Item Response Theory - AKA Item-Pattern Scoring
 - More accurate than no. (%) correct scoring
 - Takes more information into account
 - Number or percent correct
 - Which items correct (i.e., the pattern)
 - Guessing
 - Item characteristics - item difficulty, discrimination
 - Only method that can simultaneously calibrate SR & CR items on same scale
 - Provides more accurate estimates for low- & middle-ability examinees

Consistency of Measurements: Comparability

- ✓ Necessary when there are multiple test forms for a subject area, e.g., CSAP Writing
 - Mean scores
 - Correlation of the 2 writing standards
 - Correlation of each standard with total test score
- ✓ Necessary for year-to-year comparison

Comparability of Writing Forms A, B, & C

- ✓ Raw score means
 - A = 37.69
 - B = 36.51
 - C = 37.71
- ✓ 3 forms equated
 - Expected scale score same regardless of writing form
- ✓ Correlations between standards same across forms

The Writing Standards

- ✓ Standard 2
 - Students write and speak for a variety of purposes and audiences.
- ✓ Standard 3
 - Students write and speak using conventional grammar, usage, sentence structure, punctuation, capitalization, and spelling.

Correlation Among the Writing Standards

- ✓ Correlation of standards with total test
 - Standard 2 & Test: $r = 0.89$
 - Standard 3 & Test: $r = 0.95$
- ✓ Correlation of Standards 2 & 3: $r = 0.69$
- ✓ The 2 Writing standards are separate, but related, aspects of Writing

The Reading Standards

- ✓ Standard 1
 - Students read and understand a variety of materials.
- ✓ Standard 4
 - Students apply thinking skills to their reading....
- ✓ Standard 5
 - Students read to locate, select, and make use of relevant information from a variety of media, reference, and technological sources.
- Standard 6
- Students read and recognize literature as a record of human experience.

Correlations Among the Reading Standards

- ✓ Correlations among standards higher than for Writing
 - 0.74 - 0.81
- ✓ Correlations of standards with total test
 - Standard 1 & Test: $r = 0.94$
 - Standard 4 & Test: $r = 0.92$
 - Standard 5 & Test: $r = 0.89$
 - Standard 6 & Test 4 = 0.91
- ✓ All 4 Reading standards are measuring the same construct: Reading
- ✓ Test is a very good measure of how we are defining Reading in the standards

Reliability \Leftrightarrow Consistency

- ✓ Reliability: Definition
 - Index of consistency of test results
 - Reliability coefficient ranges from 0.00 to 1.00
 - Reliability of most NRTs are ~ 0.85 to 0.92
- ✓ Reliability: CSAP, grade 4
 - Writing: 0.91, all forms
 - Reading: 0.94
- ✓ Reliability of CSAP higher than most multiple choice NRTs

Validity

- ✓ Colloquially: Does the test measure what it's supposed to measure?
- ✓ Each item was selected, revised, or developed to measure a standard/benchmark
- ✓ *Should vs. Is*
 - Designed to measure what students *should* be able to do, not what current students *are* doing
 - Designed to serve as impetus for curricular/ instructional change & alignment to standards

Validity Study

- ✓ Just completed validity workshop with teachers not involved in test development
- ✓ Teachers studied passages, prompts, & items

Purposes of Validity Workshop

- ✓ Evaluate alignment of test items to standards
 - Gave their judgments of which standard/ benchmark each item was measuring
 - Compared teacher judgments with test design
- ✓ Evaluated difficulty of items by standard
- ✓ Determined adequacy of standards coverage
- ✓ Explicit statement of constructs measured by the assessments

Evaluation of Construct Validity

- ✓ Evaluative statement on the construct validity of the assessments
- ✓ “Are the assessments valid measures of 4th grade reading and writing and 3rd grade reading?”

Results: Grade 3 Reading

- ✓ 100 % agreement between teacher assignment and intended alignment
- ✓ Adequate coverage of benchmarks
- ✓ Appropriate range of difficulty

Teacher Statement of Validity: Grade 3 Reading

“We feel it does assess State Standards 1.1, 1.3, and is a valid assessment.... The test does have an appropriate range of difficulty. It was balanced between the strategies, and bounced back and forth between strategies. Range of difficulty was mixed appropriately between easy, moderate and difficult. It can be an effective tool to help teachers align instruction with standards.”

Results: Grade 4 Writing

- ✓ 100 % agreement between teacher assignment and intended alignment
- ✓ Adequate coverage of benchmarks
- ✓ Appropriate range of difficulty

Teacher Statement of Validity: Grade 4 Writing

“We have concluded that CSAP accurately tests the standards set by the state. It addresses all components of standards # 2 and #3.”

Results: Grade 4 Reading

- ✓ Overall, good agreement between teacher assignment and intended alignment (85 %)
 - Standard 1: 96 % agreement
 - Standard 4: 100 % agreement
 - Standard 5: 100 % agreement
 - Standard 6: 36 % agreement
- ✓ Most benchmarks covered, but not all
 - New items can fill in gaps

So What's the Problem with Standard 6?

- ✓ Partly due to way Standard 6 & benchmarks written (see last 2 pages of handout, particularly last page)
- ✓ At 4th grade, most questions related to literature are reading comprehension items, even those related to:
 - “...responding to, and discussing a variety of literature such as folk tales, legends, myths, fiction, ... poems...”
 - “using literary terminology such as setting, plot, character, problem, and solution...”
 - “using new vocabulary from literature in other context.”
- ✓ Still need to reexamine how to design a better assessment of components of literature

Teacher Statement of Validity: Grade 4 Reading

“We have concluded that CSAP addresses the following standards

#1

#4

#5.

We had to stretch to meet standard #6.”

Checklist: Properties of A Good Large-Scale Standards-Based Assessment

- Aligned with content standards
- Challenging
- Developmentally appropriate
- Unbiased
- Provide access & equal opportunities
- Administered under standardized conditions
- Accurate measure of what student **should** know & be able to do
- Provide accurate estimates of students' abilities
- Minimize effects of factors unrelated to knowledge/skill tested
- Reliable
- Valid

Colorado Model Content Standards

Reading & Writing

STANDARD 1:

Students read and understand a variety of materials.

In order to meet this standard, students will

- use comprehension skills such as previewing, predicting, inferring, comparing and contrasting, re-reading and self-monitoring, summarizing, identifying the author's purpose, determining the main idea, and applying knowledge of foreshadowing, metaphor, simile, symbolism, and other figures of speech;
- make connections between their reading and what they already know, and identify what they need to know about a topic before reading about it;
- adjust reading strategies for different purposes such as reading carefully, idea by idea; skimming and scanning; fitting materials into an organizational pattern, such as reading a novel chronologically; finding information to support particular ideas; and finding the sequence of steps in a technical publication;
- use word recognition skills and resources such as phonics, context clues, picture clues, word origins, and word order clues; reference guides; roots, prefixes, and suffixes of words for comprehension; and
- use information from their reading to increase vocabulary and enhance language usage.

RATIONALE:

The goal for students at all levels is that they know and can use strategies--various ways of unlocking the meaning of words and larger blocks of text--to become successful readers. The strategies are applied in increasingly difficult reading material at each grade level. At all levels, students should be challenged to read literature and other materials that stimulate their interests and intellectual abilities. Reading from a wide variety of texts, both assigned and student selected, provides experience in gaining information and pleasure from diverse forms and perspectives.

GRADES K-4

In grades K-4, what the students know and are able to do includes

- using a full range of strategies to comprehend materials such as directions, nonfiction material, rhymes and poems, and stories.

GRADES 5-8

As students in grades 5-8 extend their knowledge, what they know and are able to do includes

- using a full range of strategies to comprehend technical writing, newspapers, magazines, poetry, short stories, plays, and novels in addition to the types of reading material mentioned above. Students extend their thinking and understanding as they read stories about people from similar and different backgrounds.

GRADES 9-12

As students in grades 9-12 extend their knowledge, what they know and are able to do includes

- using a full range of strategies to comprehend essays, speeches, autobiographies, and first-person historical documents in addition to the types of literature mentioned above.

For students extending their English/Language Arts education beyond the standards, what they know and are able to do **may** include

- using a full range of strategies to comprehend literary criticism and literary analysis, professional and technical journals, and professional-level reading materials that match their career or academic interests.

STANDARD 6: Students read and recognize literature as a record of human experience.

In order to meet this standard, students will

- know and use literary terminology;
- read literature to investigate common issues and interests;
- read literature to understand places, people, events, and vocabulary, both familiar and unfamiliar;
- read literature that reflects the uniqueness and integrity of the American experience;
- read classic and contemporary literature, representing various cultural and ethnic traditions from throughout the world; and
- read classic and contemporary literature of the United States about the experiences and traditions of diverse ethnic groups.

RATIONALE:

Literature records human expression in such forms as speeches, poems, novels, stories, nonfiction, essays, plays, films, biographies, and autobiographies by male and female speakers and writers. The study of literary traditions offers a perspective on enduring questions, a glimpse into human motives and conflicts, and a sense of the beauty and power of spoken and written language. In addition, literature transmits and transforms culture; it also enables students to think, communicate, and participate in society. The study of literature and writers of the United States honors the heritage and cultures of all people who live or have lived in America, and it thus helps students develop an understanding of our national experience. A comprehensive literature program fosters habits of reading that carry over into adult life.

GRADES K-4

In grades K-4, what students know and are able to do includes

- reading, responding to, and discussing a variety of literature such as folk tales, legends, myths, fiction, rhymes and poems, non-fiction, and content-area reading;
- reading, responding to, and discussing literature as a way to explore the similarities and differences among stories and the ways in which those stories reflect the ethnic background of the author and the culture in which they were written;
- recognizing the concept of classic or enduring literature, and reading and listening to classic works;
- using literary terminology such as setting, plot, character, problem, and solution; and
- using new vocabulary from literature in other context.

GRADES 5-8

As students in grades 5-8 extend their knowledge, what they know and are able to do includes

- reading, responding to, and discussing a variety of novels, poetry, short stories, non-fiction, content-area and technical material, and plays;
- reading, responding to, and discussing literature that represents points of view from places, people, and events that are familiar and unfamiliar;
- distinguishing the elements that characterize and define a literary "classic";
- comparing the diverse voices of our national experience as they read a variety of United States literature;
- using literary terminology accurately, including setting, character, conflict, plot, resolution, theme, foreshadowing, and figurative language; and
- using new vocabulary from literature in other context.

GRADES 9-12

As students in grades 9-12 extend their knowledge, what they know and are able to do includes

- reading, responding to, and discussing novels, poetry, short stories, non-fiction, content-area and technical material, plays, essays, and speeches;
- using literary terminology accurately, such as theme, mood, diction, idiom, perspective, style, and point of view;
- identifying recurrent themes in United States literature; and
- developing and supporting a thesis about the craft and significance of particular works of literature, both classic and contemporary, from a variety of ethnic writers.

For students extending their English/Language Arts education beyond the standards, what they know and are able to do **may** include

- comparing and contrasting stories, novels, poems, and other forms of literature from different countries, time periods, or cultures;
- using novels, poetry, short stories, non-fiction, autobiographies, plays, essays, speeches, literary criticisms and analyses, and any other literature to explore academic, personal, or career issues; and
- understanding the common themes in the literature of the United States and in world literature.