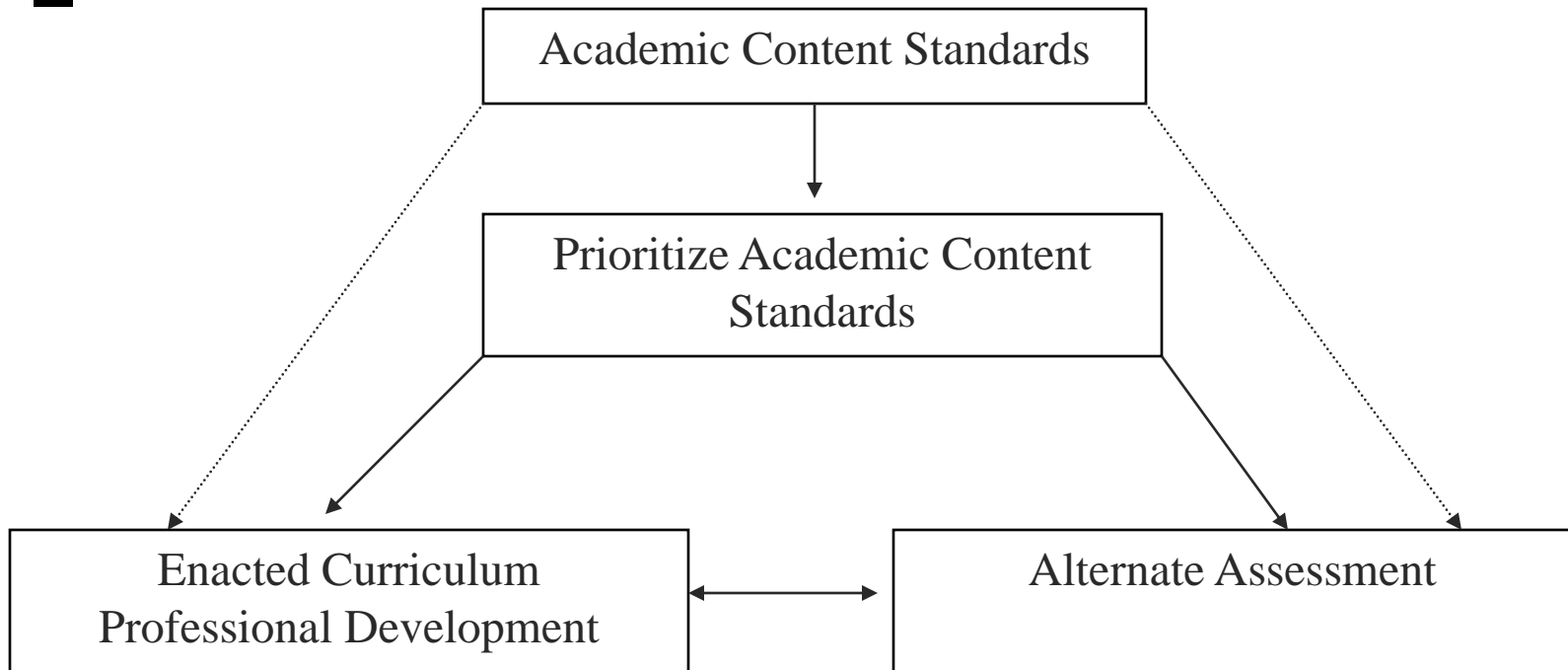


Links for Academic Learning: Measurement of Criteria

Claudia Flowers

Educational Components



[Alignment Indices]

- Content analysis
- Expert review
 - Content experts
 - Special educators
- Descriptive and alignment statistics
 - Norman Webb's work
 - Achieve Model of alignment
- Holistic judgment

Interrater Reliability

- More than one rater needed
- Training
 - Raters need to know the student population, even the content experts
 - Practice rating and capture rules
- Plan for double coding ~ 20%
 - Check agreement during the process

[Sampling Considerations]

- Portfolios
 - Teachers decide content
- Pool of Items/Task
 - Teachers select
- Checklist
 - Supporting evidence

[Criterion 1]

- The content is academic and includes the major domains/strands of the content area as reflected in state and national standards (e.g., ELA, math, science).
- **What is rated:** Extended standards (if applicable) & AA items/tasks
- **Rated by:** Content experts (Special educators may review nonacademic items)

Definition of Academic

- National standards are defined according to the national content organizations
 - National Council of Teachers of English,
 - National Council of Teachers of Math,
 - National Research Council

[Example of Non-Academic]

Turn the page of a book



Non-academic

[Example of Academic]

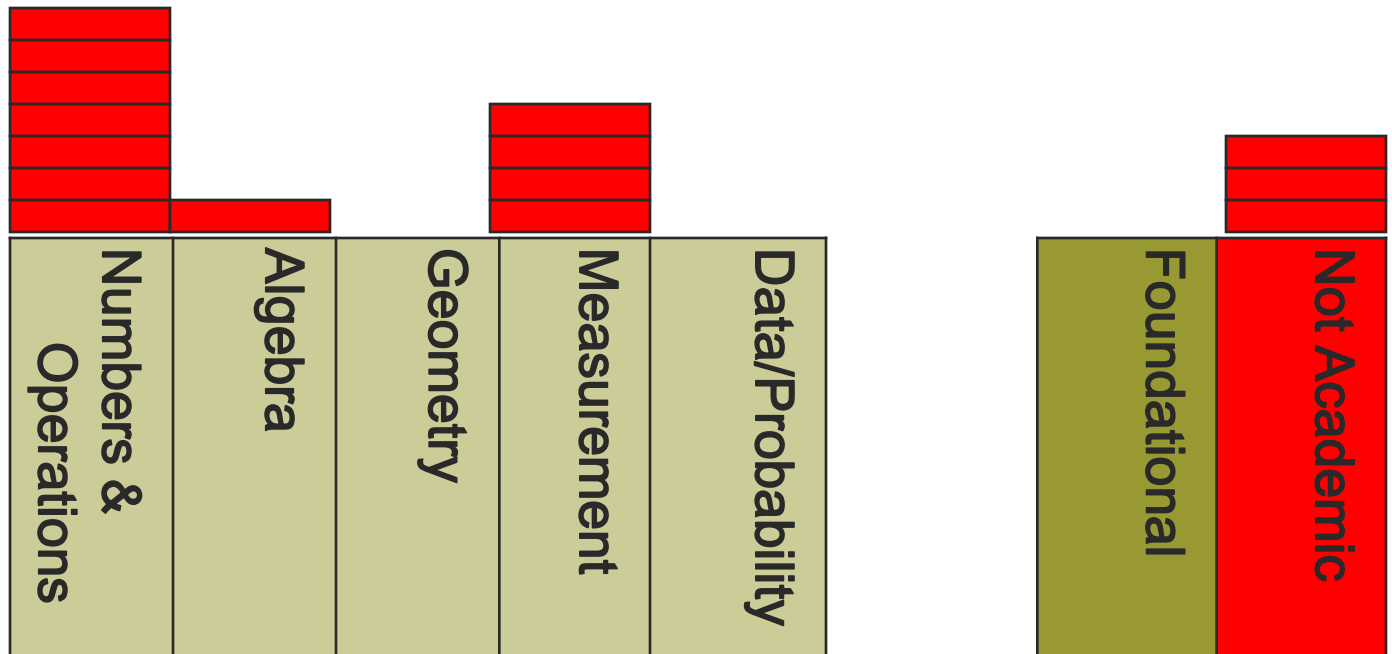
Count one-to-one correspondence



Academic

Criterion 1

15 Items/Task



NCTM Strands

September 25, 2007

Primary & Secondary Hits

Reading	Writing	Communication	Research

[Not Academic]

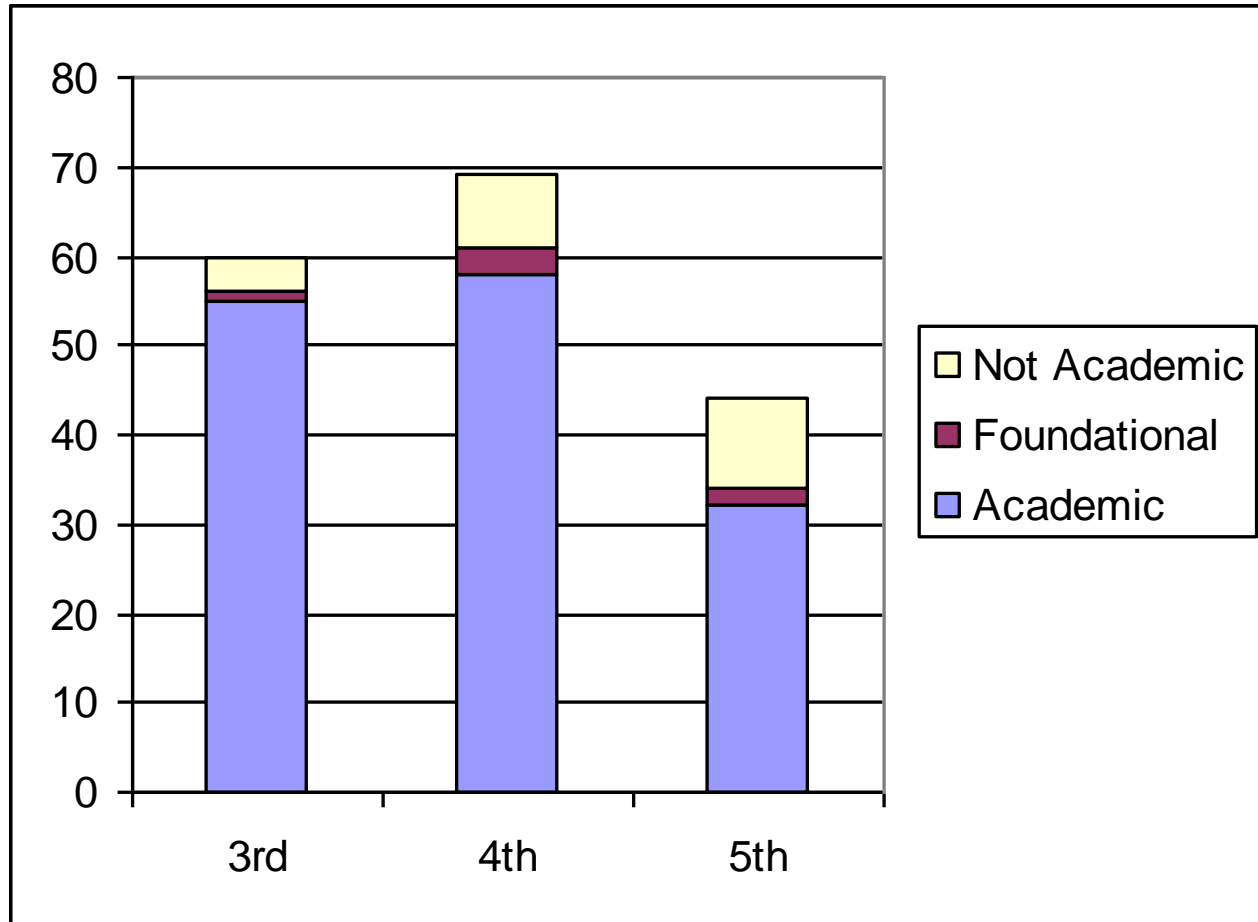
- Extended standards and assessment items that are not academic are not included in any further analyses EXCEPT:
 - Foundational
 - Symbolic level of item

Results Reported

	3 rd		4 th		5 th	
	<u>N</u>	<u>%</u>	<u>N</u>	<u>%</u>	<u>N</u>	<u>%</u>
Not Academic	4	6.7	8	11.6	10	22.7
Foundational	1	1.7	3	4.3	2	4.5
Academic	55	91.7	58	84.1	32	72.7
Academic CI (95%)						
Lower CI		85.0		76.9		62.5
Upper CI		95.8		89.4		81.1

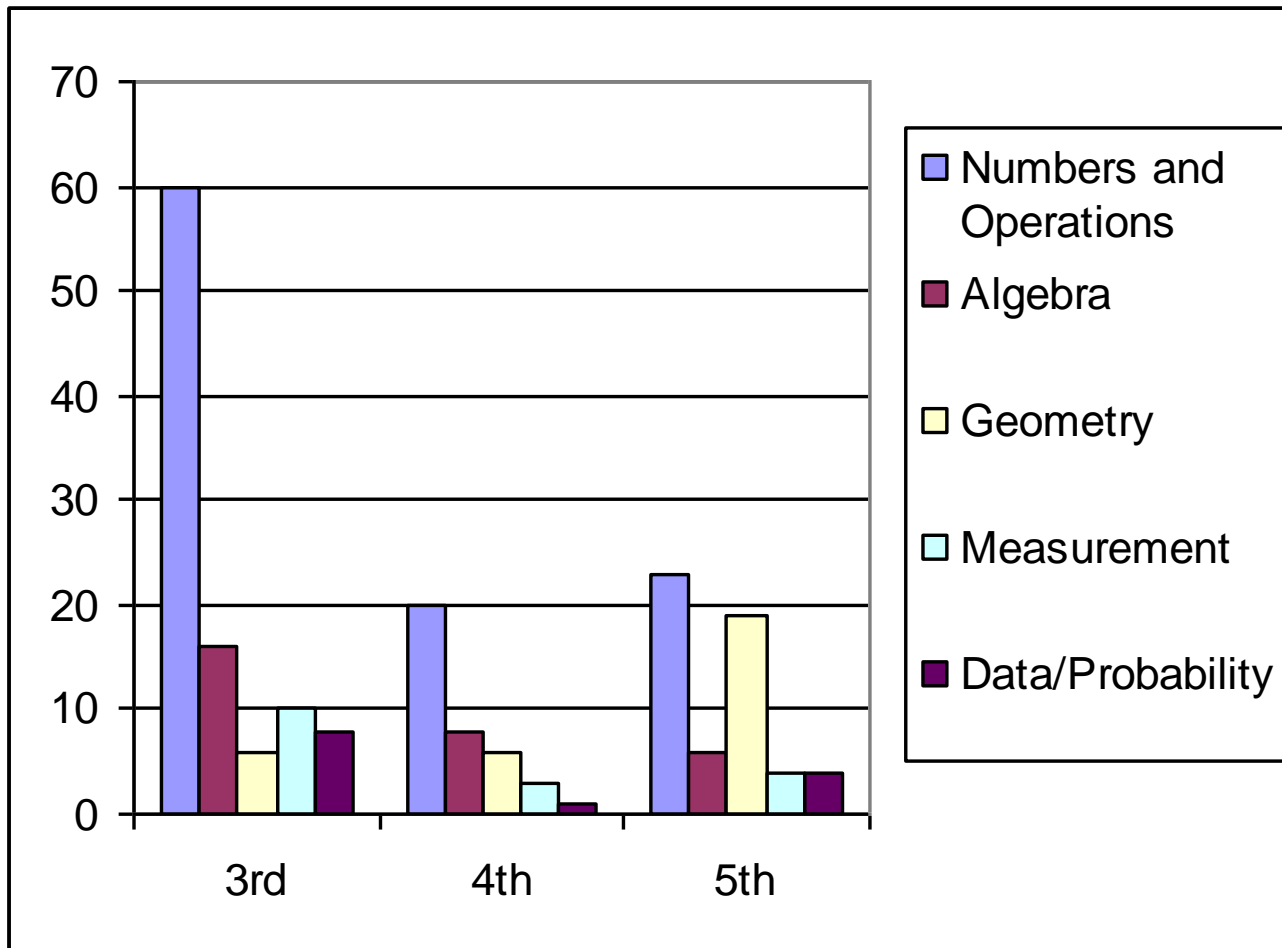
Note. These results are based on a sample of portfolios

[Graphic]



September 25, 2007

Items by Strand



[Evaluative Judgment]

- Most items/tasks should be rated academic
- Items rated not academic should be foundational or accessible for students with presymbolic or concrete levels of communication

[Special Considerations]

- Nonacademic items including foundational are eliminated from all further analysis.
- When statistics are based on a sample of student evidence or AA items, confidence intervals should be provided.
- A list of standards and/or assessment items that were rated as nonacademic should be provided to the state.

Criterion 2

- The content is referenced to the student's assigned grade level (based on chronological age).
- **What is rated:** Extended standards (if applicable) & AA items
- **Rated by:** Alignment Leader

[Intended Links]

- Summarize the link between the components (extended standards and AA items) and grade level content standards
- Illustrates the state's intended emphasis on grade level content standards
- Outcome of this step in the alignment is essentially a reverse engineered test blueprint that includes the frequency distribution of AA items/extended standards by strand and grade.

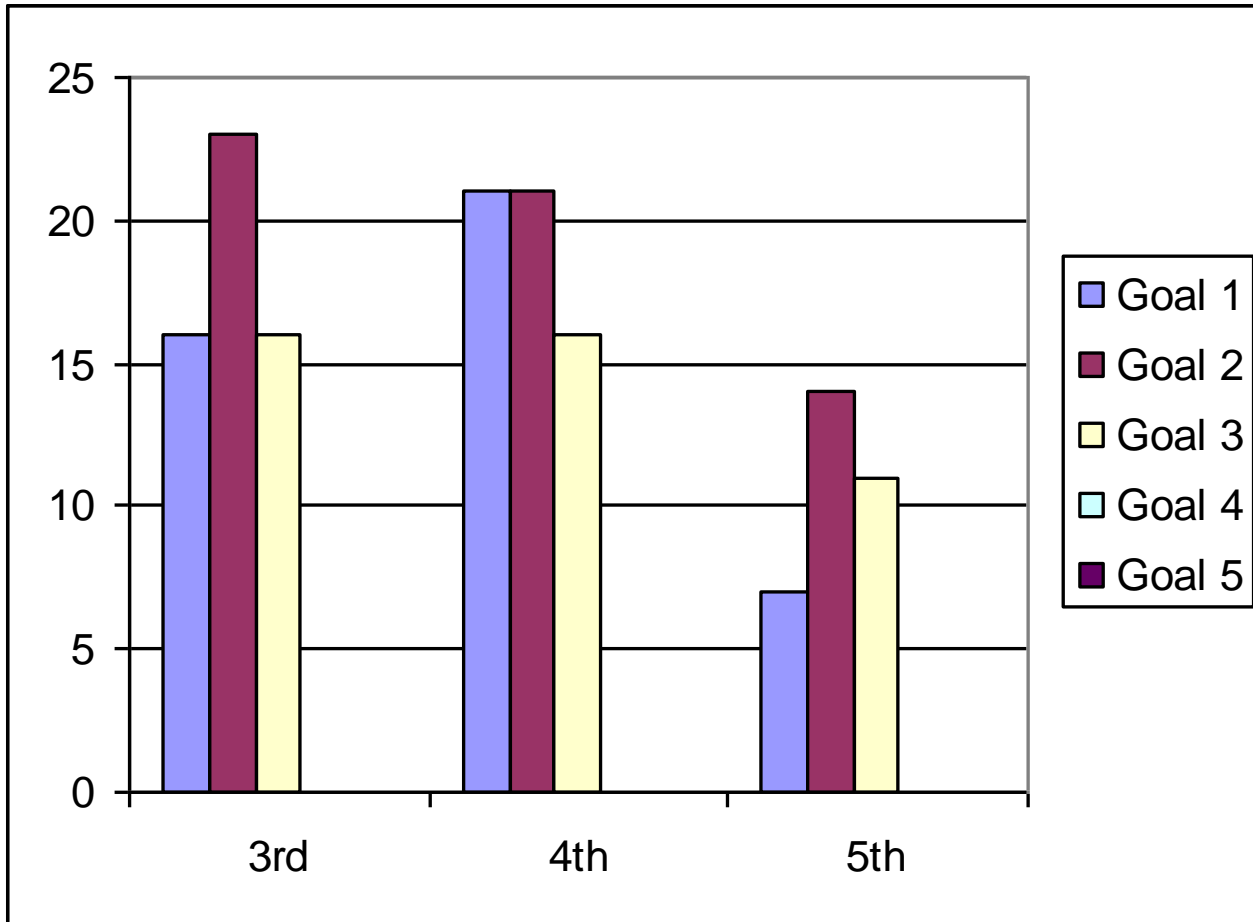
[Results Reported]

- The number and percentage of extended standards and AA items referenced to grade level content standards, by grade.
- When substantial gaps are noted in the strands and grade levels to which the components are intentionally organized, narrative statements about those gaps are included in the report.

Results

Content Goal (State's ELA Content Standards)	3rd		4 th		5 th	
	<u>N</u>	<u>%</u>	<u>N</u>	<u>%</u>	<u>N</u>	<u>%</u>
1. The learner will develop and apply enabling strategies and skills to read and write.	16	29.1	21	36.2	7	21.9
2. The learner will develop and apply strategies and skills to comprehend text that is read, heard, and viewed.	23	41.8	21	36.2	14	43.8
3. The learner will make connections through the use of oral language, written language, and media and technology.	16	29.1	16	27.6	11	34.4
4. The learner will apply strategies and skills to create oral, written and visual texts.	0	.0	0	.0	0	.0
5. The learner will apply grammar and language conventions to communicate effectively.	0	.0	0	.0	0	.0

[Graphic]



September 25, 2007

[Evaluative Judgment]

- Gaps are noted in the content standards and grade levels to which the components are intentionally organized, narrative statements about those gaps are included in the report.
- Recommendations about the gaps may also be made and the rationale that the state provides for the gaps should be considered.

Criterion 3

- The focus of achievement maintains fidelity with the content of the original grade level standards (content centrality) and when possible, the specified performance (performance centrality).
- **What is rated:** Extended standards (if applicable), AA items, supporting evidence
- **Rated by:** Content experts (special education experts follow up on selected items)

Content and Performance Centrality

- Content experts rate content centrality on a three-point scale measuring the degree of “link” (no, far, near)
- Performance centrality is the level of expected performance rated on a 3-point scale (none, some, all)
 - Based on the work of the Achieve Model ((Rothman, Slattery, Vranek, & Resnick, 2002).

[Example-Content Centrality]

Extended Standard:

Identify common sight words

AA Item:



Identify the name of another student
out a choice of three names

Near

[Example-Content Centrality]

Extended Standard:

Observe and investigate patterns of
weather over time using tools

AA Item:



Identify use of weather tools

Far

[Example-Content Centrality]

Extended Standard:

Demonstrate knowledge of microorganisms
in human disease

AA Item:



Follow the steps of washing hands

No Link

[No Content Centrality Link]

- Content experts review the “no link” items/tasks and provide a rationale
 - Overstretched
 - Back fitted skill
 - Mismatch to the standard
- Items rated as “no link” (no content centrality) are omitted from further analysis.

[Performance Centrality]

Compare different types of sounds based on characteristics such as pitch and volume



Identify and compare loud and soft sounds

Some

[Performance Centrality]

Describe how clouds are formed



Made a cumulus cloud using cotton balls

None

Results Reported

- The distribution of ratings of extended standards based on content and performance centrality
- The distribution of ratings of AA items based on content and performance centrality

Results

Content Centrality	3rd	4th	5th
No Link	10.9	10.3	0.0
Far Link	14.5	22.4	15.6
Near Link	74.5	67.2	84.4
CI (95%) for No Link			
Lower CI	5.7	5.0	NA
Upper CI	16.7	14.7	NA
N	55	58	32

[Evaluative Judgment]

- Most items should be rated as having *far* or *near* links to the standards for content centrality
- Most items should be rated as having *some* or *all* of the performance level found in the standards

Special Considerations

- AA item centrality ratings are made directly to grade level content standards if the state does not have extended standards
- If ratings were based on a random sample of items or evidence, confidence intervals should be reported for the “none” and “no link” items.
- Items rated with “no link” (no content centrality) are omitted from analysis in subsequent criteria, but are content analyzed to determine the type of problem that led to the rating.

[Criterion 4]

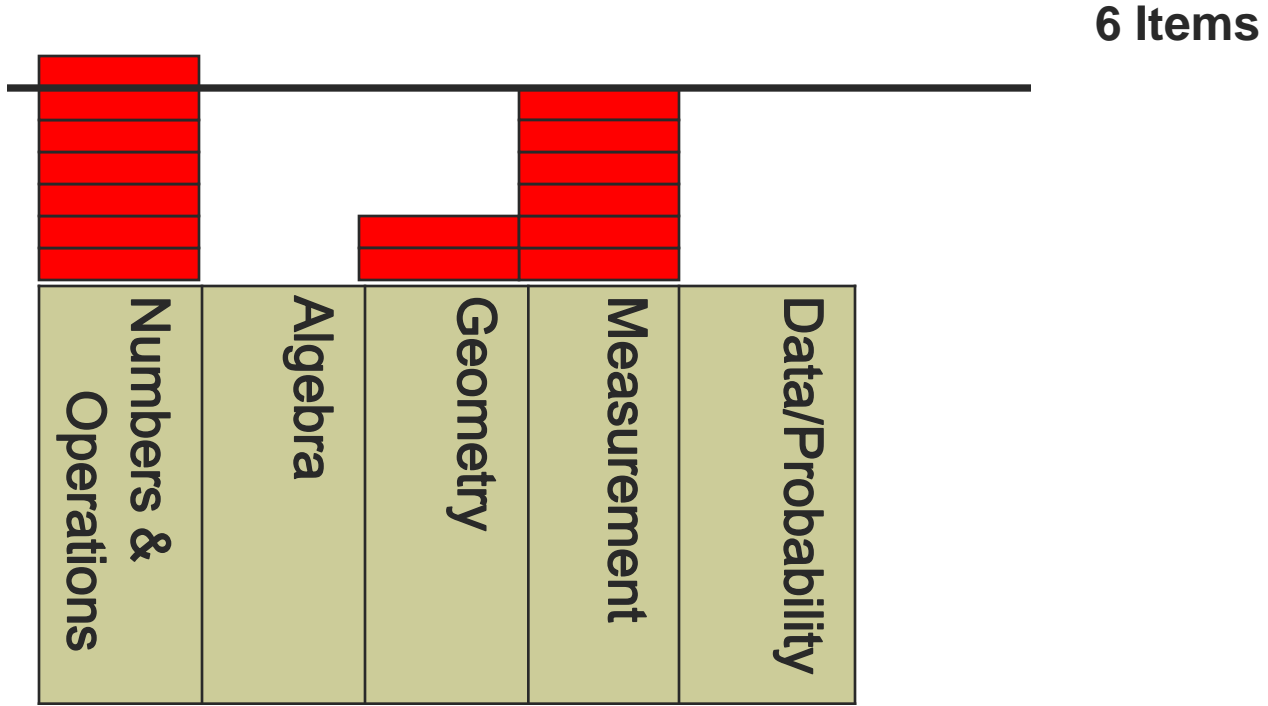
- The content differs from grade level in range, balance, and DOK, but matches high expectations set for students with significant cognitive disabilities.
- **What is rated:** Extended standards (if applicable), AA items
- **Rated by:** Content experts

[Webb's Alignment Dimensions]

- Content Concurrence
- Depth of Knowledge
- Range of Knowledge
- Balance of Representation

○ Based on the work of Webb (1997)

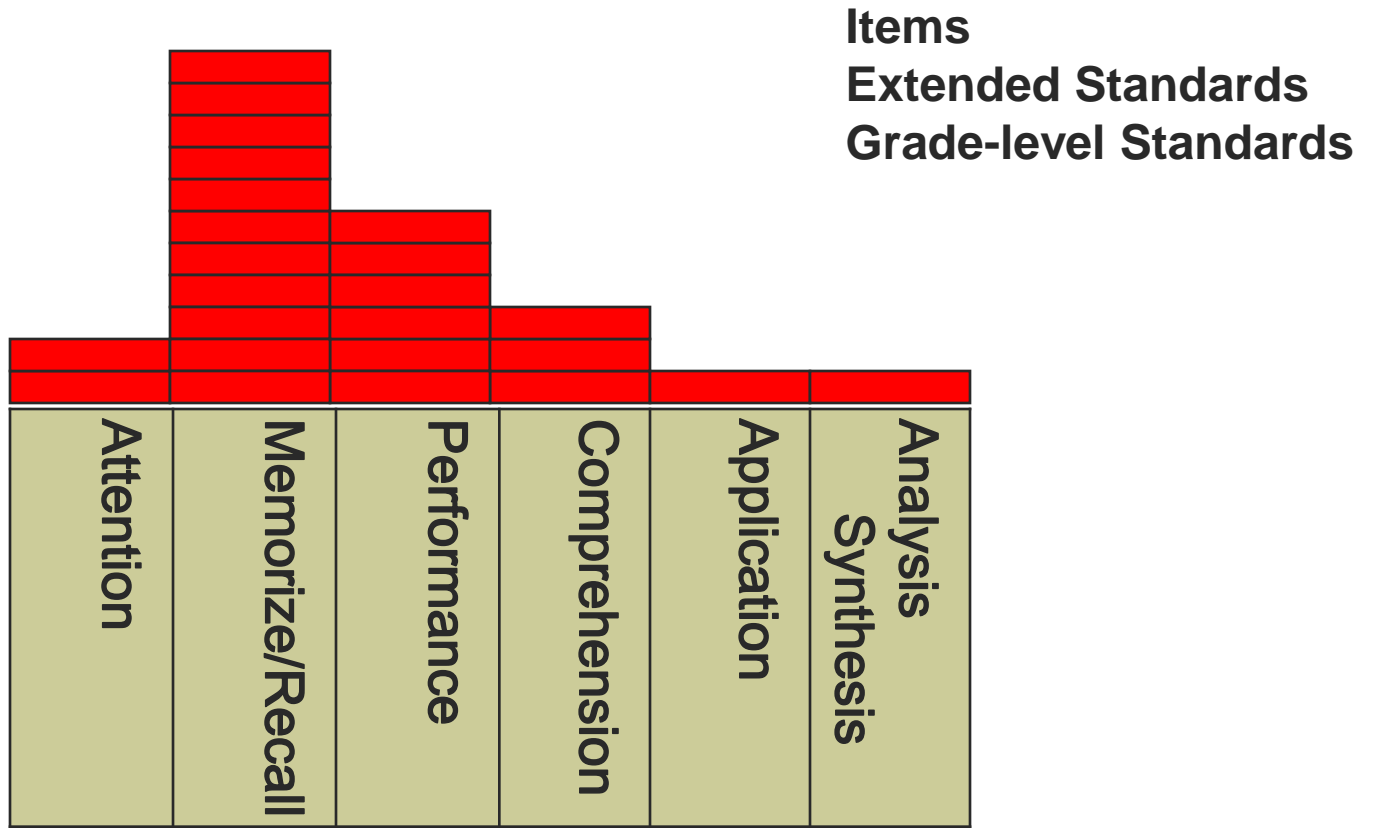
Categorical Concurrence



Categorical Concurrence

Grade	% of Standards w/ at least one item	Categories not represented*	% of standards w/ six or more items	Categories that are under- represented
3 rd	60%	#4, 5	60%	#4, 5
4 th	60%	#4, 5	60%	#4, 5
5 th	60%	#4, 5	60%	#4, 5

Depth of Knowledge



[Reporting DOK]



DOK Levels	3rd	4th	5th
Attention	7.3	6.9	.0
Memorize/Recall	70.9	65.5	59.4
Performance	7.3	19.0	12.5
Comprehension	12.7	6.9	28.1
Application	1.8	1.7	.0
Analysis/Synthesis/Evaluation	.0	.0	.0
Number of Items	55	58	32

Reporting DOK

Grade	% Below	% At	% Above		% At or Above
3 rd	89.1	10.9	.0		10.9
4 th	93.1	5.2	1.7		6.9
5 th	75.0	25.0	.0		25.0
6 th	94.6	5.4	.0		5.4
7 th	78.4	21.6	.0		21.6
8 th	80.0	10.0	10.0		20.0
10 th	65.7	25.4	9.0		34.3

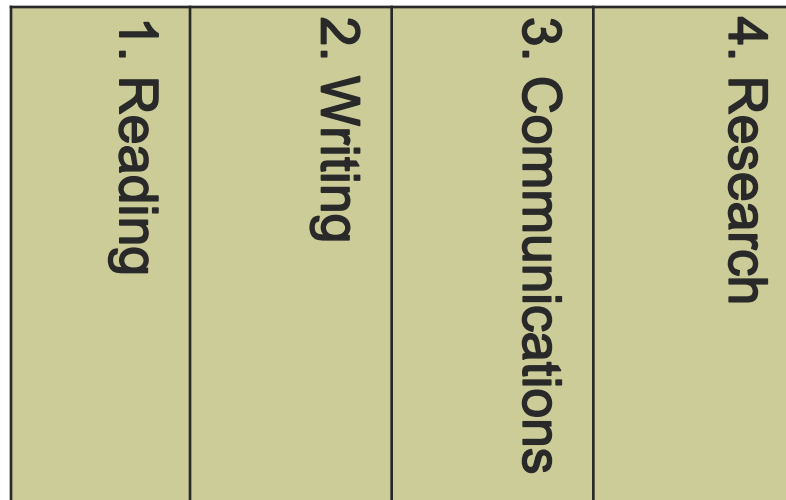
[Balance of Representation]

$$\textit{Balance} = 1 - \left(\sum_{i=1}^o \left| \frac{1}{O} - \frac{I_k}{H} \right| \right)$$

Indices > .7 indicate criterion is met

Range-of-Knowledge Correspondence

50% of objectives have a corresponding item



2. 1 The Writing Process-The student will apply a process approach to writing.

2.2 Writing Purposes: The student will write for a variety of purposes.

2.3 Responding to Texts-The student will respond to texts written by others.

2.4 Legibility-The student will create legible texts.

[Evaluative Judgment]

- AA-AAS should not meet Webb's criteria
- Purpose and intended reduction of scope and depth should mediator the relationship

[Special Considerations]

- Common items vs. vary by student
- As a general rule, more items will improve alignment but may not improve the quality of AA-AAS

[Criterion 5]

- There is some differentiation in content across grade levels or grade bands.
- **What is rated:** Extended standards (if applicable), AA items
- **Rated by:** Content experts

[Differentiation]

- This criterion evaluates the extent of differentiation of academic content across multiple grades.
 - Change in content standards
 - Change in DOK
 - Vertical relationship (broader, deeper, prerequisite, new, & identical) (Webb, 2005)
- Age and grade appropriate

Example of Differentiation

Strand	Differentiation- broader, deeper, new or used prerequisites at lower grades	Lack of Differentiation- identical items occur	Examples of redundant items for undifferentiated strands or points of interest	Overall
Algebra	Clear evidence of deeper applications. Partial evidence was found for broader applications, prerequisite knowledge/skills. Limited evidence of any new skills.	Partial.	e.g., extend a pattern	Needs Improvement.
Geometry	Clear relationship for deeper applications. Partial evidence of broader applications and prerequisite skills. Limited evidence of new skills.	Limited.		Strong Differentiation.

Example of Age Appropriateness Rating

Supporting Evidence Age Appropriateness

Age Appropriateness Rating	N	%
Not age appropriate	196	12.4
Age neutral	853	54.1
Age appropriate	362	22.9
Not enough detail to rate	78	4.9
Missing data	75	4.8
Unable to view video	14	.9

[Results]

- The percentage of AA items and extended standards within a content standard
- The number and percentage of items that are age inappropriate. A list of items should be provided.
- Holistic rating by content experts describing the vertical relationship between grade level standards

[Evaluative Judgment]

- There should be changes in content standards across the grade levels.
- All ratings of vertical relationships are acceptable except for *identical* (i.e., higher-grade standard appear identical to one of the lower-grade standards).
- All items should be age appropriate..

Criterion 6

- The expected achievement for students is for the students to show learning of grade referenced academic content.
- **What is rated:** Scoring procedures (criteria) for alternate assessment
- **Rated by:** Special education experts, alignment team

[Results]

- Accuracy
- Level of independence
- New learning
- Complexity
- Generalization (setting)
- Conceptual generalization

[Evaluative Judgment]

- Scoring criteria should be based on what the student knows and can do in the academic domain.
- While the purpose of some AAs is the improvement of instructional programs for students with significant cognitive disabilities, program improvement criteria should not be part of the student scores.

[Special Considerations]

- If the purpose of AAs is to measure progress, baseline or pretest should be new learning.
- Rationale and policy for inclusion of some scoring criteria should be reported.

[Criterion 7]

- The potential barriers to demonstrating what students know and can do are minimized in the assessment.
- **What is rated:** AA items, AA administration manual
- **Rated by:** Special education experts

Barriers for Student Performance

- Provides information about accessibility of the alternate assessment to the wide range of students with significant cognitive disabilities
 - Review:
 - assessment materials
 - special education experts' ratings of each academic AA item to determine whether or not symbolic communication is required to answer the item
 - special education review of nonacademic items

[Results]

- No provision
- Flexibility built into tasks
- Accommodations
- Modification

[Evaluative Judgment]

- All students, regardless of disability, should have access to the general curriculum. Provisions for the inclusion of students with different characteristics should be documented and evident in the AA.

Special considerations

- Materials needed to complete the ratings for criterion 7 vary considerably depending upon the state's system of documentation related to the AA-AAS. The needed documents can be identified through the review of publicly available materials and the interview with the state's key informants or liaison for the study.

Criterion 8

- The instructional program promotes learning in the general curriculum.
- **What is rated:** Professional development materials
- **Rated by:** Special education experts (Content experts rate CIS strands during first step of their coding responsibilities)

Other Sources of information

- Professional Development Checklist
- Program Quality Indicators Checklist
- Curriculum Indicators Survey

Results Reported

- Narrative summary of findings from Program Quality Indicators Checklist
 - Opportunities for instruction in general education classroom
 - Opportunities for instruction with typical peers
 - Opportunities for students to make choices, problem solve...
 - Provision of assistive technology
 - Access and use of typical classroom resources
 - Literacy being promoted across the content areas
 - Meaningful linking of academic skills in functional contexts

[Professional Development]

- Instructional alignment to grade level standards
 - Examples
- Achievement descriptors and target skills
- Assistive technology and other supports

Curriculum Indicator Survey

- Percent of items within each cell of a content x DOK table
- Discrepancy between content x DOK configuration reported on CIS and other elements of the system (where strand matches are possible; varies by state)
 - Adapted from the work of Porter & Smithson, 2001
 - Created by Meagan Karvonen and Shawnee Wakeman

Evaluative Judgment

- Professional development should demonstrate a clear link to the major academic strands.
- Academic content should be referenced to student's assigned grade level.
- The focus of achievement maintains fidelity with the content of the grade level content standards.
- There is a differentiation in content across grades.
- The emphases in the enacted curriculum should match the emphases in the content standards (as measured with the CIS).

Special considerations

- CIS topics are standardized and do not always reflect the unique curricula in each state. Reporting results at the strand level allows for comparisons with the state's curriculum.
- CIS results provide a snapshot of teachers' self-reported instructional practices and the enacted academic curriculum for a sample of students. However, there are limits to generalizability depending on sample size and response rate.

[The End!!!]

■ Selected References

- Porter, A. C. & Smithson, J. L. (2001). *Defining, developing, and using curriculum indicators* (CPRE Research Report Series RR-048). University of Pennsylvania: Consortium for Policy Research in Education.
- Rothman, R., Slattery, J. B., Vranek, J. L., Resnick, L. B. (2002). *Benchmarking and alignment of standards and testing*. National Center for Research on Evaluation, Standards, and Student Testing. Retrieved June 2, 2005, from <http://www.cresst.org/>
- Webb, N. L. (1997). *Criteria for alignment of expectation and assessments in mathematics and science education*. (NISE Research Monograph No. 6). Madison: University of Wisconsin-Madison, National Institute for Science Education. Washington, DC: Council of Chief State School Officers.