

State C Alternate Assessment Based on Alternate Achievement Level Alignment
Report

Links for Academic Learning:

Report to the State Department of Education

Prepared by

Claudia Flowers, Diane Browder, Shawnee Wakeman, and Do-Hong Kim

University of North Carolina at Charlotte

Meagan Karvonen

Western Carolina University

August 2007

Contact:

Claudia Flowers

Department of Educational Leadership

University of North Carolina at Charlotte

cpflower@email.uncc.edu

704.687.8862

This report supported by the Office of Special Education Programs (OSEP grant No. #H324U040001). Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the view of the supporting agency.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	3
INTRODUCTION.....	7
DESCRIPTION OF THE STATE C ALTERNATE ASSESSMENT.....	7
DESCRIPTION OF EXPERTS.....	8
RELIABILITY.....	8
CRITERIA FOR ALIGNING ALTERNATE ASSESSMENTS TO GRADE LEVEL CONTENT.....	10
CRITERION 1.....	13
READING.....	13
MATH.....	15
SCIENCE.....	17
CRITERION 2.....	21
CRITERION 3.....	22
READING.....	22
MATH.....	24
SCIENCE.....	26
CRITERION 4.....	27
READING.....	29
MATH.....	34
SCIENCE.....	39
CRITERION 5.....	43
CRITERION 6.....	51
CRITERION 7.....	52
CRITERION 8.....	53
SUMMARY AND RECOMMENDATIONS.....	56
REFERENCES.....	57

EXECUTIVE SUMMARY

This report details findings from an investigation of the alignment of State C's alternate assessments based on alternate achievement standards in reading, math, and science. The criteria used in this alignment study are being evaluated as part of the UNC Charlotte partnership in the *National Alternate Assessment Center* (NAAC). This report is organized by the eight criteria developed by a collaboration of content experts, special educators, and measurement experts at UNC Charlotte (Browder, Wakeman, Flowers, Rickleman, Pugalee, & Karvonen, 2006). While some of the alignment criteria are similar to other alignment methods (e.g., Webb, Surveys of Enacted Curriculum, and Achieve), some of the criteria (criteria 5-8) were designed specifically as value indicators for students with significant cognitive disabilities (see Table 1). An additional difference between this alignment protocol and other alignment methods is the examination of the targeted standards (i.e., standards intentionally selected for students with significant cognitive disabilities) and grade-level content standards. This summary describes how well the interpretation of state standards (Grade level and Alternate Content Standards), the alternate assessment and instruction (professional development) met the eight criteria for alignment.

Alignment Results by Criterion

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., reading, math, science).*

Outcome: Content experts rated alternate content standards, alternate achievement standards, and alternate assessment items as academic or nonacademic.

For reading and math, almost all of the alternate content standards and alternate achievement standards were rated academic. Most (87%) of the reading items were rated academic, and those that were not were either foundational or presymbolic/concrete items. Given the characteristics of students with significant disabilities, the numbers of foundational and presymbolic/concrete items are within an acceptable range. Almost all of the math items (98%) were rated academic.

All but one of the science alternate content standards were rated academic and 87% of the alternate achievement standards were rated academic. Science items appeared more problematic. Approximately 59% of the AA items were rated academic. The reason for the low rating appeared to be the passive expectations of the AA items or the language was too vague for the raters to understand the student expectations. Most of the nonacademic items were rated at a presymbolic/concrete symbolic level of communication.

Criterion 2: *The content is referenced to the student's assigned grade level (based on chronological age).*

Outcome: All alternate assessment items were referenced to the student's assigned grade level. In reading, all items were aligned to the four grade level indicators of reading. For math, the grade level strands of algebra, geometry, measurement, number sense, and statistics and probability were referenced. Science items were referenced to the strands of

(a) nature of science, (b) physical science, (c) life science, (d) earth/space science, and (e) science, technology, environment, and society.

Criteria 3: *The focus of achievement maintains fidelity with the content of the original grade level standards (content centrality) and when possible, the specified performance.*

Outcome: All but two of the reading alternate content standards were rated as having a far or near content and a partial or full performance centrality. Most of the reading items were rated as having a far or near content centrality (82%) and partial or full performance centrality (82%). Supporting evidence for reading had similar percentages of alignment to standards. Approximately 83% of all the supporting evidence had some or all link to the AA item.

In math, all the alternate content standards were rated as having a far or near content and a partial or full performance centrality. Most of the math items had a far or near content centrality (96%) or partial or full performance centrality (96%). Approximately 83% of all the supporting evidence had some or all link to the AA item.

In science, all the alternate content standards were rated as having a far or near content and a partial or full performance centrality. All but one of the science items were rated as having a far or near content (99%) and a partial or full performance centrality (99%). Approximately 63% of the science supporting evidence was rated as linked to AA item.

It should be noted that these percents only apply to items that were rated academic (criteria 3). For example, the 41% of science AA items that were rated nonacademic could not be analyzed for content and performance centrality.

Criterion 4: *The content differs from grade level in range, balance, and DOK, but matches high expectations set for students with significant cognitive disabilities.*

Outcome: The alternate assessment across all academic domains had acceptable categorical concurrence, depth of knowledge, range-of-knowledge, and balance of representation. There were items distributed across all levels of depth of knowledge. The range-of-knowledge and balance of representation suggested the alternate assessments had exceptional content coverage.

Criterion 5: *There is some differentiation in achievement across grade levels or grade bands.*

Outcome: Overall the majority of strands for the alternate content standards and AA items demonstrated a strong differentiation in achievement across grade levels. This state provides exemplary differentiation expected in academic achievement across grade levels. Only three strands of math (i.e., algebra, statistics and probability, and number sense) were found to have more than a few repeated items across grade levels. Also most of the supporting evidence submitted by the teachers was rated age appropriate (23%) or age neutral (54%).

Criterion 6: *The expected achievement for students is for the students to show learning of grade referenced academic content.*

Outcome: Overall this system has the potential for high student inference about student learning academic content. In the rating scale used for the checklist tasks and pieces of evidence, more credit is given to higher levels of accuracy with less prompting. Credit also is awarded for applying skills across settings. In standard setting, consideration should be given to the inference that can be made from where cut off scores are established.

Criterion 7: *The potential barriers to demonstrating what students know and can do are minimized in the assessment.*

Outcome: The alternate assessment items were accessible to students at all levels of symbolic communication and many of the items allowed flexibility in response mode. The assessment system allows for accommodations and adaptations for all students.

Criterion 8: *The instructional program promotes learning in the general curriculum.*

Outcome: While there clearly is training on instructional alignment provided, further development of this may be needed to help teachers link to grade level content, increase expectations across grade bands, and utilize inclusive educational opportunities whenever appropriate. More information on promoting student mastery (fading prompts) and teaching across materials/activities may also be needed if recommendations are followed to change expectations for proficiency.

Overall Analysis of Alignment

There is clear evidence that the alternate assessment was aligned to the grade level content standards. Almost all of the reading and math alternate standards and AA items were rated academic. Science items were more problematic with only 59% of the AA items rated academic and it is recommended that these items be reviewed for improvement. The reason for the low rating appeared to be the passive expectations of the AA items or the language was too vague for the raters to understand the student expectations.

Alignment to grade level content standards was considered in the development of the alternate assessment. All of the alternate content standards and AA items were referenced to the grade level content standards and demonstrated excellent content coverage.

Overall, the alternate assessment demonstrated strong differentiation in content and expectations across the grade levels, although some additional items in some of the strands of

math may be needed in future development of the alternate assessment. Additionally, most of the supporting evidence was rated as age neutral or age appropriate.

The alternate assessment items were accessible to students at all levels of symbolic communication and many of the items allowed flexibility in response mode but one recommendation should be considered. Additional information or examples on how all students have access is recommended.

Teacher training was aligned to the grade level content standards. To improve teacher training, it is recommended that more information and training be developed on promoting student mastery and other areas of best practice.

**ALTERNATE ASSESSMENT ALIGNMENT
REPORT TO THE STATE C STATE DEPARTMENT OF EDUCATION**

This alignment study was conducted on the basis of information obtained on the 2007 State C State Test of Educational Progress-Alternate Assessment (STEP-A). The criteria in this alignment study are being evaluated as part of the UNC Charlotte partnership in the *National Alternate Assessment Center* (NAAC). This report is organized by the eight criteria developed by a collaboration of content experts, special educators, and measurement experts at UNC Charlotte (Browder, Wakeman, Flowers, Rickleman, Pugalee, & Karvonen, 2006). While some of the alignment criteria are similar to other alignment methods (e.g., Webb, Surveys of Enacted Curriculum, and Achieve), some of the criteria (criteria 5-7) were designed specifically as value indicators for students with significant cognitive disabilities (see Table 2). An additional difference between this alignment protocol and other alignment methods is the examination of the targeted standards (i.e., standards intentionally selected for students with significant cognitive disabilities) and grade-level content standards. These results are intended to inform decision makers about what content standards are being emphasized and what standards are being excluded for students with significant cognitive disabilities when compared with the general education population.

All reviewers were instructed on the purpose of alternate assessments and reviewed all the testing materials and academic content standards provided by the state of State C. The content reviewers rated the alignment of AA items to content standards as a team until there was consensus. After both the content experts had consensus, they rated subsequent items independently. Independent ratings of some common items were used to evaluate inter-rater agreement. Special education experts rated the professional development materials, and the age appropriateness and symbolic levels of the standards and alternate assessment items.

DESCRIPTION OF STATE C STANDARDS AND ALTERNATE ASSESSMENT

The State C Alternate Content Standards

The Alternate Content Standards based on the developmental spectrum of the State C Content Standards are aligned to the general education goals/strands and indicators. The Alternate Content Standards were developed to allow all students the opportunity to access the general education curriculum. The Alternate Content Standards encompass grades K–12 in reading, grades K–8 and grade spans for grades 9–12 in math and science. The Alternate Standards provide alternate achievement descriptors and target skills for each performance level and for each grade. Alternate achievement descriptors are organized into performance levels, indicating how a student at that level would be expected to perform on the Alternate Content Standards. Target skills, defined as higher level skills that enable students with disabilities to individually utilize the Alternate Content Standards, serve as entry points to the Alternate Standards, providing a range of options at which a student with a disability can access the learning standards at a challenging level. Students working in the Alternate Content Standards take the alternate assessment, the State C State Test of Educational Progress-Alternate (STEP-A) (<http://doe.sd.gov/contentstandards/alternatestandards/>).

The State C Alternate Assessment

The State C State Test of Educational Progress-Alternate (STEP-A) consists of two components: the Rating Scale and Supporting Evidence. The State C alternate assessment consists of rating forms that are aligned to the Alternate Content Standards and are grade specific focusing on reading and math, at grades 3-8 and 11, with additional of science in grades 5, 8, and 11. Items were developed based on alternate content standards that were developed based on the developmental spectrum of the State C content standards. The rating scale consists of seven items for each standard under an indicator on the assessment. Two raters, primary special education teacher and colleague who knows and works with the student, are responsible for rating a student's performance on a series of skills. Ratings are based on each rater's knowledge of the student's current performance level. The scale consists of five performance levels (i.e., nonexistent, minimal, emerging, progressing, and accomplished). In addition, special education teachers provided one piece of supporting evidence, including work sample, data collection form, media (photo, video and audio), and other sources, for each of the indicators to support the rating for the task selected on the rating scale. Supporting evidence is needed for 4 reading, 5 math, and 5 science, one for each content indicator/strand (<http://doe.sd.gov/oess/specialed/news/altassessment.asp>).

EXPERT REVIEWERS

The alignment team consisted of two English Language Arts (ELA) experts, two Mathematics experts, two Science experts, four experts in the education of students with significant cognitive disabilities, and two measurement experts. Content experts had a range of experience in their content area of 9 to 31 years and special education experts had a range of 7-30 years. The level of education ranged from a bachelor's degree (1 content expert) to a doctoral degree or individuals participating in a PhD program (2 content experts, 4 special education experts, 2 measurement experts). All experts had participated in conducting professional development related to their content area. Five of the experts have been involved in curriculum writing on their district, state, or national level (e.g., textbook writing). One of the experts was nationally board certified in their content area. Five experts had a licensure in curriculum and instruction. Six experts taught higher education classes. Finally, eight of the experts had been item writers for their state's general and special education assessments.

At the beginning of each alignment activity, the team worked together to come to a consensus on the alignment of educational components. When experts disagreed, decision rules were made to ensure consistency. Then the reviewers independently rated a subset of items/standards/alternate standards and agreement between raters was examined. When the raters agreed 90%, each rater was given specific tasks. Reliability was checked periodically throughout the tasks to ensure consistent ratings.

RELIABILITY OF CODING

The interrater reliability for each dimension coded by experts is report in Table 1. At least 15% of the items for each content area and level of alignment were double coded to establish interrater reliability. The range of interrater agreement for ELA content experts for grade level

standards overall was 95%, 90% for alternate standards overall, and 92% for alternate assessment items overall. Interrater agreement for math content experts for grade level standards overall was 96%, 89% for alternate standards overall, and 89% for alternate assessment items overall. Interrater agreement for Science content experts for grade level standards overall was 98%, 96% for alternate standards overall, and 95% for alternate assessment items overall. Overall interrater agreement for special education experts on AA items was 85%. Finally, the dataset was cleaned with 20% of the entries checked for data entry errors. Data entry was 94-100% accurate for all items related to ELA, 93-100% accurate for all items related to math, and 94-99% accurate for Science, and 96% accurate for special education.

Table 1: Interrater Agreement for Components Rated

Component Rated	N Double Coded	% Double Coded	% Agreement
<i>ELA</i>			
<i>Grade level</i>			
National Standards	64	30%	100%
DOK	47	26%	89%
<i>Alt Standards</i>			
Academic?	56	22%	100%
National Standards	56	23%	100%
Content Centrality	56	23%	96%
Performance Centrality	56	23%	96%
DOK	56	23%	86%
<i>AA items</i>			
Academic?	57	29%	98%
National Standards	65	33%	97%
Foundational	11	44%	91%
Content Centrality	47	28%	92%
Performance Centrality	47	28%	92%
DOK	47	28%	92%
<i>Math</i>			
<i>Grade level</i>			
National Standards	56	26%	100%
DOK	56	26%	88%
<i>Alt Standards</i>			
Academic?	117	17%	100%
National Standards	117	17%	100%
Content Centrality	117	17%	96%
Performance	117	17%	96%

Centrality			
DOK	117	17%	83%
<i>AA items</i>			
Academic?	126	23%	100%
National Standards	125	23%	100%
Foundational	1	9%	100%
Content Centrality	125	23%	99%
Performance Centrality	125	23%	99%
DOK	125	23%	79%
<i>Science</i>			
<i>Grade level</i>			
National Standards	44	32%	82%
DOK	29	29%	97%
<i>Alt Standards</i>			
Academic?	63	15%	100%
National Standards	59	16%	98%
Content Centrality	59	16%	100%
Performance Centrality	59	16%	100%
DOK	59	16%	93%
<i>AA items</i>			
Academic?	59	28%	100%
National Standards	43	27%	85%
Foundational	28	33%	86%
Content Centrality	30	24%	97%
Performance Centrality	30	24%	93%
DOK	30	24%	90%
<i>Special Education</i>			
Accessibility	234	16%	92%
Age Appropriateness	234	17%	83%
Link of Evidence	234	17%	81%

CRITERIA FOR ALIGNING ALTERNATE ASSESSMENTS TO GRADE LEVEL ACADEMIC CONTENT

Non-regulatory guidance has specified that alternate assessments “should be clearly related to grade-level content, although it may be restricted in scope or complexity or take the

form of introductory or prerequisite skills” (U.S. Department of Education, 2005, p.26). As stated in this regulation, there should be a clear *link* to the content standards for the grade in which the student is enrolled. While this gives states flexibility to determining the scope and breadth of content of alternate assessments, it does not exempt states from designing assessments that measure an academic domain with interpretable results and accurately reflecting what the student knows and can do within that academic domain. For this reason, the authors believe that the investigation of alignment between academic content, academic performance, alternate assessments, and instructional practices and resources should be as strenuous as those used for the assessment of students in the general population. In contrast, it is also expected there would be some differences in the depth, breadth or complexity of content addressed when the achievement target is an alternative to grade level achievement. Because of the unique characteristics and needs of students with significant cognitive disabilities (e.g., testing formats and instructional practices), additional alignment criteria also need to be considered for alternate assessments.

In our conceptual framework, we propose eight criteria for linking to grade-level academic content standards (see Table 2). To be linked to grade level standards, the target for achievement must be academic content (e.g., reading, math, science) that is referenced to the student’s assigned grade based on chronological age. Functional activities and materials may be used to promote understanding, but the target skills for student achievement are academically-focused. Some prioritization of the content will occur in setting this expectation, but it should reflect the major domains of the curricular area (e.g., strands of math) and have fidelity with this content and how it is typically taught in general education. The alternate expectation for achievement may focus on prerequisite skills or some partial attainment of the grade level, but students should still have the opportunity to meet high expectations, to demonstrate a range of cognitive demand, to achieve within their level of symbolic communication, and to show growth across grade levels or grade bands.

Table 2: Criteria for Instruction and Assessment that Links to Grade Level Content

1. The content is academic and includes the major domains/strands of the content area as reflected in state and national standards (e.g., reading, math, science).
2. The content is referenced to the student's assigned grade level (based on chronological age).
3. The focus of achievement maintains fidelity with the content of the original grade level standards (content centrality) and when possible, the specified performance.
4. The content differs from grade level in range, balance, and DOK, but matches high expectations set for students with significant cognitive disabilities.
5. There is some differentiation in content across grade levels or grade bands.
6. The expected achievement for students is for the students to show learning of grade referenced academic content.
7. The potential barriers to demonstrating what students know and can do are minimized in the assessment.
8. The instructional program promotes learning in the general curriculum.

The following sections report the results of the pilot alignment method organized around the eight criteria, as applied to State C alternate assessment system. In each section, results are reported for measurement guidelines and alternate assessments in ELA and math.

ALIGNMENT RESULTS

Criterion 1: The content is academic and includes the major domains/strands of the content areas as reflected in national standards as defined by the National Council of Teachers of English (NCTE), National Council of Teachers of Mathematics (NCTM), and the National Research Council.

Criterion 1 criterion requires content experts to determine if the AA items, Alternate content standards, and grade level alternate achievement descriptors are academic. AA items that are not academic are not included in any further analyses. In order to be rated nonacademic, content experts must agree (either by identical, independent ratings or by consensus after different ratings) that the item was indeed nonacademic. To make these decisions, reviewers are asked whether the item/standard/descriptor can be logically defined by a national standard for that content area. National standards are defined according to the national content organizations (National Council of Teachers of English, National Council of Teachers of Math, and National Research Council). If the standard or assessment item can be labeled according to a strand within the national standards, the reviewer codes it as academic.

Some skills may be *foundational skills* or those skills which are an assumed competence across all grade levels specific to an academic context such as turning the pages of a book. While these skills are not academic skills (i.e., cannot be defined by a national strand), they may be appropriate for some students with significant cognitive disabilities. Special educators review the items/standards/descriptors that were rated by content experts as nonacademic to identify those components that may be foundational skills.

Nonacademic items are also coded for symbolic level of communication. A student at the awareness/presymbolic level communicates through gestures and objects (e.g., holding up a cup when thirsty). In the concrete symbolic level, the student uses some symbols to communicate (e.g., provides teacher with a picture of a cup when thirsty). At the abstract symbolic level, the student communicates with multiple symbols (e.g., concrete and abstract picture symbols) or words (e.g. provides teacher with written word “cup” when thirsty).

The following analyses indicate the number and percentage of AA items, alternate content standards, and alternate achievement standards that were rated academic. To make these decisions, raters were asked whether the item can be logically defined by a national standard for that content area. National standards are defined according to the national content organizations (National Council of Teachers of English, National Council of Teachers of Math, and National Research Council).

Reading Results

Alternate content standards. All the Alternate Content Standards (N=49) were rated academic. All but 1 of the 197 grade-level alternate achievement descriptor were rated academic. The descriptor rated as nonacademic was “Students are able to identify a book and look at a book.”

AA items. A total of 193 AA reading items were rated by content experts as academic or nonacademic. Overall, 87% ($n=168$) of the AA reading items were rated academic. Of the 25 nonacademic items, 68% ($n=17$) were rated foundational (i.e., skills that students are assumed to be competent—“turn the page”) and 96% ($n=24$) of the nonacademic items were rated at the presymbolic/concrete level of symbolic communication. Given the characteristics of students with significant disabilities, the numbers of foundational and presymbolic/concrete items are within a reasonable range. The number of AA items, number rated academic and percentage of items that were academic are reported in Table 3.

Table 3: Number and Percentage of AA Items Rated Academic

Grade	Items Total	Academic	%
3 rd	28	22	78.6
4 th	27	21	77.8
5 th	28	23	82.1
6 th	28	24	85.7
7 th	26	23	88.5
8 th	28	27	96.4
11 th	28	28	100.0
All	193	168	87.0

Note. There should be 28 items across all grade levels. Three items (1 for 4th grade and 2 for 7th grade) were inadvertently missed during coding.

A list of the items that were rated nonacademic is reported in Table 4.

Table 4: AA Items Rated Nonacademic

Grade	Item	Foundational	Presymbolic/ Concrete
3	Explores and responds to a book (e.g., holds a book, turns pages).	Yes	Yes
3	Attends to pictures, functional symbols, and objects that are pertinent to a story	Yes	Yes
3	Imitates "real" sound effects (e.g., "What does the pig say?")	No	Yes
3	Attends to a story.	Yes	Yes
3	Identifies a familiar object or person by eye gaze, pointing, or assistive technology.	No	Yes
3	Demonstrates a differentiated awareness to access visual, tactile, auditory, and olfactory stimuli (e.g., shows a different response to his/her mother than to a stranger).	No	Yes
4	Responds to an auditory or visual presentation of a book.	Yes	Yes
4	While a story is read aloud, attends to the story from beginning to end.	Yes	Yes
4	Recognizes and responds to his/her name	No	Yes
4	When given simple questions regarding gender, responds appropriately (e.g., "Are you a boy?")	No	No
4	Shares something about himself/herself that is true (e.g., "I have a kitten.")	Yes	Yes
4	Shows interest in books	Yes	Yes
5	When experiencing difficulties, seeks adult help by any means	Yes	Yes
5	Attends to books on tape from various time periods	Yes	Yes

5	Attends to books on tape from various cultures	Yes	Yes
5	Attends to a variety of informational texts.	Yes	Yes
5	When moving to various classes, locates his/her own class schedule for the day	No	Yes
6	Attends to text containing similes	No	Yes
6	Repeats likenesses between characters after they are explained by the teacher	Yes	Yes
6	Attends to books on tape from various time periods	Yes	Yes
6	Attends to books on tape that reflect various cultures	Yes	Yes
7	Attends to pictures/symbols/objects pertinent to a story	Yes	Yes
7	Attends to pictures that show the setting or characters in a story	Yes	Yes
7	Attends to music and lyrics from different cultures and time periods	No	Yes
	Attends/responds to a newspaper weather report or an Internet weather report to make decisions about transportation and clothing	Yes	Yes
8			

AA items that were academic were aligned to national standards (as defined by NCTE). The alignment of AA items to national standards is reported in Table 5. Most of the items were aligned to reading (62.6%).

Table 5: Number and Percentage of AA Items in ELA Components

Components	AA Items		AA Standards	
	N	%	N	%
Reading	102	62.6	191	77.3
Writing	5	3.1	1	.4
Speaking	3	1.8	9	3.6
Listening	5	3.1	10	4.0
Viewing/Visual	21	12.9	1	.4
Research	27	16.6	35	14.2

Note. 5 AA items were inadvertently not rated.

Math Results

Alternate content standards. All the math Alternate Content Standards ($N=135$) were rated academic. All but 1 of the 681 grade level alternate achievement descriptor was rated academic. The nonacademic descriptor was “Students are able to recognize the difference between morning, afternoon, and night.”

AA items. A total of 560 mathematic AA items were rated by content experts as academic or nonacademic. Almost all (98%) of the AA items were rated academic. The number of AA mathematic AA items, the number rated academic, and the percentage rated academic are reported in Table 6.

Table 6: Number and Percentage of Mathematic AA Items Rated Academic

Grade	N of Items	Academic	%
3 rd	81	78	96.3
4 th	77	75	97.4

5 th	84	83	97.8
6 th	77	76	98.7
7 th	82	81	98.8
8 th	82	80	97.6
11 th	77	77	100.0
All	560	550	98.1

A list of the items that were rated nonacademic is reported in Table 7.

Table 7: AA Items Rated Nonacademic

Grade	Item	Foundational	Presymbolic/ Concrete
3	Follows simple one-step directions	Yes	Yes
3	Identify the cause of a simple event (e.g., darkness in the room=lights turned out)	No	No
3	Predicts when the class bell will ring or when it is almost lunchtime or the end of the day.	No	Yes
4	Plays with real money	No	Yes
4	Predicts when the class bell will ring or when it is almost lunchtime or the end of the day	No	Yes
5	Predicts when the class bell will ring or when it is almost lunchtime or the end of the day	No	Yes
6	Predicts when the class bell will ring or when it is almost lunchtime or the end of the day	No	Yes
7	Predicts when the class bell will ring or when it is almost lunchtime or the end of the day	No	Yes
8	Predicts the outcome of a given event such as flipping a light switch	No	No
8	Predicts when the class bell will ring or when it is almost lunchtime or the end of the day	No	Yes

AA items that were academic were aligned to national standards (as defined by NCTM). The alignment of AA items to national standards is reported in Table 8. Items were distributed across all strands of math.

Table 8: Number and Percentage of AA Items in National Math Strands

Components	AA Items		Alternate Standards	
	N	%	N	%
Numbers and Operations	165	29.7	181	26.5
Algebra	163	29.3	187	27.4
Geometry	96	17.3	120	17.6
Measurement	47	8.5	96	14.1
Data/Probability	85	15.3	96	14.1

Science Results

Alternate content standards. All (N=61) but one of the Alternate Content Standards were rated academic. The following standard was rated nonacademic —“Students are able to label the colors found in the spectrum of light.” A total of 310 grade level alternate achievement descriptor review. Most of the descriptors were rated academic (n=270, 87%), with some rated nonacademic (n=34, 11%) and 5 (1.6%) were too vague to code. The nonacademic descriptors are listed in Table 9.

Table 9: Alternate Achievement Descriptors Rated Nonacademic

Grade	Descriptor
5	Students are able to label the colors found in the spectrum of light.
5	Students are able to recognize the colors found in the spectrum of light.
5	Students are able to explore visual and or tactile aids of plants.
5	Students will respond to illustrations of parents and their offspring.
5	Students are able to attend to stimuli of ecosystems.
5	Students are able to identify items related to a season.
5	Students are able to explore items related to seasons.
5	Students are able to show a response to the sun, moon, and Earth.
5	Students attend to facts and predictions.
8	Students will attend to the activities about the Periodic Table.
8	Students are able to observe matter.
8	Students are introduced to different forms of matter.
8	Students are able to experience different weather conditions.
8	Students are able to manipulate objects that have been eroded.
8	Students are able to sequence order of the planets according to size.
8	Students are able to attend to the concept of planets.
8	Students will attend to activities that involve objects that meet their social needs.
8	Students are able to attend to problems.
8	Students will observe a simple scientific experiment.
11	Students observe safe laboratory techniques.
11	Students are able to access the Periodic Table.
11	Students are exposed to different substances.
11	Students are exposed to different reactions.
11	Students are exposed to different types of matter.
11	Students will observe change.
11	Students are exposed to time and distance through activities.
11	Students will participate in movement activities.
11	Students will observe different types of waves.
11	Students attend to the concept of cells.
11	Students will attend to presentation on extinct animals.
11	Students attend to living and non-living.
11	Students respond to yes/no questions.
11	Students observe various technological devices.
Vague	
4	matches identifiable pictures of the sun, the moon, and Earth (e.g., a picture of the solar system, individual images)
4	Observes and/or participates in discussions about the difference between rotation and

	revolution.
5	Differentiates between a scientific fact and opinion (e.g., "I think he is tall." vs. "He is 6'5.").

AA Items. A total of 217 science AA items were rated by content experts as academic or nonacademic. A little over half of the AA items ($n=127$, 59%) were rated academic with the remaining items rated as nonacademic ($n=87$, 40%) or too vague to judge ($n=3$, 1%) Of the 87 items rated nonacademic, 59 were rated as foundational and 81 were rated at the presymbolic or concrete level. The number of AA science AA items, the number rated academic, and the percentage rated academic for each grade level are reported in Table 10.

Table 10: Number and Percentage of Science AA Items Rated Academic

	N of Items	Academic	%
5 th	70	41	58.6
8 th	63	38	60.3
11 th	84	48	57.1
All	217*	127	58.5

Note. *3 items did not provide enough information to rate.

A list of the items that were rated nonacademic is reported in Table 11.

Table 11: AA Items Rated Nonacademic

Grade	Item	Foundational	Presymbolic/ Concrete
5	Attend to a teacher led presentation showing that matter has weight	No	Yes
5	Uses his/her senses to respond to various weights (e.g., holds various weights, views similar objects with diff weights)	No	Yes
5	Attends to a teacher-led presentation about simple machines.	No	Yes
5	Attends to a teacher-led presentation showing how objects stop on different surfaces.	No	Yes
5	Participates in demonstrations involving moving objects and surfaces of varying resistance.	Yes	Yes
5	Observes stopping distances by experimenting with objects and different surfaces as provided by the teacher.	Yes	Yes
5	By experimenting with object on ramps, observes stopping distances.	Yes	Yes
5	Attends to a teacher-led presentation showing how a thermometer works.	No	Yes
5	Attends to a teacher-led presentation about the various colors in the spectrum of light.	No	Yes
5	Selects colors found in the presented spectrum of light (e.g., a color wheel, a list of colors found in the classroom)	Yes	Yes
5	Attends to a presentation about illustrations of plants and how plants get food	No	Yes
5	Observes his/her teacher demonstrating how a plant absorbs water (e.g., put a celery stick in water with food coloring)	Yes	Yes
5	Experiments with watering a plant and not watering a plant	Yes	Yes
5	Experiments with fertilizing a plant and not fertilizing a plant	Yes	Yes
5	Attends to a teacher-led presentation of parents and their offspring	No	Yes

5	Attends to a teacher-led presentation about the parts of a plant.	No	Yes
5	Attends to a teacher-led presentation about how humans react to seasonal changes	No	Yes
5	Attends to a teacher-led presentation about living things relying on each other within an energy pyramid	No	Yes
5	Participates in establishing a set ecosystem (e.g., an aquarium, a terrarium)	Yes	Yes
5	Attends to a teacher-led presentation about Earth's crust.	No	Yes
5	Attends to a presentation about Earth's rotation causing day and night	No	Yes
5	Imitates a teacher's demonstration of Earth's rotation (e.g., spins a globe).	Yes	Yes
5	Attends to a presentation about problems in the environment and the effects of the problems (e.g., flood, drought)	No	Yes
5	Attends to a presentation about different types of transportation	No	Yes
5	Attends to problems within the environment (e.g., looks at other students engaged in problems: student missed the bus)	No	Yes
5	Attends to presentations about animals in their specific habitats (e.g., videos of animals interacting in their habitats)	No	Yes
5	By using his/her senses, explores wildlife (e.g., touches/smells wildlife at a petting zoo, a farm, a taxidermist)	Yes	Yes
5	Manipulates pictures of animals within the animal's specific habitat	Yes	No
8	Attends to a teacher-led presentation about facts and predictions in scientific investigations	No	Yes
8	Attends to the facts and predictions of an experiment/discussion from a scientific investigation	No	Yes
8	Attends to a demonstration as part of a scientific investigation	No	Yes
8	Attends to the predictions of an experiment/discussion from a scientific investigation	No	Yes
8	Attends to the predictions of an experiment/discussion from a scientific investigation	No	Yes
8	Attends to the facts and predictions of an experiment/discussion from a scientific investigation	No	Yes
8	Attends to teacher-led presentation about changes in matter	No	Yes
8	Attends to teacher-led presentations about the periodic table of the elements	No	Yes
8	By using his/her senses, explores and/or manipulates mixtures (e.g., touches dough, touches nuts and bolts)	Yes	Yes
8	Matches the color-coded categories on the periodic table of the elements (e.g., metals are blue, nonmetals are green)	Yes	No
8	Participates in lab activities explaining why matter changes (e.g., boiling an object such as an egg- chemical, boiling water- physical)	Yes	Yes
8	Locates the first 18 elements by utilizing a word bank and the periodic table of elements	Yes	No
8	Attends to a presentation about the effects of currents and waves in the ocean.	No	Yes
8	Attends to a teacher-led presentation about the factors that cause weather	No	Yes
8	Attends to a teacher-led presentation about the major tectonic plates.	No	Yes
8	By using media, observes currents and waves	Yes	Yes
8	Observes weather maps of wind currents that include temperatures and other factors	Yes	Yes
8	Participates in activities/demonstrations involving different plates of Earth	No	Yes
8	Participates in experiments involving factors that affect the weather (e.g., warm/cold currents)	No	Yes
8	Attends to teacher-led presentation about the solar system	No	Yes
8	Responds to the effects of direct and indirect heat	.	Yes
8	Attends to a guest speaker as the speaker explains how scientific discoveries have met the needs of society (e.g., a utility spokesman, a wildlife specialist, a construction leader).	No	Yes

8	Attends to a teacher-led presentation of how science has been influenced by social needs	No	Yes
8	Attends to a teacher-led presentation about how science has led to technological advances in the local environment	No	Yes
8	Attends to a teacher-led presentation about problems created by humans in the local environment	No	Yes
11	Attends to stories of scientific discoveries	No	Yes
11	Attends to a teacher-led presentation about stating and testing a hypothesis	No	Yes
11	Attends to a multimedia presentation about safe practices with laboratory equipment	No	Yes
11	Attends to a teacher-led presentation about a simple scientific experiment	No	Yes
11	Recognizes lab safety equipment (e.g., lab aprons, safety goggles).	Yes	Yes
11	Practices safe and effective laboratory techniques with an emphasis on safety (e.g., wears safety glasses, disposal of materials)	Yes	Yes
11	Attends to various activities involving different chemical and/or physical reactions	No	Yes
11	Observes sensory changes of chemical and physical reactions (e.g., changes in color, changes in smell, changes in taste).	Yes	Yes
11	Manipulates different types of matter to recognize that matter cannot be destroyed (e.g., cuts an apple into pieces, rips paper)	Yes	Yes
11	By using a model, matches elements with symbols from the periodic table of the elements (e.g., matches Au to Au)	Yes	No
11	Attends to a teacher-led presentation relating energy to work	No	Yes
11	Attends to a presentation about speed and distance	No	Yes
11	Attends to a teacher-led presentation about forms of energy and their effects	No	Yes
11	Attends to a teacher-led presentation about simple circuits	No	Yes
11	Participates in activities involving charges (e.g., rubs a balloon on her/his head for a charge, observes static electricity, uses an electroscope).	No	Yes
11	By using his/her senses, explores different types of plants and animals (e.g., holds different types of stuffed animals, listens to different animal sounds, looks at different plants)	No	Yes
11	Attends to a presentation about the structures and functions of different plants/animals	No	Yes
11	Attends to and/or participates in activities involving the concept of cells (e.g., building blocks as cells to create something bigger)	No	Yes
11	Attends to multimedia or movies about prehistoric times (e.g., dinosaurs).	No	Yes
11	Attends to a teacher-led presentation about food chains	No	Yes
11	Attends to various activities involving food chains.	No	Yes
11	Uses a switch to activate programs on a computer about food chains.	No	Yes
11	Observes a food chain by attending a field trip (e.g., a zoo, a museum, a national park, etc.).	Yes	Yes
11	Attends to a teacher/multimedia presentation about cycles between living and non-living systems	No	Yes
11	Attends to a teacher/multimedia presentation about changes in the environment due to human activity.	No	Yes
11	Explains the concept of contamination (e.g., tastes water from a glass and compares it to tasting salt water).	Yes	No
11	Conducts an experiment involving a type of cycle (e.g., observes a glass of ice water-evaporation and condensation)	Yes	Yes
11	Attends to a teacher/multimedia presentation about the motion of planets (e.g., rotation and revolution).	No	Yes
11	Participates in activities involving rotation (e.g., a merry-go-round, a pinwheel).	Yes	Yes
11	Participates in sensory activities involving rotation and revolution (e.g., round objects, textured objects that are either rough or smooth, a pinwheel).	Yes	Yes

11	Attends to a teacher/multimedia presentation about the impact of science on individual people and the community.	No	Yes
11	Responds to true/false questions relating to science (e.g., "Your hair is black.", "Lunch is at 5: P.M.").	No	No
11	Attends to a teacher/multimedia presentation about technological limits.	No	Yes
11	Attends to a teacher/multimedia presentation about the benefits of recycling.	No	Yes

The alignment of AA items to national standards is reported in Table 12. Items were distributed across all strands except History and Nature of Science.

Table 12: Number and Percentage of AA Items in National Math Strands

Components	AA Items		Alternate Standards	
	N	%	N	%
Science as Inquiry	7	5.5	46	12.3
Physical Science	32	25.2	113	30.3
Life Science	25	19.7	85	22.8
Earth & Space Science	27	21.3	66	17.7
Science & Technology	16	12.6	27	7.2
Personal and Social Perspectives	20	15.7	36	9.7
History and Nature Science				

Summary:

For reading and math, almost all of the alternate content standards and alternate achievement standards were rated academic. Most (87%) of the reading items were rated academic, and those that were not were either foundational or presymbolic/concrete items. Almost all of the math items (98%) were rated academic.

All but one of the science alternate content standards were rated academic and 87% of the alternate achievement standards were rated academic. Science items appeared more problematic. A little over half (59%) of the AA items were rated academic. This may be related to the student expectations (e.g., passive expectations in items such as attending to a presentation or watching a glass for evaporation), language that may be vague or undeterminable as to what the students would actually be doing (e.g., participating in an activity), or misconceptions (e.g., salt water is not contaminated water). Most of the nonacademic items were rated at a presymbolic/concrete symbolic level of communication.

Criterion 2: The content is referenced to the student’s assigned grade level (based on chronological age).

All Alternate Content Standards and AA items across all the grades were referenced to grade level standards. In ELA, all Alternate Content Standards and AA items were referenced to

the content domain of reading; none of the Alternate Content Standards or AA items were referenced to Writing or Listening, Viewing, and Speaking. In Math, Alternate Content Standards and AA items for all grades were referenced to the content domains of (a) algebra, (b) geometry, (c) measurement, (d) number sense, and (e) statistics and probability. For science, Alternate Content Standards and AA items were referenced to all five of the content domains. The following Table lists the domains and content indicators, number of items, and number of supporting evidence required.

Table 13: Alternate Assessment Domains, Content Indicators, Number of Items and Supporting Evidence Per Indicator

Domain	Content Indicators	Items	Evidence
Reading (3rd-8th, 11th)	Reading Comprehension Strategies	7	1
	Response to Literacy	7	1
	Reading of Diverse Works, Cultures, and Time Periods	7	1
	Reading Informational Text	7	1
Mathematics (3rd-8th, 11th)	Algebra	21	1
	Geometry	14	1
	Number Sense	21	1
	Measurement	7	1
Science (5 th , 8 th , 11 th)	Statistics and Probability	14	1
	Nature of Science (8 th & 11 th only)	14	1
	Physical Science	21	1
	Life Science (5 th and 11 th only)	21	1
	Earth/Space Science	14	1
	Science, Technology, Environment, and Society	14	1

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

Content experts rate content centrality on a three-point scale measuring the degree of alignment (none, far, near) between the grade level content standards, alternate content standards and AA items. Performance centrality concerns the level of expected performance of the standards. Ratings are made on a 3-point scale (none, some, all) based on the degree of match between the focal (i.e., Alternate Content Standards and AA Items) and reference components (Grade level content standards). For example, an alternate content standard of “identify” would have some of the same performance as “analyze and identify.” Items that are rated as nonacademic are excluded from this analysis.

Reading

Alternate Content Standards content and performance centrality ratings are reported in Table 14. All but two the standards were rated as having far or near content centrality. The reason for the lack of content centrality was a mismatch to the standard. The two Alternate Content Standards rated none were: (1) Students are able to state one purpose of a book; (2) Students are able to identify main idea and supporting details

Table 14: Content and Performance Centrality Ratings of Alternate Content Standards

	Content Centrality		Performance Centrality	
	<i>N</i>	%	<i>N</i>	%
None	2	4.2	None	2 4.2
Far	14	29.2	Partial	21 43.8
Near	32	66.7	Full	25 52.1

AA items content and performance centrality ratings are reported in Table 15. Thirty AA items were rated as having no content centrality. The reason for the lack of content centrality was mismatch to standard ($n=28$) or overstretching ($n=1$). Thirty of the AA items were rated as having none of the performance centrality.

Table 15: Content and Performance Centrality Ratings of AA Items

	Content Centrality		Performance Centrality	
	<i>N</i>	%	<i>N</i>	%
None	30	17.9	None	30 17.9
Far	65	38.7	Partial	68 40.5
Near	73	43.5	Full	70 41.7

Table 16: AA Items with Content Centrality Rating of None

Grade	AA Item	Reason
3	Initiates either verbal or nonverbal communication that requires a response	Mismatch
3	When given two pictures, identifies the one that shows a real person, object, or situation	Mismatch
3	By browsing multicultural reading material in a purposeful manner, shows an interest in reading.	Overstretch
3	Shares a personal holiday experience.	Mismatch
3	Attends to and responds to a variety of multicultural texts read aloud.	Mismatch
3	Indicates (e.g., body language, head shakes) likes or dislikes about a multicultural story	Mismatch
4	Recognizes top to bottom and left to right (e.g., holds book upright, follows text with finger or eyes)	Mismatch
4	When given a prompt, indicates the correct picture.	Mismatch
4	While a story is read aloud, attends to the pictures	Mismatch
4	Follows directions in pictorial format	Mismatch
4	By using prior knowledge or pictures, makes predictions about a book	Mismatch
5	Categorizes objects according to one attribute	Mismatch
5	Categorizes objects according to two attributes (e.g., color, size, function traits)	Mismatch
5	Uses a timeline to indicate daily routine	Mismatch
5	Uses a calendar to indicate past, present, and future	Mismatch
5	By looking at photos of family members and connecting them to various time periods, identifies past, present, and future	Mismatch

5	Reads the prices of meals on various menus	Mismatch
5	Indicates that a library is a source of information	Mismatch
6	With teacher assistance, identifies an unknown word	Mismatch
6	Indicates the words before and after an unknown word	Mismatch
6	Identifies/states the definition of the term "main idea"	Mismatch
6	Uses a timeline to indicate daily routine	Mismatch
6	Identifies images or models as being from the past, present, or future	Mismatch
7	Locates a desired picture symbol on a map in a nonfiction book	Mismatch
8	Matches letter to letter	Mismatch
8	Attends/responds to a story on a computer or a tape	Mismatch
8	Identifies pictures/objects of common symbols (e.g., a flag, a church, a heart, etc.)	Mismatch
8	Attends/responds to a computer-assisted reading program	Other
8	identifies (verbal or nonverbally) survival signs in the environment (e.g., stop, restroom, exit, caution)	Mismatch
11	Recognize the beginning sounds of words	Mismatch

Supporting Evidence

Special educators reviewed the 566 pieces of student evidence in 142 AAs. Results of the special educators' ratings are reported in Table 17; 83% of the evidence had some or all of the evidence linked to the AA item; the remaining evidence was either rated as not linked (7%), too vague to rate (6%), no evidence provided (2%), missing data (2%), or unable to view video (<1%).

Table 17: Number and Percentage of Supporting Evidence that was Linked AA Items

Rating	<i>N</i>	%
No evidence provided	11	1.9
None of the evidence links	42	7.4
Some of the evidence links	64	11.3
All of the evidence links	404	71.4
Too vague	32	5.7
Missing data	9	1.6
Unable to view video	4	.7

Math Results

Alternate Content Standards content and performance centrality ratings are reported in Table 18. All of the standards were rated as having far or near content centrality.

Table 18: Alternate Content Standards

	<u>Content Centrality</u>		<u>Performance Centrality</u>	
	<u><i>N</i></u>	<u>%</u>	<u><i>N</i></u>	<u>%</u>
None			None	
Far	50	37.0	Partial	49 36.3
Near	85	63.0	Full	84 63.6

AA items content and performance centrality ratings are reported in Table 19. Twenty AA items were rated as having no content centrality. The reason for the lack of content centrality was backmapping ($n=4$), mismatch to standard ($n=5$), or overstretching ($n=10$). Eighteen of the AA items were rated as having none of the performance centrality.

Table 19: AA Items

	Content Centrality		Performance Centrality		
	<i>N</i>	%		<i>N</i>	%
None	20	3.6	None	18	3.3
Far	414	75.4	Partial	411	75.0
Near	115	20.9	Full	116	21.2

The following Table lists the AA items that were rated *none* for content centrality and the reason.

Table 20: AA Items with Content Centrality Rating of None

Grade	AA Item	Reason
3	Divides a given set of objects into equal groups	Mismatch
3	After completing a task, checks for accuracy using fact families with manipulatives and teacher support	Other
4	Identifies circles	Other
5	Recognizes that multiplication and division are opposite operations	Mismatch
6	Copies number patterns	Backmapping
6	Extends number patterns on a graph to solve problems	Backmapping
6	Distinguishes between increasing and decreasing patterns	Backmapping
6	When identifying a pattern, uses a letter to represent an unknown	Backmapping
7	Removes objects from a set of objects	Overstretch
7	Creates a whole unit from 2 halves (e.g., uses a 2-part puzzle)	Overstretch
7	Discriminates between a whole unit and half of that unit (e.g., a whole pie of half of a pie)	Overstretch
7	Discriminates between a whole unit and a quarter of that unit (e.g., a whole pie of 1/4 of a pie)	Overstretch
7	Uses words such as "all/half."	Overstretch
8	Copies number patterns	Overstretch
8	Names the days in a week or the months in a year	Overstretch
8	Demonstrates the relationship of days to a week	Overstretch
8	Demonstrates the relationship of months to a year	Overstretch
8	Counts backward from 10	Overstretch
11	Converts a number written scientific notation to a number with an exponent	Mismatch
11	Adds to a collection of items	Mismatch

Supporting Evidence

Special educators reviewed the 705 pieces of student evidence in 142 AAs. Results of the special educators' ratings are reported in Table 21; 82% of the evidence had some or all of the evidence linked to the AA items; the remaining evidence was either rated as not linked (12%),

too vague to rate (3%), no evidence provided (<1%), missing data (2%), or unable to view video (<1%).

Table 21: Number and Percentage of Student Evidence that was Linked to AA Items

Rating	<i>N</i>	%
No evidence provided	3	.4
None of the evidence links	83	11.8
Some of the evidence links	93	13.2
All of the evidence links	485	68.8
Too vague	22	3.1
Missing data	14	2.0
Unable to view video	5	.7

Science Results

Alternate Content Standards content and performance centrality ratings are reported in Table 22. All of the standards were rated as having far or near content centrality and partial or full performance centrality.

Table 22: Alternate Content Standards

	<u>Content Centrality</u>		<u>Performance Centrality</u>		
	<u><i>N</i></u>	<u>%</u>		<u><i>N</i></u>	<u>%</u>
None			None		
Far	1	1.7	Partial	1	1.7
Near	58	98.3	Full	58	98.3

AA items content and performance centrality ratings are reported in Table 23. Only one AA item was rated as having no content centrality. The reason for the lack of content centrality was overstretching. The item was “identifies different wildlife habitats (e.g. a desert, a prairie, a forest, a rainforest, a pond).” Only one of the AA items was rated as having none of the performance centrality.

Table 23: AA Items

	<u>Content Centrality</u>		<u>Performance Centrality</u>		
	<u><i>N</i></u>	<u>%</u>		<u><i>N</i></u>	<u>%</u>
None	1	.8	None	1	.8
Far	16	12.6	Partial	31	24.4
Near	110	86.6	Full	95	74.8

Approximately 64% of the student evidence was rated as linked (some or all) to AA items; the remaining evidence was either rated as not linked (7%), too vague to rate (6%), no evidence provided (4%), missing data (17%), or unable to view video (<2%).

Table 24: Number and Percentage of Student Evidence Linked to AA Items

Rating	<i>N</i>	%
No evidence provided	13	4.2
None of the evidence links	22	7.1
Some of the evidence links	19	6.1
All of the evidence links	180	58.3
Too vague	18	5.8
Missing data	52	16.8
Unable to view video	5	1.6

Summary

In reading, all but two of the alternate content standards were rated as having a far or near content and a partial or full performance centrality. Most of the reading items were rated as having a far or near content (82%) and performance centrality (82%). Supporting evidence for reading had similar percentages of alignment to standards. Approximately 83% of all the supporting evidence had some or all link to the AA items.

In math, all the alternate content standards were rated as having a far or near content and a partial or full performance centrality. Most of the math items had a far or near content centrality (96%) or partial or full performance centrality (96%). Approximately 82% of all the supporting evidence had some or all link to AA items.

In science, all the alternate content standards were rated as having a far or near content and a partial or full performance centrality. All but one of the science items were rated as having a far or near content (99%) and a partial or full performance centrality (99%). Approximately 64% of the science supporting evidence was rated as linked to AA items.

Criterion 4: The content differs from grade level in range, balance, and DOK, but matches high expectations set for students with significant cognitive disabilities

Alignment indices for criterion 4 are based on Webb’s alignment method (Webb, 1997). The following analyses include only AA items that were rated academic and had a content centrality rating of near or far. Four alignment indices are calculated: (a) categorical concurrence, (b) depth of knowledge, (c) range-of-knowledge, and (d) balance of representation. Below is a short description.

Categorical concurrence is the consistency of categories of content in the standards and assessments. The criterion of categorical concurrence between standards and assessment is met if the same or consistent categories of content appear in both the assessment and the standards. For example, if a content standard (or stand) is *measurement* in math, the

evaluative decision is “Does the assessment have items that target *measurement*?” It is possible for an assessment item to align to more than one content standard. For example, if an assessment item requires students to calculate surface area, which is aligned to the content standard of *measurement*, to successfully answer the question the student needs to be able to multiply numbers, which is aligned to the content standard of *operations*. In this case the item is aligned to both content standards.

Depth of knowledge (DOK) examines the consistency between the cognitive demands of the standards and cognitive demands of assessments (Webb, 1997). Important aspects of learning go beyond academic topics and include students’ organization of knowledge, problem representations, use of strategies, and self-monitoring skills (Glaser, Linn, & Bohrnstedt, 1997). Completely aligned standards and assessments requires an assessment system designed to measure in some way the full range of cognitive complexity within each specified content standard. Rated on a 6-point scale, DOK¹ provides a measure of performance complexity required to perform the skill listed in the standard or item. DOK ratings are guided by a list of verbs (e.g., identify, state) that reflect the response that would be required of the student. Experts consider the verb in conjunction with the content when determining DOK.

Range-of-knowledge correspondence criterion examines the alignment of assessment items to the multiple objectives within the content standards. Range-of-knowledge correspondence is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer assessment items. The range-of-knowledge numeric value is the percentage of content standards with at least 50% of the objectives having one or more hits. For example, if there are five objectives (e.g., length, area, volume, telling time, and mass) included in the content standard of measurement, a minimum expectation is at least one assessment item is related to at least three of the objectives. According to Webb (1997), 40-50% of the objectives for a standard could be considered weakly met.

The *balance of representation* criterion is used to indicate the extent to which items are evenly distributed across the content standards and the objectives under the content standards. In practice educational agencies may place greater emphasis on specific objectives and content standards. In this case the assumption of an even distribution would be replaced with the expected proportion, or emphasis, as specified by the educational agency. The formula used to compute the balance of representation index is the following

$$Balance = 1 - \left(\sum_{i=1}^O \left| \frac{1}{O} - \frac{I_k}{H} \right| \right) / 2,$$

where O is the total number of objectives hit (i.e., item has been judged to be aligned) for the content standard, I_k is the number of items hit corresponding to objective k , and H is the total number of items hit for the content standard. The balance index can range from 0

¹ There are many scales for determining DOK levels. The one provided in this manual is just one of many rating scales available.

(indicating unbalanced representation) to 1.0 (indicating balance representation) with values from .6 to .7 considered a weak acceptable balance and values .7 or greater considered acceptable.

Reading Results

Categorical Concurrence

The following analyses are based on the old 2004 reading standards. The following Table reports the grade level, content indicators, number of AA items that aligned (i.e., rated as academic and rated a content centrality of *far* or *near*), and number of AA items that were referenced to the content indicator. Across all grades, the AA was designed to have 100% categorical concurrence.

Table 25: Number of AA Items across Grade Level and Content Indicator

Grade	Indicator	Aligned	Intended
		<i>N</i>	<i>N</i>
3rd	apply various reading strategies to comprehend and interpret text	4	7
	evaluate text structures, literary elements, and literary devices within various genres...	4	7
	interpret and respond to diverse works from various cultures and time periods	2	7
	retrieve, analyze, synthesize, and evaluate a variety of informational texts.	6	7
4th	apply various reading strategies to comprehend and interpret text	4	7
	evaluate text structures, literary elements, and literary devices within various genres...	5	6
	interpret and respond to diverse works from various cultures and time periods	4	7
	retrieve, analyze, synthesize, and evaluate a variety of informational texts.	3	7
5th	apply various reading strategies to comprehend and interpret text	6	7
	evaluate text structures, literary elements, and literary devices within various genres...	5	7
	interpret and respond to diverse works from various cultures and time periods	2	7
	retrieve, analyze, synthesize, and evaluate a variety of informational texts.	3	7
6th	apply various reading strategies to comprehend and interpret text	4	7
	evaluate text structures, literary elements, and literary devices within various genres...	5	7
	interpret and respond to diverse works from various cultures and time periods	3	7
	retrieve, analyze, synthesize, and evaluate a variety of informational texts.	7	7
7th	apply various reading strategies to comprehend and interpret text	5	6
	evaluate text structures, literary elements, and literary devices within various genres...	6	7
	interpret and respond to diverse works from various cultures and time periods	5	6
	retrieve, analyze, synthesize, and evaluate a variety of informational texts.	6	7
8th	apply various reading strategies to comprehend and interpret text	5	7
	evaluate text structures, literary elements, and literary devices within various genres...	6	7
	interpret and respond to diverse works from various cultures and time periods	6	7
	retrieve, analyze, synthesize, and evaluate a variety of informational texts.	5	7
11th	apply various reading strategies to comprehend and interpret text	6	7
	evaluate text structures, literary elements, and literary devices within various genres...	7	7

interpret and respond to diverse works from various cultures and time periods	7	7
retrieve, analyze, synthesize, and evaluate a variety of informational texts.	7	7

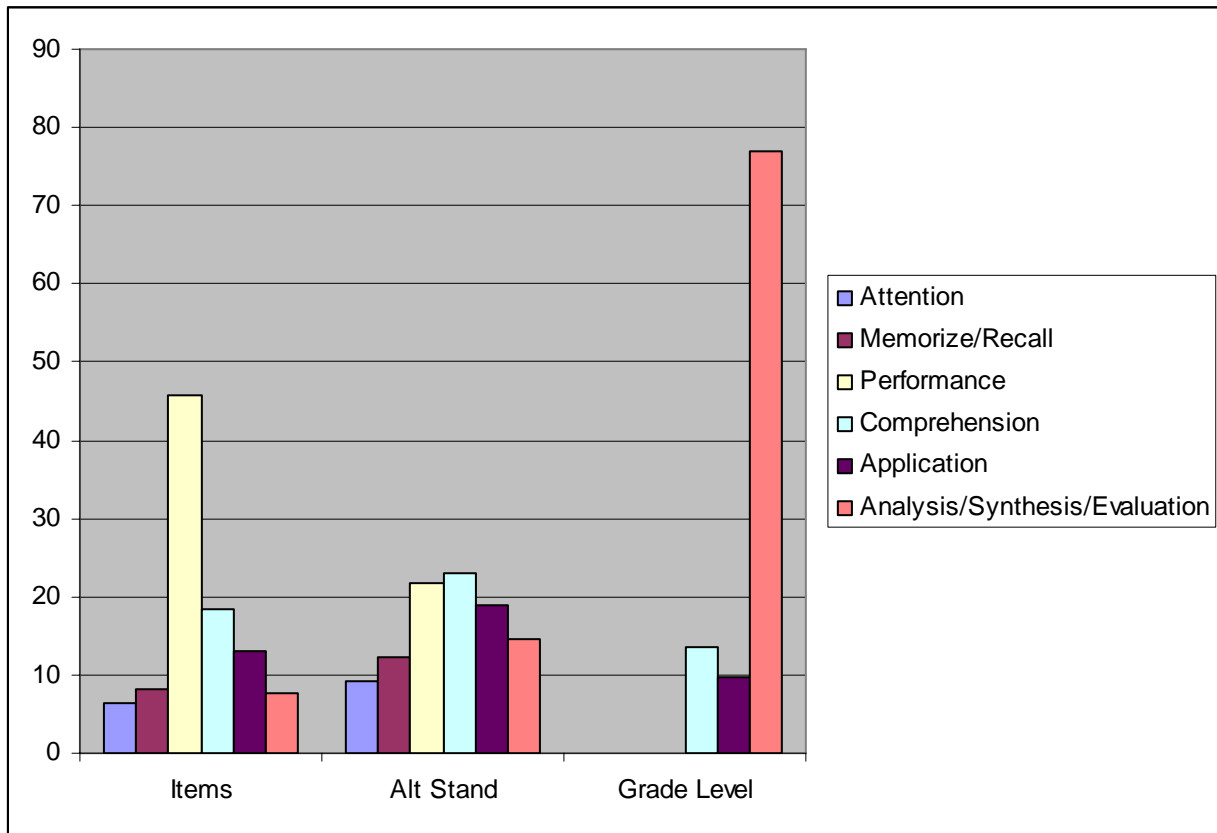
Depth of Knowledge

The following Table reports the depth of knowledge (based on a modified Bloom’s taxonomy) for all the AA items, Alternate Content Standards, and Grade Level Standards. There were AA items and Alternate Content Standards are all DOK levels.

Table 26: Depth of Knowledge of All AA Items, Alternate Content Standards, and Grade Level Standard

	<i>AA Items</i>		<i>Alternate Standards</i>		<i>Grade Level Standards</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Attention	11	6.5	22	9.2		
Memorize/Recall	14	8.3	29	12.2		
Performance	77	45.8	52	21.8		
Comprehension	31	18.5	55	23.1	7	13.5
Application	22	13.1	45	18.9	5	9.6
Analysis/Synthesis/Evaluation	13	7.7	35	14.7	40	76.9

Figure 1: Percentage of Depth of Knowledge of All AA Items, Alternate Content Standards, and Grade Level Standard



The following Table reports the depth of knowledge by grade level for AA items, Alternate Content Standards, and Grade Level Standards.

Table 27: Depth of Knowledge by Grade Level of AA Items, Alternate Content Standards, and Grade Level Standard

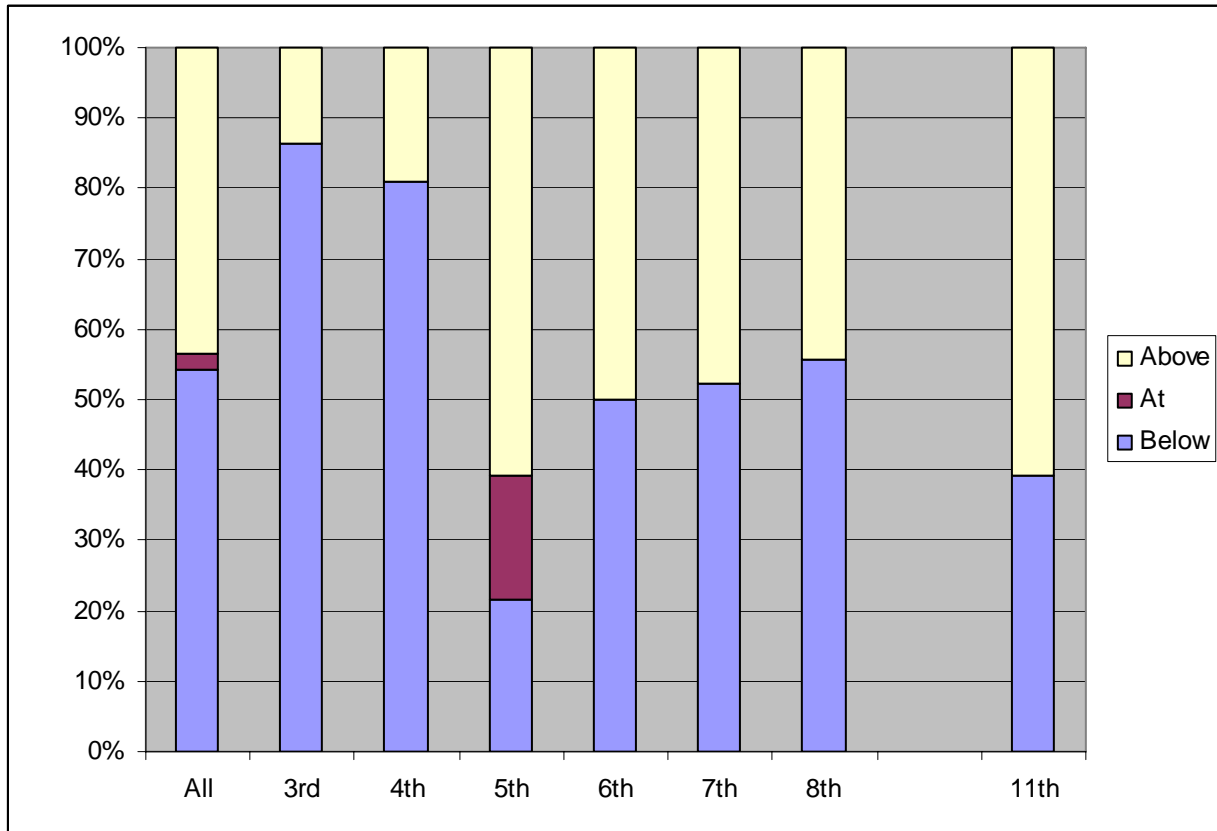
	AA Items		Alternate Standards		Grade Level Standards	
	N	%	N	%	N	%
3 rd						
Attention	6	27.3	1	2.6		
Memorize/Recall	3	13.6	6	15.4		
Performance	9	40.9	8	20.5		
Comprehension	1	4.5	6	15.4	1	12.5
Application	3	13.6	12	30.8	1	12.5
Analysis/Synthesis/Evaluation	6	27.3	6	15.4	6	75.0
4 th						
Attention	2	9.5	1	2.8		

Memorize/Recall	2	9.5	5	13.9		
Performance	12	57.1	4	11.1		
Comprehension	1	4.8	10	27.8	1	16.7
Application	2	9.5	5	13.9		
Analysis/Synthesis/Evaluation	2	9.5	11	30.6	5	83.3
5 th						
Attention			1	2.9		
Memorize/Recall			6	17.1		
Performance	12	52.2	9	25.7		
Comprehension	7	30.4	12	34.3	1	12.5
Application	4	17.4	6	17.1		
Analysis/Synthesis/Evaluation			1	2.9	7	87.5
6 th						
Attention	1	4.2	5	15.6		
Memorize/Recall	4	16.7	6	18.8		
Performance	9	37.5	2	6.3		
Comprehension	1	4.2	7	21.9	1	12.5
Application	4	16.7	7	21.9	1	12.5
Analysis/Synthesis/Evaluation	5	20.8	5	15.6	6	75.0
7 th						
Attention			4	11.4		
Memorize/Recall	1	4.3	2	5.7		
Performance	14	60.9	12	34.3		
Comprehension	5	21.7	11	31.4	1	12.5
Application	2	8.7	2	5.7	1	12.5
Analysis/Synthesis/Evaluation	1	4.3	4	11.4	6	75.0
8 th						
Attention	2	7.4	6	17.1		
Memorize/Recall	2	7.4	2	5.7		
Performance	10	37.0	10	28.6		
Comprehension	6	22.2	6	17.1	1	14.3
Application	4	14.8	10	28.6	1	14.3
Analysis/Synthesis/Evaluation	3	11.1	1	2.9	5	71.4
11 th						
Attention			4	15.4		
Memorize/Recall	2	7.1	2	7.7		
Performance	11	39.3	7	26.9		
Comprehension	10	35.7	3	11.5	1	14.3
Application	3	10.7	3	11.5	1	14.3
Analysis/Synthesis/Evaluation	2	7.1	7	26.9	5	71.4

Table 28: Percentage of AA Items Below, At, or Above Grade Level Content Standards DOK Level

Grade	Below	At	Above
All	54.2	2.4	43.5
3 rd	86.4	0	13.6
4 th	81.0		19.0
5 th	21.7	17.4	60.9
6 th	50.0		50.0
7 th	52.2		47.8
8 th	55.6		44.4
11 th	39.3		60.7

Figure 2: Percentage of AA Items Below, At, or Above Grade Level Depth of Knowledge



Range of Knowledge

The AA were designed to have at least one item for each grade level standards indicating 100% of all indicators have 50% of the grade level standards with at least one item.

Balance of Representation

The balance of representation values for each grade is reported in the following Table. The values ranged from .81 to .95 indicating that items were balanced across all the content indicators.

Table 29: Balance of Representation Across Grades

Grade	Reading
3 rd	.87
4 th	.93
5 th	.81
6 th	.87
7 th	.95
8 th	.95
11 th	.85

Math Results

Categorical Concurrence

The following Table reports the grade level, content indicators, number of AA items that aligned (i.e., rated as academic and rated a content centrality of *far* or *near*), and number of AA items that were referenced to the content strand. Across all grades, the AA was designed to have 100% categorical concurrence.

Table 30: Number of AA Items across Grade Level and Content Strand

Grade	Strand	Aligned	Intended
3 rd	Algebra	25	28
	Geometry	13	13
	Measurement	5	5
	Number Sense	21	21
	Statistics and Probability	12	14
4 th	Algebra	21	21
	Geometry	13	14
	Measurement	6	7
	Number Sense	21	21
	Statistics and Probability	13	14
5 th	Algebra	28	28
	Geometry	14	14
	Measurement	7	7
	Number Sense	20	21
	Statistics and Probability	17	19
6 th	Algebra	16	21

	Geometry	14	14
	Measurement	8	8
	Number Sense	21	21
	Statistics and Probability	13	14
7 th	Algebra	27	28
	Geometry	13	13
	Measurement	7	7
	Number Sense	16	20
	Statistics and Probability	13	14
8 th	Algebra	27	28
	Geometry	14	14
	Measurement	4	7
	Number Sense	19	19
	Statistics and Probability	11	14
11 th	Algebra	21	21
	Geometry	14	14
	Measurement	7	7
	Number Sense	20	21
	Statistics and Probability	13	14

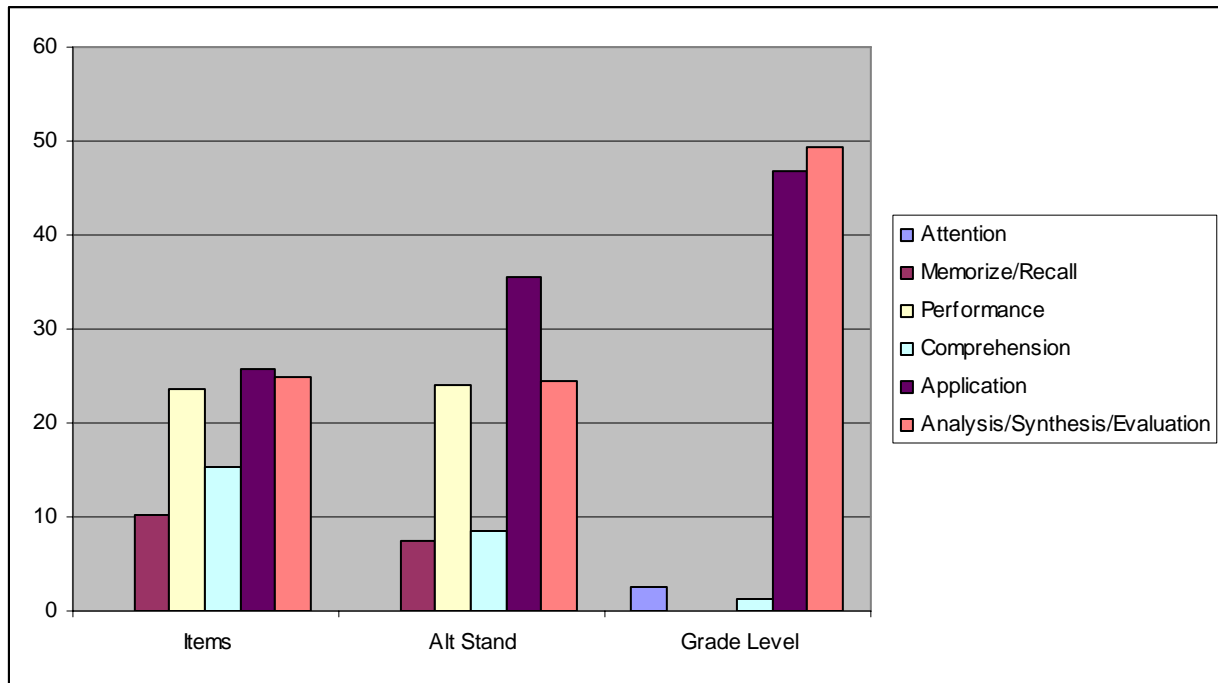
Depth of Knowledge

The following Table reports the depth of knowledge (based on a modified Bloom's taxonomy) for all the AA items, Alternate Content Standards, and Grade Level Standards. There were AA items and Alternate Content Standards are all DOK levels.

Table 31: Depth of Knowledge of All AA Items, Alternate Content Standards, and Grade Level Standard

	<i>AA Items</i>		<i>Alternate Standards</i>		<i>Grade Level Standards</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Attention					2	2.6
Memorize/Recall	57	10.3	50	7.4		
Performance	131	23.6	163	24.0		
Comprehension	85	15.3	58	8.5	1	1.3
Application	143	25.8	242	35.6	36	46.8
Analysis/Synthesis/Evaluation	139	25.0	166	24.4	38	49.4

Figure 3: Percentage of Depth of Knowledge of All AA Items, Alternate Content Standards, and Grade Level Standard



The following Table reports the depth of knowledge by grade level for AA items, Alternate Content Standards, and Grade Level Standards.

Table 32: Depth of Knowledge by Grade Level of AA Items, Alternate Content Standards, and Grade Level Standard

	AA Items		Alternate Standards		Grade Level Standards	
	N	%	N	%	N	%
3rd						
Attention						2
Memorize/Recall	12	15.4		21	19.3	
Performance	24	30.8		22	20.2	
Comprehension	7	9.0		3	2.8	1
Application	18	23.1		33	30.3	4
Analysis/Synthesis/Evaluation	17	21.8		30	27.5	5
4th						
Attention						
Memorize/Recall	11	14.7		18	15.4	
Performance	18	24.0		19	16.2	
Comprehension	8	10.7		10	8.5	
Application	10	13.3		34	29.1	6
Analysis/Synthesis/Evaluation	28	37.3		36	30.8	6

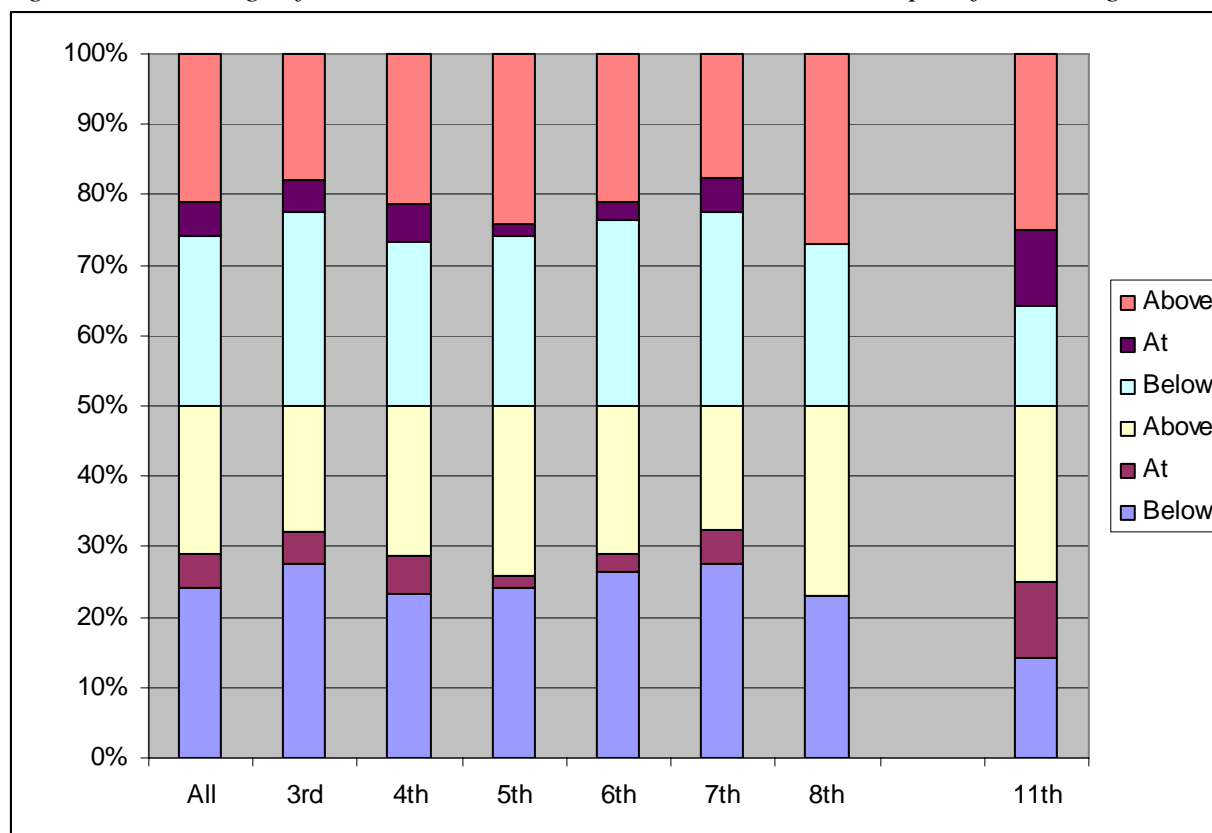
5 th					
Attention					
Memorize/Recall	9	10.3	4	3.7	
Performance	23	26.4	48	44.0	
Comprehension	17	19.5	4	3.7	
Application	19	21.8	31	28.4	6
Analysis/Synthesis/Evaluation	19	21.8	22	20.2	6
6 th					
Attention					
Memorize/Recall	12	15.8	2	2.5	
Performance	22	28.9	28	34.6	
Comprehension	10	13.2	13	16.0	
Application	21	27.6	23	28.4	6
Analysis/Synthesis/Evaluation	11	14.5	15	18.5	6
7 th					
Attention					
Memorize/Recall	8	9.8	1	1.1	
Performance	22	26.8	17	18.9	
Comprehension	18	22.0	9	10.0	
Application	18	22.0	48	53.3	6
Analysis/Synthesis/Evaluation	16	19.5	15	16.7	6
8 th					
Attention					
Memorize/Recall	4	5.0	18	24.0	
Performance	17	21.3	7	9.3	
Comprehension	9	11.3	30	40.0	
Application	29	36.3	20	26.7	7
Analysis/Synthesis/Evaluation	21	26.3			5
11 th					
Attention					
Memorize/Recall	1	1.3	4	4.1	
Performance	5	6.5	11	11.2	
Comprehension	16	20.8	12	12.2	
Application	28	36.4	43	43.9	1
Analysis/Synthesis/Evaluation	27	35.1	28	28.6	4

Table 33: Percentage of AA Items Below, At, or Above Grade Level Content Standards DOK Level

Grade	Below	At	Above
All	48.5	9.2	42.3
3 rd	55.1	9.0	35.9
4 th	46.7	10.7	42.7
5 th	48.3	3.4	48.3

6 th	52.6	5.3	42.1
7 th	54.9	9.8	35.4
8 th	46.3		53.8
11 th	35.1	27.3	62.3

Figure 4: Percentage of AA Items Below, At, or Above Grade Level Depth of Knowledge



Range of Knowledge

The AA were designed to have at least one item for each grade level standards indicating 100% of all strands have 50% of the grade level standards with at least one item.

Balance of Representation

The balance of representation values for each grade is reported in the following Table. The values ranged from .79 to .89 indicating that items were balanced across all the content indicators.

Table 34: Balance of Representation Across Grades

Grade	Math
3 rd	.79
4 th	.83
5 th	.84
6 th	.89
7 th	.83
8 th	.79
11 th	.85

Science Results

Categorical Concurrence

The following Table reports the grade level, content indicators, number of AA items that aligned (i.e., rated as academic and rated a content centrality of *far* or *near*), and number of AA items that were referenced to the content strand. Across all grades, the AA was designed to have 100% categorical concurrence.

Table 35: Number of AA Items across Grade Level and Content Strand

Grade	Strand	Aligned	Intended
5 th	Nature of Science	0	0
	Physical Science	11	21
	Life Science	12	21
	Earth/Space Science	10	14
	Science, Technology, Environmental, and Society	7	14
8 th	Nature of Science	8	14
	Physical Science	8	14
	Life Science	0	0
	Earth/Space Science	12	21
	Science, Technology, Environmental, and Society	10	14
11 th	Nature of Science	8	14
	Physical Science	12	21
	Life Science	13	21
	Earth/Space Science	6	14
	Science, Technology, Environmental, and Society	9	14

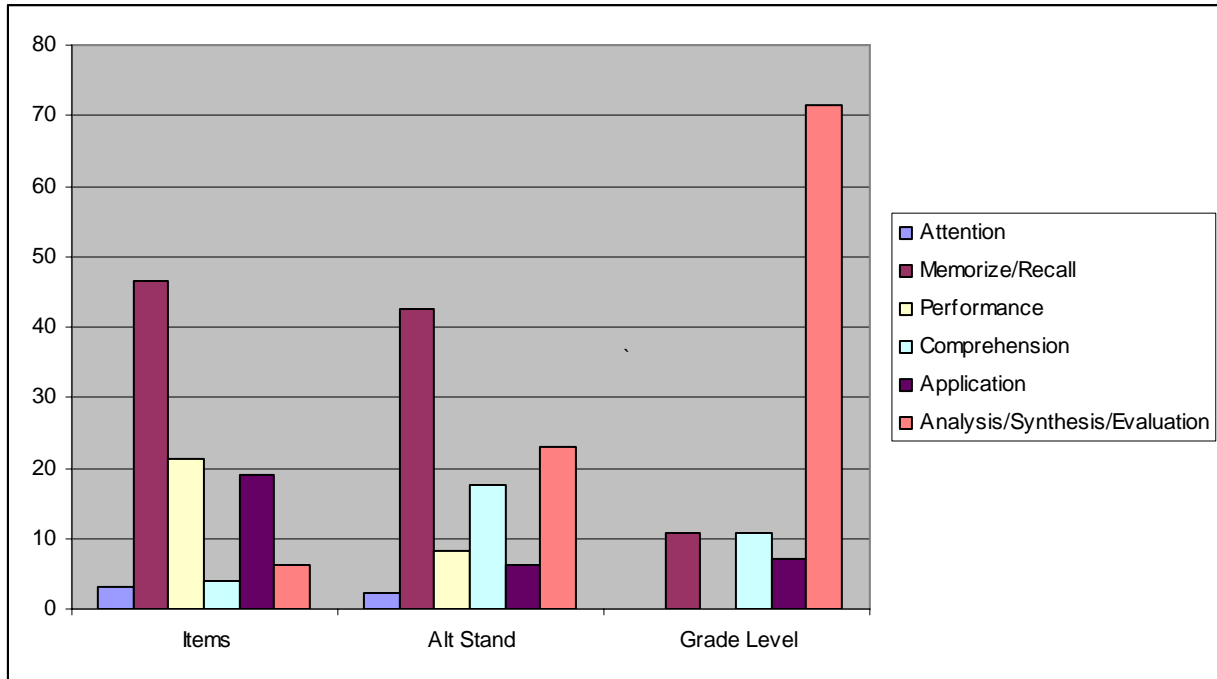
Depth of Knowledge

The following Table reports the depth of knowledge (based on a modified Bloom’s taxonomy) for all the AA items, Alternate Content Standards, and Grade Level Standards. There were AA items and Alternate Content Standards are all DOK levels.

Table 36: Depth of Knowledge of All AA Items, Alternate Content Standards, and Grade Level Standard

	AA Items		Alternate Standards		Grade Level Standards	
	N	%	N	%	N	%
Attention	4	3.1	8	2.2		
Memorize/Recall	59	46.5	158	42.5	3	10.7
Performance	27	21.3	31	8.3		
Comprehension	5	3.9	66	17.7	3	10.7
Application	24	18.9	23	6.2	2	7.1
Analysis/Synthesis/Evaluation	8	6.3	86	23.1	20	71.4

Figure 5: Percentage of Depth of Knowledge of All AA Items, Alternate Content Standards, and Grade Level Standard



The following Table reports the depth of knowledge by grade level for AA items, Alternate Content Standards, and Grade Level Standards.

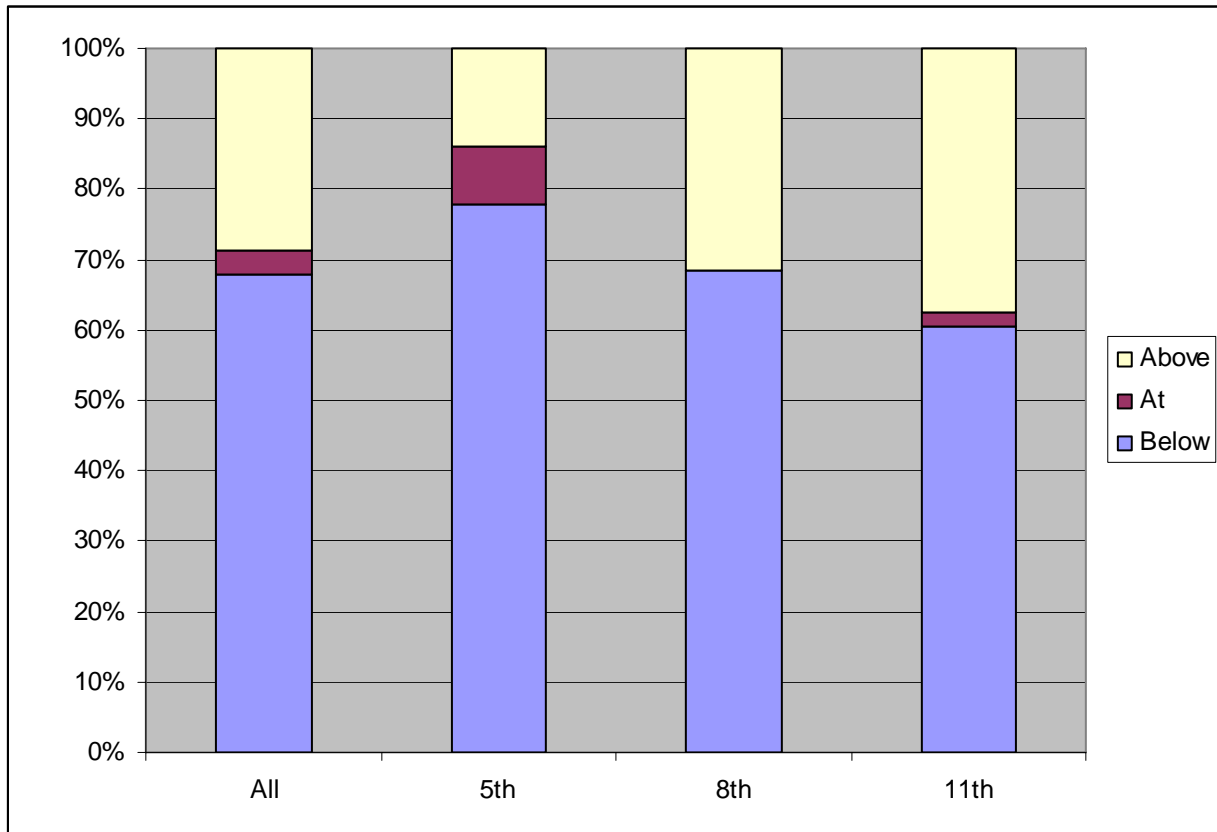
Table 37: Depth of Knowledge by Grade Level of AA Items, Alternate Content Standards, and Grade Level Standard

	<i>AA Items</i>		<i>Alternate Standards</i>		<i>Grade Level Standards</i>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
<i>5th</i>						
Attention	2	4.9	1	0.9		
Memorize/Recall	27	65.9	61	57.5	1	9.1
Performance	6	14.6	9	8.5		
Comprehension	2	4.9	12	11.3	1	9.1
Application	3	7.3	3	2.8	1	9.1
Analysis/Synthesis/Evaluation	1	2.4	20	18.9	8	72.7
<i>8th</i>						
Attention			1	1.2		
Memorize/Recall	17	44.7	35	41.2	1	14.3
Performance	9	23.7	4	4.7		
Comprehension			12	14.1	1	14.3
Application	9	23.7	7	8.2	1	14.3
Analysis/Synthesis/Evaluation	3	7.9	26	30.6	4	57.1
<i>11th</i>						
Attention	2	4.2	6	3.3		
Memorize/Recall	15	31.3	62	34.3	1	10.0
Performance	12	25.0	18	9.9		
Comprehension	3	6.3	42	23.2	1	10.0
Application	12	25.0	13	7.2		
Analysis/Synthesis/Evaluation	4	8.3	40	22.1	8	80.0

Table 38: Percentage of AA Items Below, At, or Above Grade Level Content Standards DOK Level

<i>Grade</i>	<i>Below</i>	<i>At</i>	<i>Above</i>
All	68.0	3.3	28.7
5 th	77.8	8.3	13.9
8 th	68.4		31.6
11 th	60.4	2.1	37.5

Figure 6: Percentage of AA Items Below, At, or Above Grade Level Depth of Knowledge



Range of Knowledge

The AA were designed to have at least one item for each grade level standards indicating 100% of all strands have 50% of the grade level standards with at least one item.

Balance of Representation

The balance of representation values for each grade is reported in the following Table. The values ranged from .80 to .88 indicating that items were balanced across all the content indicators.

Table 39: Balance of Representation Across Grades

Grade	Science
5 th	.85
8 th	.80
11 th	.88

Summary

The alternate assessment across all academic domains had acceptable categorical concurrence, depth of knowledge, range-of-knowledge, and balance of representation. There were items distributed across all levels of depth of knowledge. The range-of-knowledge and balance of representation suggested the alternate assessments had exceptional content coverage.

Criterion 5: There is some differentiation in content across grade levels.

One way to examine the change in content across the grade levels demonstrated in the alternate assessment items is to use Webb's (2005) definitions for change across content. Those definitions are:

- (a) *broader*—higher-grade standards or items reflect broader application of target skill or knowledge;
- (b) *deeper*—higher-grade standards or items reflect deeper mastery of the target skill or knowledge;
- (c) *prerequisite*—lower-grade standards or items reflects a different by prerequisite skill for mastery of the higher grade standard;
- (d) *new*—the higher-grade has a new skill or knowledge unrelated to skills or knowledge covered at prior grades; and
- (e) *identical*—higher-grade standards or items appear identical to one of the lower-grade standards.

Content experts were trained using these definitions and examples to review the AA items. Each strand within each content area was then rated using the definitions and a rating system (not evident-0%, limited- 25%, partial-50%, and clear- 75%) with experts noting information of particular interest and examples.

Reading Results

Results of the differentiation across grade levels for the reading alternate standards and AA items are reported in Tables 40 and 41. For the reading alternate standards, all strands of reading demonstrated very strong differentiation across the grade levels.

Table 40: Reading Alternate Standards

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
Reading	Clear relationship for	None.	Experts noted that there was “virtually	Very Strong Differentiation.

	broader and deeper applications, prerequisite and new skills.		no repetition from one grade to the next. The early grades build basic skills while subsequent grades build analytical skills.”	
--	---	--	---	--

Items aligned to the first three indicators of reading demonstrated strong differentiation across grade levels. Items aligned to the four indicators (i.e., Retrieve, Analyze, Synthesize, and Evaluate a Variety of Informational Texts) were rated as having partial evidence of deeper application and new skills and limited evidence of broader application of knowledge and skills.

Table 41: Reading AA Items

Indicator	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
Interpret and Respond to Diverse Works	Clear evidence of broader and deeper applications, prerequisite, and new knowledge/skills.	None.		Very Strong Differentiation.
Evaluate Text Structures, Literary Elements, and Literary Devices	Clear evidence of broader and deeper applications, prerequisite, and new knowledge/skills.	None.		Very Strong Differentiation.
Apply Various Reading Strategies to Comprehend	Clear evidence of broader and deeper applications, prerequisite, and new knowledge/skills.	None.	The interaction with text becomes more complex and progresses from decoding text to making meaning out of text. The experts also noted that the majority of items at each grade level were new.	Very Strong Differentiation.

Retrieve, Analyze, Synthesize, and Evaluate a Variety of Informational Texts	Clear evidence of the relationship of prerequisite skills. Partial evidence was noted for deeper applications and new skills. Limited evidence was noted for broader applications of knowledge/skills	Partial.		Needs Improvement with a continued focus on reducing number of repeating items across grades.
--	---	----------	--	---

Math Results

Results of the differentiation across grade levels for the math alternate standards and AA items are reported in Tables 42 and 43. For the math alternate standards, all strands of math demonstrated very strong differentiation across the grade levels except for Statistics and Probability, which demonstrated limited evidence of new and broader skills being introduced in the higher grade levels.

Table 42: Math Alternate Standards

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 vertical relationships for broader applications, deeper applications, and prerequisite skills. Limited evidence of new skills.	Limited.		Strong Differentiation.
Geometry	Clear relationship for deeper applications and	Limited.		Strong Differentiation.

	prerequisite skills. Partial evidence of broader applications and new skills.			
Statistics and Probability	Clear evidence of deeper applications and prerequisite skills. Limited evidence of new skills and broader applications.	Very few/None.	Gather/organize data.	Adequate with a continued focus on including new skills and broader applications of knowledge.
Number Sense	Clear evidence of deeper applications and prerequisite skills. Broader applications and new skills were identified as having evidence of a partial relationship.	Limited.	Experts repeatedly noted in this strand that as math is a building block content area, the covering of prerequisite skills from earlier grades is a <i>good</i> thing. Example of identical item: identify and describe angles in grades 5 and 6.	Strong Differentiation.
Measurement	Clear evidence of broader and deeper applications. Partial evidence of new and prerequisite skills.	Limited.		Strong Differentiation.

Items in the strands of geometry and measurement were rated as having strong differentiation across grades. Items associated with the remaining strands tended to repeat across grade levels indicating a lack of differentiation.

Table 43: Math AA Items

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.
Statistics and Probability	Partial evidence of all.	Partial.	Experts also noted that several items repeated across each grade level.	Needs Improvement with a continued focus on reducing number of repeating items across grades.
Number Sense	Clear evidence of deeper applications. Broader applications and new skills were	Partial.	The experts noted that while there were five identical items from grade 3 to grade 4 for this strand, there were	Needs Improvement with a continued focus on reducing

	identified as having evidence of a partial relationship. Prerequisite skills had limited evidence.		four identical items from grades 3 and 4 to high school.	number of repeating items across grades.
Measurement	Clear evidence of new and broader applications. Partial evidence of deeper applications. Limited evidence of prerequisite skills.	Limited.		Strong Differentiation.

Science Results

Results of the differentiation across grade levels for the science alternate standards and AA items are reported in Tables 44 and 45. For the math alternate standards, all strands of science demonstrated very strong differentiation across the grade levels.

Table 44: Science Alternate Standards

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
Nature of Science	Clear for all.	None.		Very Strong Differentiation
Physical Science	Clear evidence for broader, prerequisite and new skills. Limited evidence for deeper knowledge.	None.		Strong Differentiation.
Life Science	Clear evidence of deeper, prerequisite and new skills. Partial	Limited.	Experts identified a limited amount of Life Science at the grade 8 which	Strong Differentiation.

	evidence for broader applications.		requires prerequisite learning at grade 5 without reinforcement until grade 11.	
Earth and Space Science	Clear vertical relationships with broader applications, prerequisite skills, and new skills. Partial evidence for deeper applications.	None.		Strong Differentiation.
Science, Technology, Environment, and Society	Clear vertical relationships with broader applications, prerequisite skills, and new skills. Partial evidence for deeper applications.	None.		Strong Differentiation.

Science AA items were rated as demonstrating a very strong differentiation across grade levels except for those items aligned to Life Science. The primary concern for the content experts was the lack of AA items in the strand of Life Science for the 8th graders.

Table 45: Science AA Items

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
Nature of Science	Clear for deeper applications, prerequisite, and new skills. Partial evidence of broader applications.	None.	This strand was not present in grade 5.	Very Strong Differentiation
Physical Science	Clear evidence for broader, deeper, prerequisite and	None.		Very Strong Differentiation.

	new skills.			
Life Science	Limited relationship for broader and deeper applications of knowledge/skills across the grades as well as for new knowledge/skills. Partial evidence was found for prerequisite skills.	Partial.	As this strand was missing from the grade 8 alternate standards, it was also missing from the alternate assessment for grade 8 which troubled the experts in that the expectation of retention of information from grade 5 to grade 11 may be even more challenging for this population.	Needs Improvement.
Earth and Space Science	Clear vertical relationships across the grade levels for broader and deeper applications, prerequisite, and new knowledge/skills.	None.		Very Strong Differentiation.
Science, Technology, Environment, and Society	Clear evidence of a relationship across the grade levels for broader and deeper applications and new knowledge/skills. Partial evidence was found for prerequisite skills.	None.	Experts did wonder about the inclusion of the topic of habitat in this strand and recommended that the link from grade 5 to grade 8 be stronger for prerequisite skills.	Very Strong Differentiation.

Age Appropriateness of Supporting Evidence

A total of 1578 pieces of support evidence across 142 student AAs were examined for age appropriateness. Approximately half (54%) of the supporting evidence was age neutral, 23% was age appropriate, and 12% was not age appropriate.

Table 46: 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

Summary:

Overall the majority of strands within the adapted standards and AA items were rated as strong to very strong for differentiation. Very few strands were noted as needing improvement by reducing the number of repeating items across grades. Most of the supporting evidence submitted by the teachers was rated age appropriate (23%) or age neutral (54%).

Criterion 6: The expected achievement for students is for the students to show learning of grade referenced academic content.

Since inferences from scores on alternate assessments are to be made about what students know or can do in an academic domain, the achievement standards should consider those indicators which provide information about student performance rather than system or teacher performance. The strongest inference can be made that the student learned the content if: (a) there is evidence the student did not already have the skill (e.g., through use of pretest, baseline or previous year’s learning), (b) the skill is performed without teacher prompting, and (c) the skill is performed across materials/lessons to show mastery of the concept versus rote memory of one specific response. This inference can also be strengthened by considering the difficulty level of the skills performed. The inference is weakened if the student’s score is augmented by program level criteria. In contrast, these program level criteria are essential to the question of instructional alignment (Criterion 8). Evaluation of this criterion is done through special education review of state documents related to alternate achievement standards and application of a rubric.

Degree of Inference

Accuracy. The state had incorporated consideration of accuracy. Students receive more credit for higher levels of accuracy.

Level of independence. Students also receive more credit for performing skills with minimal prompting. Students also receive some credit for performing skills with more extensive prompting. Thus student performance scores may be influenced by teacher assistance.

New learning. This system does not include a baseline or pretest nor is it necessary. Instead, the introduction of new items across years (see criterion 5) supports inferences about new learning.

Complexity. The checklist has items at varying levels of complexity. All students within the AA-AAS are assessed on all items.

Generalization. The system requires students to perform skills across more than one setting at the progressing level. This increases the inference about student learning. In contrast, the level of conceptual generalization is less clear because the evidence only contains demonstration of the skill with on specific activity/material.

Summary

This system has the potential for high student inference concerning student learning academic content depending on the cutoff scores established in standard setting. If the level established for adequate yearly progress includes minimal levels of prompting, higher levels of accuracy, and generalization a strong inference could be made about student learning of grade referenced academic content. In contrast, for students who perform with extensive prompting at low levels of accuracy the inference would be weaker. One thought for future development of the system, some consideration might also be given to providing additional credit for student demonstration of conceptual generalization. For example, a student might learn to select a picture for “sad” to show how the main character felt after an event in a story. It might take the student many repetitions of the story to learn this response. The inference about student learning could be further strengthened if the student could then generalize identifying a main character as “sad” in a new story. This is a stronger inference about learning academic content than if the student could identify the main character as “sad” in the *same* story in a new setting (e.g., library).

<p>Criterion 7: The potential barriers to demonstrating what students know and can do are minimized in the assessment.</p>

Special education experts reviewed the alternate assessment and the accompanying administration manual to examine the extent to which a wide range of students to complete tasks within the assessment within the level of independence and accuracy expected by the state. Experts used four definitions to describe how students were able access the assessment items.

No provision: This type of student would not be able to demonstrate knowledge/skill on the assessment; needed supports are nonexistent or insufficient to help this type of student demonstrate learning.

Flexibility built into tasks: This type of student would be able to demonstrate knowledge/skill because of flexibility in administration. Flexibility is built into the items (e.g., teacher choice/design in portfolio, scaffolding in scripted performance events).

Accommodations: This type of student would be able to demonstrate knowledge/skill because of allowable accommodations. Accommodations are not built into items/tasks, but are described in the test administration materials and may be applied to this type of student. Accommodations do not change the construct being measured.

Modifications: This type of student would be able to demonstrate knowledge/skill because of modifications in assessment materials, administration procedures, etc. Modifications are not built into items/tasks, but are described in the test administration materials and may be applied to this type of student. Modifications do change the construct being measured.

Results

There is a great deal of flexibility with how items are interpreted and taught including examples of students with a variety of needs and supports (e.g., augmentative communication, pictures symbols, adapted books, etc.). Flexibility was designed for nonverbal students by responds using printed words, pictures, manual sings, and eye gaze. While experts did not identify specific information regarding allowable accommodations and adaptations for all students (e.g., students with vision and hearing needs), it was noted that the assessment was flexible to allow teachers to assess (and potentially provide instruction) based upon student needs.

The checklist and training materials provide specific information regarding data collection of the evidence and rating of the items in a standardized format. The limitation to standardization is in how the teacher interprets and provides instruction on the item itself. The materials used for instruction and assessment vary across students as found in the supporting evidence.

Symbolic Level of Supporting Evidence

Special educators reviewed the supporting evidence to rate if they were accessible to students with presymbolic and symbolic levels of communication. Of the 1496 items reviewed, 20% ($n=299$) were rated at the presymbolic level and 80% ($n=1197$) rated at the symbolic level.

Summary

The alternate assessment items were accessible to students at all levels of symbolic communication and many of the items allowed flexibility in response mode. Accommodations and adaptations are allowed for all students. As experts noted limited information on how students with hearing or visual impairments would have access, we recommend some information or examples on how to adapt items for sensory impairments to the administration manual for the alternate assessment.

Criterion 8: The instructional program promotes learning in the general curriculum.
--

Special education experts reviewed the alternate assessment rating scales, professional development/ training materials, and the alternate standards for evidence of overall program quality indicators. The Table below summarizes their findings.

Table 47: Program Quality Indicators

Does the alternate assessment and professional development promote:	
1. opportunities for instruction in general education classrooms for students with significant cognitive disabilities?	Minimal- While there was no overt evidence found, experts could find limited support for promoting inclusive opportunities in the rating scale evidence.
2. opportunities for instruction with typical peers for students with significant cognitive disabilities?	Minimal- While there was no overt evidence found, experts could find limited support for this indicator in the rating scale evidence.
3. opportunities for students with significant cognitive disabilities to make choices, problem solve, self-advocate, self-evaluate?	Minimal- There was limited evidence that this practice was occurring in the professional development but this evidence was NAAC materials (1 example from a NAAC powerpoint had this information but it was not emphasized).
4. the provision of assistive technology for students who need it?	Yes- This indicator was evident in the in all three sources (e.g., examples of access for 6 th grade math).
5. the access and use of typical classroom resources within instruction (e.g., science kits, grade level books, textbooks)?	Yes- Experts were able to find evidence for this in all three resources reviewed (e.g., the Giver in the Access and Alternate Content Standards powerpoint).
6. literacy being promoted across the content areas for students with significant cognitive disabilities (e.g., the pairing of text with picture symbols and objects)?	Yes- Experts were able to find some examples of the use of literacy skills in the alternate standards (e.g., life science- use of pictures, objects, and stories to answer questions about animals and their habitats).
7. the meaningful linking of academic skills in functional contexts?	Yes- All three sources provided evidence of this indicator (e.g., Access and Alternate Content Standards powerpoint)

Professional Development

Special education experts reviewed the professional development materials including the grade level math rollout materials, the alternate assessment rating scales and alternate standards, the Access and Alternate Content Standards powerpoint presentation, and a document title “Example of Access to the General Curriculum: 6th Grade Math”.

Alternate Assessment and Alternate Content Standards: The alternate content standards clearly provide the links to the state standards. The overarching state goals and indicators can be found for all tested grade levels for Math and Science. Only Reading was provided for English Language Arts (ELA). No other ELA strands were found (e.g., Writing, Listening, Viewing/Visually Representing). The general education content standards to which the alternate content standards are linked are also evident across domains except for ELA (only Reading). The Reading Alternate Content Standards do, however, address the components of reading (i.e., comprehension, fluency, vocabulary, phonics, and phonemic awareness).

The alternate content standards also provide detailed performance information for each alternate standard at each grade level including achievement descriptors and target skills for those descriptors. These descriptors and target skills provide examples that are different from grade level achievement and can be used to track the vertical alignment across the different grade levels of the content and performance expectations for students who participate in the alternate assessment.

Professional Development Materials: Background information on standards based instruction can be found in the Access PowerPoint and in the math roll out materials (i.e., IDEA, NCLB). Detailed IEP guidance can be found in the Access powerpoint. The math access document provides examples of assistive technology that can be used to promote independent responses by students. The math roll out materials and Access powerpoint also provide examples of grade level (math roll out) and alternate (Access ppt) student work products.

The math roll out materials provide in depth descriptions and examples of grade level performance standards. While these materials do not address performance standards for students with disabilities, the Access powerpoint and alternate content standards do. General education classroom activities can be found in the math roll out materials, but there is limited evidence of adapting the grade level academic activities for students with significant cognitive disabilities (e.g., The Giver example in the Access ppt and the math access document). There is also limited information for special education teachers to use to learn how to maintain the fidelity of the content standard within their instruction. While the Access PowerPoint provides a body of evidence rubric that will be used to score the alternate assessment supporting evidence, no other information is provided regarding how to promote concept generalization (i.e., using the same skill or concept across materials or content areas) or student mastery of skills. Finally, there is limited information regarding how to facilitate instruction in an inclusive context and/or with typical peers.

Recommendations:

While there clearly is training on instructional alignment provided, further development of this may be needed to help teachers link to grade level content, increase expectations across grade bands, and utilize inclusive educational opportunities whenever appropriate. More information on promoting student mastery (fading prompts) and teaching across materials may also be needed if recommendations are followed to change expectations for proficiency.

Overall Summary and Recommendation

There is clear evidence that the alternate assessment was aligned to the grade level content standards. Almost all of the reading and math alternate standards and AA items were rated academic. Science items were more problematic with only 59% of the AA items rated academic and it is recommended that these items be reviewed for improvement. The reason for the low rating appeared to be the passive expectations of the AA items or the language was too vague for the raters to understand the student expectations.

Alignment to grade level content standards was considered in the development of the alternate assessment. All of the alternate content standards and AA items were referenced to the grade level content standards and demonstrated excellent content coverage.

Overall, the alternate assessment demonstrated differentiation in content and expectations across the grade levels. One area that might need to be examined is the math items in the strand of algebra, statistics and probability, and number sense. Content experts reported repeating items across all the grade levels. Most of the supporting evidence was rated as age neutral or age appropriate.

The alternate assessment items were accessible to students at all levels of symbolic communication and many of the items allowed flexibility in response mode but one recommendation should be considered. Accommodations and adaptations are allowed for all students.

Teacher training was aligned to the grade level content standards. To improve teacher training, it is recommended that more information and training be developed on promoting student mastery.

References

- Browder, D. M., Wakeman, S. Y., Flowers, C., Rickelman, R., Pugalee, D., Karvonen, M. (2007). Creating access to the general curriculum with links to grade level content for students with significant cognitive disabilities: An explication of the concept. *Journal of Special Education, 41*, 2-16.
- Glaser, R., R. Linn, and G. Bohrnstedt 1997 *Assessment in Transition: Monitoring the Nation's Educational Progress*. Stanford, CA: National Academy of Education.
- U. S. Department of Education. (2005). *Alternate achievement standards for students with the most significant cognitive disabilities*. Washington, DC: Author. Retrieved August 12, 2005, from <http://www.ed.gov/policy/elsec/guid/altguidance.doc>
- 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.
- Webb, N. L. (2005, November). *Alignment, depth of knowledge, and change*. Paper presented at the annual meeting of the Florida Educational Research Association, Miami, FL.