National Center and State Collaborative

Appendix 3-B Study 3

NCSC GSEG Research Report

Alignment of NCSC English Language Arts and Mathematics Operational Items to Academic Grade-Level Content Targets

(page is intentionally blank)

National Center and State Collaborative

NCSC GSEG Research Report

Alignment of English Language Arts and Mathematics Operational Items to Academic Grade-Level Content Targets

September 2015

All rights reserved. Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

National Center and State Collaborative.[1] (2015). *NCSC GSEG Research Report: Alignment of English Language Arts and Mathematics Operational Items to Academic Grade-Level Content Targets.* Minneapolis, MN: University of Minnesota, National Center and State Collaborative.

[1] Contributing partners include in alphabetical order: Lauren Deters, Claudia Flowers, Jessica McCarrick, Lori Nebelsick-Gullet, Rachel Quenemoen, Martha Thurlow, Liz Towles, and Shawnee Wakeman.

National Center and State Collaborative

The National Center and State Collaborative (NCSC) is applying the lessons learned from the past decade of research on alternate assessments based on alternate achievement standards (AA-AAS) to develop a multi-state comprehensive assessment system for students with the most significant cognitive disabilities.

NCSC is a collaborative of 24 states and five organizations. The NCSC states participating in the Spring 2015 NCSC operational assessment are: Arizona, Arkansas, Connecticut, District of Columbia, Idaho, Indiana, Pacific Assessment Consortium (PAC-6)[2], Maine, Montana, New Mexico, Rhode Island, South Carolina, South Dakota, and US Virgin Islands. As of Spring 2015, additional states are members of the NCSC Consortium, representing varying levels of participation. They are: California, Delaware, Florida, Louisiana, Maryland, New York, Oregon, Pennsylvania, Tennessee, and Wyoming.

The five NCSC partner organizations include: National Center on Educational Outcomes (NCEO) at the University of Minnesota, National Center for the Improvement of Educational Assessment (Center for Assessment), University of North Carolina at Charlotte, University of Kentucky, and edCount, LLC.

IDEAs that Work
U.S. Office of Special Education Programs

This work was supported by a grant from the U.S. Department of Education, Office of Special Education Programs (H373X100002, Project Officer: Susan.Weigert@ed.gov). The contents do not necessarily represent the policy of the U.S. Department of Education, and no assumption of endorsement by the Federal government should be made.

The University of Minnesota is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

This document is available in alternative formats upon request.

[2] The Pacific Assessment Consortium (including the entities of American Samoa, Commonwealth of the Northern Mariana Islands, Federated States of Micronesia, Guam, Republic of Palau, and Republic of the Marshall Islands) partner with NCSC as one state, led by the University of Guam Center for Excellence in Developmental Disabilities Education, Research, and Service (CEDDERS).

Table of Contents

Executive Summary i

Alignment of English Language Arts and Mathematics Operational Items to Academic Grade-Level Content Targets 1

Introduction 1

The LAL Methodology 1

Application of the LAL Methodology to NCSC Operational Item Alignment Study 2

Overview of the Operational Items 3

Methodology 3

Panelists 3

Rating Rubrics 5

Data Collection Process 9

Inter-rater Agreement 10

Participant Evaluation 10

Considerations for Analysis and Interpretation of Results 11

Results 12

Academic Content of the Operational Items 12

Alignment of Operational Items to the Academic Grade-Level Content Targets 12

Alignment of ELA Items and Academic Grade-Level Content Targets 12

Mathematics Items and Academic Grade-Level Content Targets Alignment 16

Summary of Alignment of Operational Items to the Academic Grade-Level Content Targets 20

Alignment of Academic Grade-Level Content Targets, Focal KSAs, and Essential Understandings 20

ELA Results 21

Mathematics Results 25

Summary of the Alignment of Content Targets 29

Differentiation Across Grades 30

Barriers 35

Overall Analysis of Alignment 35

Limitations and Recommendations 36

References 37

Appendix A: Alignment Panelist Background Questionnaire 38

Appendix B: Differentiation Form **40**
Appendix C: Minimizing Barriers Form **41**
Appendix D: Item Alignment Study Agenda **42**
Appendix E: Panelist Evaluation Form **44**
Appendix F: Alignment of Items to Focal KSAs or Essential Understandings **45**

List of Exhibits

Exhibit 1. Alignment Aspects Addressed in the Operational Item Alignment Study..... 3
Exhibit 2. Demographic Characteristics of Expert Panelists..... 4
Exhibit 3. Experience with Professional Development Activities 5
Exhibit 4. National Standards in ELA and Mathematics 6
Exhibit 5. Content Centrality Rubric 7
Exhibit 6. Performance Centrality Rubric..... 7
Exhibit 7. DOK Rubric 8
Exhibit 8. Vertical Relationships 9
Exhibit 9. Item Level Interrater Agreement, by Panel 10
Exhibit 10. Panelist Evaluation Results 11
Exhibit 11. Number of Academic Grade-Level Content Targets and Items Reviewed by Grade 12
Exhibit 12. ELA Content Centrality of Items to Relation to Academic Grade-Level Content Targets, by Grade..... 13
Exhibit 13. ELA Performance Centrality of Items to Relation to Academic Grade-Level Content Targets, by Grade..... 14
Exhibit 14. ELA DOK Ratings: Academic Grade-Level Content Targets..... 15
Exhibit 15. ELA DOK Ratings: Items 15
Exhibit 16. Comparison of the DOK Ratings of ELA Items and the DOK Ratings of Academic Grade-Level Content Targets 16
Exhibit 17. Mathematics Content Centrality of Items in Relation to Academic Grade-Level Content Targets, by Grade 17
Exhibit 18. Mathematics Performance Centrality of Items in Relation to Academic Grade-Level Content Targets, by Grade 17
Exhibit 19. Mathematics DOK Ratings: Academic Grade-Level Content Targets 19
Exhibit 20. Mathematics DOK Ratings: Items..... 19
Exhibit 21. Comparison of DOK Ratings of Mathematics Items and DOK Ratings of Academic Grade-Level Content Targets 20
Exhibit 22. ELA Content Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade..... 21
Exhibit 23. ELA Content Centrality of Essential Understandings in Relation to Academic Grade-Level Content Targets, by Grade 22
Exhibit 24. ELA Performance Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade 22

Exhibit 25. ELA Performance Centrality of Essential Understandings in Relation to the Academic Grade-Level Content Targets, by Grade 23

Exhibit 26. ELA DOK Ratings: Focal KSAs 24

Exhibit 27. ELA DOK Ratings: Essential Understandings 24

Exhibit 28. Comparison of DOK Ratings of Focal KSAs and DOK Ratings of Academic Grade-Level Content Targets 25

Exhibit 29. Comparison of DOK Ratings of the Essential Understandings and DOK Ratings of Academic Grade-Level Content Targets 25

Exhibit 30. Mathematics Content Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade 26

Exhibit 31. Mathematics Content Centrality of Essential Understandings in Relation to Academic Grade-Level Content Targets, by Grade 26

Exhibit 32. Mathematics Performance Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade 27

Exhibit 33. Mathematics Performance Centrality of Essential Understandings in Relation to Academic Grade-Level Content Targets, by Grade 27

Exhibit 34. Mathematics DOK Ratings: Focal KSAs 28

Exhibit 35. Mathematics DOK Ratings: Essential Understandings 28

Exhibit 36. Comparison of DOK Ratings of Focal KSAs and DOK Ratings of Academic Grade-Level Content Targets 29

Exhibit 37. Comparison of DOK Ratings of Essential Understandings and DOK Ratings of Academic Grade-Level Content Targets 29

Exhibit 38. ELA Differentiation Across Grade Levels 31

Exhibit 39. Mathematics Differentiation Across Grade Levels 33

Executive Summary

To serve students with the most significant cognitive disabilities, the National Center and State Collaborative (NCSC) is designing an alternate assessment based on alternate achievement standards (AA-AAS) as part of a system that includes curriculum, instruction, and professional development resources, all aligned to college and career-ready standards. NCSC is committed to developing an assessment that allows all students to demonstrate what they know and can do in mathematics, reading, and writing and that aligns to the identified content standards.[3]

As part of the assessment development process, NCSC conducted an alignment study in 2015. The purpose of the alignment study was to examine the relationship between the spring 2015 operational assessment items and academic grade-level content targets prioritized for the NCSC AA-AAS. NCSC will use results from this study to provide content validity evidence for the NCSC summative assessment. This report summarizes the results from the alignment study for English language arts (ELA) and mathematics.

The National Alternate Assessment Center (NAAC) developed the Links for Academic Learning (LAL; Flowers, Wakeman, Browder, & Karvonen, 2007), which established the criteria used in this alignment study. The criteria are:

1. The content is academic;
2. The focus of achievement maintains fidelity with the content of the original grade-level standards (content centrality) and when possible, the specified performance (performance centrality); and
3. The depth of knowledge (DOK) differs from academic grade-level content targets, as described in the academic grade-level content targets, but maintains high expectations for students with the most significant cognitive disabilities.
4. There is some differentiation in achievement across grade levels or grade bands; and
5. The potential barriers to demonstrating what students know and can do are minimized in the assessment.

Researchers framed this executive summary with the following question: To what extent do operational items align with the academic grade-level content targets in terms of content centrality, performance centrality, and depth of knowledge?

Overview of the Operational Items

[3] For a full copy of the report, please contact the Arizona Department of Education Director of Alternate Assessment, Audra Ahumada, at Audra.Ahumada@azed.gov.

NCSC created four forms (or versions) of the mathematics operational test and four forms of the ELA operational test at each grade level. At all grade levels, the mathematics tests contained 40 items; the number of items on each ELA test depended on the number of items that accompanied each passage, so ELA test lengths varied. Each mathematics and ELA item corresponded to an academic grade-level content target. NCSC item developers created items at four levels of complexity to provide students multiple points of entry to show what they know and can do relative to each content target. Items at complexity levels 2–4 directly related to grade-level academic content, known in the NCSC system as the focal knowledge, skills, and abilities (focal KSAs) for item development. Items at complexity level 1 allowed students who were just beginning to interact with the academic content to show what they know with simplified content, known in the NCSC system as Essential Understandings. Both the mathematics and ELA tests contained items that corresponded to all four levels of complexity.

Methodology

Panelists

Researchers recruited curriculum and special education experts and distributed these experts among eight panels. Each content area, mathematics and ELA, had four panels corresponding to four grade spans: grades 3-4, grades 5-6, grades 7-8, and grade 11. Researchers recruited panelists based on the following criteria:

- Certified teacher; active faculty member; or consultant with expertise in significant cognitive disabilities, the assigned content area, or both, for the grade range of the panel to which she or he is assigned;
- No previous or current involvement in developing the items, academic grade-level content targets, focal KSAs, or Essential Understandings.

Using professional networks and contacts in the field, NCSC researchers recruited four panelists for each of the mathematics and ELA panels except the ELA panels for grades 5–6 and grades 7–8, both of which had only three panelists (*N*=30). The panelists had an average of 18 years of experience as professional educators (median = 14). Within their area of expertise, panelists had 15.9 years of experience (median = 13). The majority of panelists were female (n=22), Caucasian (n=21), and held a master's degree (n=14).

Rating Rubrics

As the first step in the process of evaluating the alignment of assessment items to both the academic grade-level content targets and the focal KSAs or Essential Understandings, panelists evaluated the academic content of the assessment items. Panelists used the national standards as defined by the National Council of Teachers of English (NCTE) and National Council of Teachers of Mathematics (NCTM) to rate the academic content of the items (see Exhibit 1).

Exhibit 1. National Standards in ELA and Mathematics

National Council of Teachers of English	National Council of Teachers of Mathematics
Reading (E1)	Numbers and operations (M1)
Writing (E2)	Algebra (M2)
Speaking (E3)	Geometry (M3)
Listening (E4)	Measurement (M4)
Viewing or visually representing (E5)	Data analysis and probability (M5)
Research (E6)	

Panelists also rated

- the degree of content centrality between the academic grade-level content target and the corresponding focal KSA,
- the degree of content centrality between the academic grade-level content target and the corresponding Essential Understanding, and
- the degree of content centrality between the operational item and the academic grade-level content target and the corresponding focal KSA or Essential Understanding (see Exhibit 2).

Content centrality is a measure of the degree of fidelity between the content of the operational assessment items and the content of the academic grade-level content targets, or the degree of fidelity between the content of the operational assessment items and either the focal KSAs or Essential Understandings. Panelists rated each pair as having *no link*, a *far link*, or a *near link*. For any rating of *no link*, researchers asked panelists to provide a rationale.

Exhibit 2. Content Centrality Rubric

Rating		Description
0	No link	The item measures NONE of the content of the academic grade-level content target.
1	Far link	The item PARTIALLY MATCHES the performance of the academic grade-level content target.
2	Near link	The same content is present in the item and the grade-level content target.

As panelists rated content centrality, they also rated

- the degree of performance centrality between the academic grade-level content target and the corresponding focal KSA,
- the degree of performance centrality between the academic grade-level content target and the corresponding Essential Understanding,
- the degree of performance centrality between the items and both the academic grade-level content targets and the corresponding focal KSA or Essential Understanding.

Performance centrality represents the degree to which the operational assessment item and the corresponding academic grade-level content target contain the same performance expectation. Panelists rated each pair as having *none*, *some*, or *all* for the performance centrality evaluated (see Exhibit 3). For any rating of *none*, researchers asked panelists to provide a rationale.

Exhibit 3. Performance Centrality Rubric

Rating		Description
0	None	None of the performance expectations found in the item are evident in the academic content standards.
1	Some	Some of the performance expectations found in the item are evident in the academic content standards.
2	All	The performance expectations found in the items are the same as the expectations in the academic content standards.

Panelists also provided depth of knowledge (DOK) ratings for the academic grade-level content targets, the focal KSAs, Essential Understandings, and the items. DOK examines the consistency between the cognitive demands of the operational assessment items and the cognitive demands of the academic grade-level content targets, focal KSAs, or Essential Understandings (Webb, 1997). Rated on a 5-point scale, DOK[4] provided a measure of the complexity of performance required to show 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 required of the student. Panelists considered the verb in conjunction with the content when determining DOK (see Exhibit 4).

[4] There are many scales for determining DOK levels. The one provided in this manual is just one of many rating scales available and is based on Bloom's taxonomy.

Exhibit 4. DOK Rubric

DOK Level		Description
1	Attention	Touch, look, vocalize, respond, attend
2	Recall/Memorize	Recall or recognition of a fact, information, concept, or procedure (identify, recall, recognize, memorize, use, classify, measure)
3	Performance	Use of information, conceptual knowledge, following or selecting appropriate procedures, two or more steps with decision points along the way, routine problems, organizing/displaying data (perform, demonstrate, follow, count, locate, read)
4	Comprehension	Requires reasoning, developing a plan or sequence of steps to approach problem; requires some decision making and justification; abstract and complex; often more than one possible answer (explain, conclude, restate, review, describe (concepts), paraphrase, infer, summarize, illustrate)
5	Application	An investigation or application to real world; requires time to research, think, and process multiple conditions of the problem or task; non-routine manipulations, across disciplines/content areas/multiple sources (compute, organize, collect, apply, classify, construct, solve, use, order, develop, generate, interact with text, implement)

Finally, panelists examined the vertical relationship of the items across grades by examining differentiation in a particular grade band (e.g., grades 3–4). They used a Differentiation Form to provide holistic judgments on the items, applying Webb's (2005) definitions of change across ELA and mathematics (see Exhibit 5 for definitions). Panelists rated the degree of change using the following rating system: (a) 0%: not evident; (b) 25%: limited evidence of change; (c) 50%: partial evidence of change; and (d) 75%: clear evidence of change. Panelists also provided examples of evidence for each of their ratings.

Exhibit 5. Vertical Relationships

Vertical Relationship	Description
Broader	Standards or items at a higher grade reflect a broader application of target skill or knowledge.
Deeper	Standards or items at a high grade reflect a deeper mastery of the target skill or knowledge.
Prerequisite	Standards or items at a lower grade reflect a different prerequisite skill for mastery of the higher-grade standard.
New	The higher grade has a new skill or knowledge unrelated to skills or knowledge covered at prior grades.
Identical	Standards or items at a higher grade appear identical to one of the lower-grade standards.

Panelists with expertise in special education provided additional feedback about whether the assessment items and materials provided a wide range of communication modes, accommodations, or modifications that would allow students access to the assessment. Four ELA special education experts and three mathematics special education experts (n=7), completed the Minimizing Barriers Form to evaluate accessibility for students characterized as follows:

- visual impairment/legally blind;
- hearing impaired;
- deaf/blind; or currently use one of the following nonverbal communication modes– respond using printed words;
- nonverbal – respond using pictures;
- nonverbal – respond using manual signs;
- nonverbal – respond using eye-gaze;
- verbal but does not use hands; or
- communicates with objects or by indicating “yes” or “no.”

Operational Item Alignment Results by Criteria

Criterion 1: The content is academic.

Researchers asked panelists to identify the national standards to which the items aligned. If panelists indicated an item did not align to a national standard, then panelists indicated if the content of the item was foundational. Both ELA and mathematics panelists indicated that 100% of the items aligned to a national standard, thus establishing that the items were academic

Criterion 2: The focus of achievement maintains fidelity with the content of the original grade-level standards (content centrality) and when possible, the specified performance (performance centrality).

Researchers examined panelists' rating of the content centrality of the operational assessment items in relation to the academic grade-level content targets (see Exhibit 5). Across all grades, panelists rated over 94% of the items in both content areas as aligned to the academic grade-level content target. Of those items that panelists rated as having no connection to the academic grade-level content targets, most were at complexity level 1 and were aligned to Essential Understandings.

Exhibit 5. Content Centrality of Items in Relation to Academic Grade-Level Content Targets

	Items	**No Link**		**Far Link**		**Near Link**		**Linked to Target**
Content	*N*	n	%	n	%	n	%	**%**
ELA	597	10	1.7	301	50.4	286	47.9	98.3
Math	699	40	5.7	263	37.6	396	56.7	94.3

Exhibit 6 displays the panelists' performance centrality ratings by content area. Nearly 90% or more of all items demonstrated some or all of the performance expectations found in the academic grade-level content targets. Because NCSC designed items at complexity level 1 using the Essential Understandings, NCSC intended the performance centrality of those items to be lower than that of items designed using the focal KSAs. Thus, NCSC expected that for all items, overall performance centrality would be lower than overall content centrality.

Exhibit 6. Performance Centrality of Items in Relation to Academic Grade-Level Content Targets

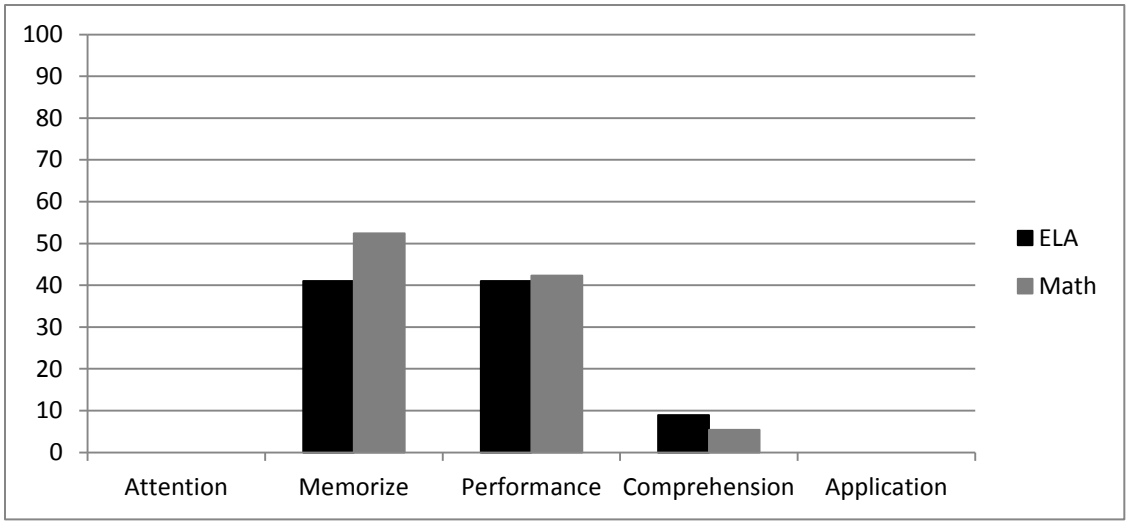
	Items	None		Some		All		Linked to Target
Content	*N*	n	%	n	%	n	%	**%**
ELA	597	42	7.0	293	49.1	262	43.9	93.0
Math	699	75	10.7	178	25.5	446	63.8	89.3

Overall, NCSC ELA and mathematics operational items met the recommended alignment standard: They were 90% aligned to content standards. As expected with alternate assessment items, the performance centrality was lower than content centrality; however, most of the items maintained the performance expectations found in the academic grade-level content targets.

Criterion 3: The depth of knowledge (DOK) differs from grade-level standards, but maintains high expectations for students with the most significant cognitive disabilities.

In examining DOK, researchers first calculated the frequency of item DOK ratings for the items by content area (see Figure 1). DOK results for the ELA and math items indicated that panelists primarily rated items at the *memorize* and *performance* levels. Panelists did not rate items at the lower (*attention*) or upper (*application*) levels, and rated very few at the *comprehension* level.

Figure 1. Percent of DOK Ratings by Content Area



Researchers also compared the item DOK ratings to the academic grade-level content target DOK ratings to investigate if panelists rated items *at*, *above*, or *below* the DOK level of the corresponding target. Researchers found that more than half of the items were rated as *at* or *above* the DOK level of the content target (see Exhibit 7).

Exhibit 7. Comparison of the DOK Ratings of Items and the DOK Ratings of Academic Grade-Level Content Targets

	Items	Above		At		Below		At/Above
Content	*N*	n	%	n	%	n	%	**%**
ELA	597	20	3.4	291	48.7	286	47.9	52.1
Math	699	73	10.4	333	47.6	293	41.9	58.0

Alignment of Focal KSAs, Essential Understandings, and Academic Grade-Level Content Targets

In addition to examining the alignment of the operational items with the academic grade-level targets and either the focal KSAs or Essential Understandings, panelists rated the alignment among the focal KSAs, Essential Understandings, and the academic grade-level content targets. Across grades and content areas, the focal KSAs had a strong link to the academic grade-level content targets. Panelists rated all of the focal KSAs as having a *far* or *near link* to the academic grade-level content targets, and panelists rated all of the focal KSAs as possessing *some* or *all* of the performance expectations found in the academic grade-level content targets. Researchers also found that the DOK levels for the focal KSAs and the academic grade-level content targets were nearly identical.

Across grades and content areas, panelists rated almost all Essential Understandings as having a *far* or *near* link to the academic grade-level content targets (95%) and having at least *some* of the performance expectations found in the academic grade-level content targets (95%). As expected, the DOK levels of most of the Essential Understandings (71%) were below the DOK levels of the academic grade-level content targets.

Criterion 4: There is some differentiation in achievement across grade levels or grade bands.

Across ELA grades 3–11, panelists indicated that the items changed to some degree in terms of all five definitions (*broader, deeper, prerequisite, new, and identical*), and they provided examples. Panelists at all grades who provided ratings reported at least limited or partial evidence of change in the *broader, deeper, prerequisite*, and *identical* relationships.

Across mathematics grades 3–11, panelists indicated that the items provided limited or partial evidence of change by all five definitions (*broader, deeper, prerequisite, new, identical*), and they provided examples. Panelists in the grades 3–4 panel indicated that the items did not measure broader skills. Panelists in grades 8 and 11 panels indicated that no skills at grade 11 were identical to skills at grade 8—the skill discrepancy is reasonable given the gap between the two grades.

Criterion 5: The potential barriers to demonstrating what students know and can do are minimized in the assessment.

Special education experts indicated whether students who were blind, deaf, deaf-blind, or who used a nonverbal mode of communication could demonstrate their knowledge on the assessment (a) as designed, with flexibility built into the tasks/items; (b) with accommodations; (c) with modifications; or (d) not at all. Panelists provided their response to all four possibilities by indicating that yes, a student with particular characteristics could demonstrate knowledge on the assessment, or no, a student with particular characteristics could not demonstrate knowledge on the assessment. Across grades and content areas, all special education experts responded that students could do the alternate assessment as designed with flexibility built into tasks and with accommodations. No special education experts indicated that students would need modifications.

Overall Analysis of Alignment

NCSC designed the operational assessment items to evaluate the knowledge and skills of a wide variety of students with the most significant cognitive disabilities. All items referenced grade-level content, and panelists rated over 90% of the items as having a *far* (partial) or *near* (full or complete) link to the content targets. A majority of the items (93% in ELA and 89% in math) maintained the performance expectations found in the academic grade-level content targets. Most of the items' DOK ratings were in the middle of the DOK distribution, with no items' ratings in the lower DOK levels. When researchers compared the DOKs of the items to the DOK levels of the academic grade-level content targets, over half of the items' DOK levels were rated *at* or *above* the level found in the academic grade-level content targets. The focal KSAs and Essential Understandings had a strong link, both in content and performance, to the academic grade-level content targets. The DOK levels for the focal KSAs were almost identical to those of the academic grade-level content targets, and the DOK levels for the Essential Understandings were below those found in the academic grade-level content targets. Overall, there was strong coherence among the operational items and the content targets for both ELA and mathematics, and there was strong vertical coherence in skills assessed by the items across the grade levels.

The study provided evidence that the assessment items allowed students in subgroups using various communication modes and with specific characteristics. Panelists indicated that the assessment items were suitable for students who used various communication modes, and panelists indicated that no modifications were necessary to enable student access to the assessment items.

Limitations

Researchers conducted the study using ELA and mathematics content experts and special education experts in the field. Researchers intended to recruit at least one special education expert per grade span, but due to availability issues, could recruit only seven special education experts; the mathematics grade 3–4 group and the mathematics grade 11 group did not have a special education expert. Though

stakeholders may want to consider the impact that a smaller sample of special education experts had on results, researchers noted that the impact was minimal.

Alignment of English Language Arts and Mathematics Operational Items to Academic Grade-Level Content Targets

Introduction

To serve students with the most significant cognitive disabilities, the National Center and State Collaborative (NCSC) is designing an alternate assessment based on alternate achievement standards (AA-AAS) as part of a system that includes curriculum, instruction, and professional development resources, all aligned to college and career-ready standards. NCSC is committed to developing an assessment that allows all students to demonstrate what they know and can do in mathematics, reading, and writing and that aligns to the content standards.

As part of the assessment development process, in early 2015 NCSC conducted a study of the alignment of operational assessment items with academic grade-level content targets. The purpose of the alignment study was to examine the relationship between the operational assessment items and academic grade-level content targets. NCSC will use results from this study to provide content validity evidence for the NCSC summative assessment. This report summarizes the results from the alignment study for English language arts (ELA) and mathematics.

The LAL Methodology

Flowers, Wakeman, Browder, and Karvonen (2007) developed the LAL alignment methodology to support the process of evaluating the alignment between alternate assessments based on alternate achievement standards (AA-AAS) and the academic content standards those assessments are meant to address. In developing the LAL method, Flowers et al. (2007) first considered the policy and practice requirements for evaluating alignment quality for AA-AAS. According to Title I—Improving the Academic Achievement of the Disadvantaged: Final Rule (2003), these assessments must:

- Be aligned with the content standards,
- Promote access to the general curriculum, and
- Reflect the highest achievement standards possible.

With these requirements in mind, Flowers et al. (2007) analyzed the alignment methods most often used in evaluations involving general assessments. Although there were many common themes among these methods, none provided a comprehensive set of criteria that would allow evaluators to address all three of the requirements noted above for AA-AAS systems. Thus, Flowers and her colleagues developed the LAL method specifically for use with AA-AAS.

The basic premises of the LAL method include the following expectations for AA-AAS (adapted from Flowers et al., 2007):

1. The assessments must be linked to grade-level content standards;
2. The target for achievement must be academic content (e.g., reading and mathematics) that corresponds to the student's assigned grade based on chronological age;
3. Functional activities and materials may be used to promote understanding, but the target skills for student achievement are academically focused;
4. Prioritization of the content, if it occurs, 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;
5. The alternate expectations for achievement may include a 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 depth of knowledge, to achieve within their symbolic level, and to show growth across grade levels or grade bands.

Thus, researchers used the following criteria from the LAL model to examine the alignment between the assessment items and the academic grade-level content targets:

1. The content is academic;
2. The focus of achievement maintains fidelity with the content of the original grade-level standards (content centrality) and when possible, the specified performance (performance centrality); and
3. The depth of knowledge (DOK) differs from grade-level content targets, as described in the academic grade-level content targets, but maintains high expectation for students with significant cognitive disabilities.
4. There is some differentiation in achievement across grade levels or grade bands; and
5. The potential barriers to demonstrating what students know and can do are minimized in the assessment.

Application of the LAL Methodology to NCSC Operational Item Alignment Study

Researchers translated the LAL premises into three aspects of alignment considered in the present study (see Exhibit 1). The following question guides this report: To what extent do operational items align with the academic grade-level content targets in terms of content centrality, performance centrality, and depth of knowledge?

Exhibit 1. Alignment Aspects Addressed in the Operational Item Alignment Study

Aspect	Evaluation questions
1. Content centrality	To what extent does the focus of achievement in the operational item maintain fidelity with the content of the associated academic grade-level content target?
2. Performance centrality	To what extent does the focus of achievement in the operational item maintain fidelity with the specified performance of the associated academic grade-level content target?
3. Depth of knowledge (DOK)	Do the DOK levels in the item reflect fidelity with the DOK levels in the associated academic grade level content targets?

Overview of the Operational Items

NCSC created four versions of the mathematics operational test and four versions of the ELA operational test at each grade level. At all grade levels, the mathematics tests contained 40 items; the number of items on each ELA test depended on the number of items that accompanied each passage, so ELA test lengths varied. Each mathematics and ELA item corresponded to an academic grade-level content target. NCSC item developers created items at four levels of complexity to provide students multiple points of entry to show what they know and can do relative to each content target. Items at complexity levels 2–4 directly related to grade-level academic content, known in the NCSC system as the focal knowledge, skills, and abilities (focal KSAs). Items at complexity level 1 allowed students who were just beginning to interact with the academic content to show what they know with simplified content, known in the NCSC system as Essential Understandings. Both the mathematics and ELA tests contained items that corresponded to all four complexity levels.

Methodology

Panelists

Researchers recruited curriculum and special education experts and distributed these experts among eight panels. Each content area, mathematics and ELA, had four panels corresponding to four grade spans: grades 3-4, grades 5-6, grades 7-8, and grade 11.Researchers recruited panelists based on the following criteria:

- Must be a certified teacher or an active faculty member or consultant with expertise in significant cognitive disabilities, the assigned content area, or both, for the grade range of the panel to which she or he is assigned;
- Must have had no previous and have no current involvement in developing the items, academic grade-level content targets, focal KSAs, or Essential Understandings.

Using professional networks and contacts in the field, NCSC researchers recruited four panelists for each of the mathematics and ELA panels except the ELA panels for grades 5–6 and 7–8, both of which had only three panelists (*N*=30).

Prior to beginning their ratings, the 30 panelists in the study completed the Panelist Background Questionnaire (see Appendix A: Alignment Panelist Background Questionnaire), which asked panelists a variety of questions about their years of experience, expertise, and demographics. The average number of years of experience panelists had in being a professional educator was 18.0, and the median was 14.0. Panelists also indicated how many years of experience they had in their area of expertise: The average number of years was 15.9, and the median number of years was 13.0 (see Exhibit 2 for other demographic characteristics). The majority of panelists were female (n=22), Caucasian (n=21), and held a master's degree (n=14).

Exhibit 2. Demographic Characteristics of Expert Panelists

Demographic Variables		Number of Panelists
Gender	Female	22
	Male	8
Ethnicity	Caucasian	23
	African American	4
	Other	2
	Missing	1
Degree	Bachelor's degree	6
	Master's degree	14
	Doctorate	10

Panelists also provided information about their experience with professional development activities related to the alignment study work (see Exhibit 3). Panelists responded with a "yes" or "no" to one of three statements about their professional experience: "I conducted professional development for teachers in their respective content area," "I have taken a leadership role in curriculum planning in their school or district," and "I taught future teachers in a higher education setting." Of the 30 panelists, the majority had conducted professional development for teachers in their respective content areas (n=23), and the majority served at the time or had previously served in a leadership role in curriculum planning for their school or district (n=22). Fewer than half of the panelists had taught future teachers in a higher education setting (n=13).

Exhibit 3. Experience with Professional Development Activities

Questionnaire Statement	Yes	No
	n	n
Please check all of the following statements that apply to you: I conduct/have conducted professional development for teachers in my content area.	23	7
Please check all of the following statements that apply to you: I have/had a leadership role in curriculum planning in my school or district.	22	8
Please check all of the following statements that apply to you: I have taught future teachers in a higher education setting.	13	17

Rating Rubrics

Panelists reviewed and provided ratings of the operational assessment items, the academic grade-level content targets, the focal KSAs, and the Essential Understandings. They provided the following ratings:

1. DOK for each of the prioritized academic grade-level content targets;
2. DOK for each of the focal KSAs and Essential Understandings associated with the prioritized academic grade-level content targets;
3. Content centrality or alignment between each prioritized academic grade-level content target and the corresponding focal KSA and Essential Understanding;
4. Performance centrality or alignment between each prioritized academic grade-level content target and the corresponding focal KSA and Essential Understanding;
5. Rating of national standard and whether each item was academic;
6. DOK for each operational assessment item;
7. Content centrality or alignment between each operational assessment item and the corresponding academic grade-level content target and focal KSA or Essential Understanding; and
8. Performance centrality or alignment between each operational assessment item and the corresponding academic grade-level content target and the focal KSA or Essential Understanding.

As the first step in the process of evaluating the alignment of assessment items to both the academic grade-level content targets and focal KSA or Essential Understandings, panelists had to evaluate the academic content of the items. Panelists used the national standards as defined by the National Council of Teachers of English (NCTE) and National Council of Teachers of Mathematics (NCTM) to rate the academic content of the assessment items (see Exhibit 4).

Exhibit 4. Categories of National Standards in ELA and Mathematics

National Council of Teachers of English	National Council of Teachers of Mathematics
Reading (E1)	Numbers and operations (M1)
Writing (E2)	Algebra (M2)
Speaking (E3)	Geometry (M3)
Listening (E4)	Measurement (M4)
Viewing or visually representing (E5)	Data analysis and probability (M5)
Research (E6)	

Content centrality is a measure of the degree of fidelity between the content of the operational assessment items and the content of the academic grade-level content targets and the focal KSAs or Essential Understandings. An operational assessment item requiring students to “identify and define a word in a sentence” and an academic grade-level content target requiring students to “identify and define grade-level concepts” would not have the same content (“parts of speech” vs. “punctuation marks”), although they do have the same performance (“identify and define”).

A content centrality rubric was applied in this study (see Exhibit 5). Panelists rated:

- the degree of content centrality between the academic grade-level content target and the corresponding focal KSA;
- the degree of content centrality between the academic grade-level content target and the corresponding Essential Understanding; and
- the degree of content centrality between the operational item and the academic grade-level content target and the corresponding focal KSA or Essential Understanding.

Panelists rated each pair as having *no link*, a *far link*, or a *near link*. For any rating of *no link*, researchers asked panelists to provide a rationale.

Exhibit 5. Content Centrality Rubric

Rating		Description
0	No link	The item measures NONE of the content of the academic grade-level content target.
1	Far link	The item PARTIALLY MATCHES the performance of the academic grade-level content target.
2	Near link	The same content is present in the item and the grade-level content target.

Performance centrality represents the degree to which the operational assessment item and the corresponding academic grade-level content target contain the same performance expectation. For example, an item requiring students to “identify and define a word in a sentence” and an academic grade-level content target requiring students to “identify and define grade-level concepts” would have the same performance expectation (“identify and define”), although the content is different. The panelists rated:

- the degree of performance centrality between the academic grade-level content target and the corresponding focal KSA;
- the degree of performance centrality between the academic grade-level content target and the corresponding Essential Understanding; and
- the degree of performance centrality between the items and the academic grade-level content targets and the corresponding focal KSAs or Essential Understandings.

Panelists rated each pair as having *none*, *some*, or *all* of the performance centrality (see Exhibit 6). For any rating of *none*, researchers asked panelists to provide a rationale.

Exhibit 6. Performance Centrality Rubric

Rating		Description
0	None	None of the performance expectations found in the item are evident in the academic content standards.
1	Some	Some of the performance expectations found in the item are evident in the academic content standards.
2	All	The performance expectations found in the items are the same as the expectations in the academic content standards.

Depth of knowledge (DOK) examines the consistency between the cognitive demands of the operational assessment items and the cognitive demands of the academic grade-level content targets, focal KSAs, or Essential Understandings (Webb, 1997).

Completely aligned standards and assessments require an assessment system designed to measure in some way the full range of cognitive complexity within each specified content standard. Rated on a 5-point scale, DOK[5] provided 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 identify the response that would be required of the student. Panelists considered the verb in conjunction with the content when determining DOK (see Exhibit 7).

Exhibit 7. DOK Rubric

DOK Level		Description
1	Attention	Touch, look, vocalize, respond, attend
2	Recall/Memorize	Recall or recognition of a fact, information, concept, or procedure (identify, recall, recognize, memorize, use, classify, measure)
3	Performance	Use of information, conceptual knowledge, following or selecting appropriate procedures, two or more steps with decision points along the way, routine problems, organizing/displaying data (perform, demonstrate, follow, count, locate, read)
4	Comprehension	Requires reasoning, developing a plan or sequence of steps to approach problem; requires some decision making and justification; abstract and complex; often more than one possible answer (explain, conclude, restate, review, describe (concepts), paraphrase, infer, summarize, illustrate)
5	Application	An investigation or application to real world; requires time to research, think, and process multiple conditions of the problem or task; non-routine manipulations, across disciplines/content areas/multiple sources (compute, organize, collect, apply, classify, construct, solve, use, order, develop, generate, interact with text, implement)

Upon completion of the centrality and DOK ratings, each group examined the vertical relationship of the items across grades by examining differentiation in a particular grade band (e.g., grades 3–4).. Differentiation refers to whether the items reflect changing expectations across grades and whether they are age appropriate. For example, achievement of the academic skills in ELA in grade 4 might reflect achievement of new skills not present in ELA in grade 3. The vertical relationship of skills should show that each grade builds upon the next, ultimately toward the exit expectations. Panelists used

[5] There are many scales for determining DOK levels. The one provided in this manual is just one of many rating scales available and is based on Bloom's taxonomy.

Webb's (2005) definition of five vertical relationships (see Exhibit 8) to capture item differentiation across grade levels (see Appendix B: Differentiation Form).

Exhibit 8. Vertical Relationships

Vertical Relationship	Description
Broader	Standards or items at a higher grade reflect broader application of target skill or knowledge
Deeper	Standards or items at a higher grade reflect deeper mastery of the target skill or knowledge
Prerequisite	Standards or items at a lower grade reflect a different prerequisite skill for mastery of the higher grade standard
New	The higher grade has a new skill or knowledge unrelated to skills or knowledge covered at prior grades
Identical	Standards or items at a higher grade appear identical to one of the lower-grade standards

Researchers asked special education experts within each group to identify barriers students with various characteristics might face in accessing the items and demonstrating what they know and can do. LAL methodology examines whether items are difficult for construct-irrelevant reasons specific to disabilities. Examining items for barriers is a method of determining whether an item excludes students with some disabilities from accurately showing what they know and can do. Special education panelists completed the Minimizing Barriers Form (refer to Appendix C: Minimizing Barriers Form) and identified barriers for a range of students with physical, sensory, and communication characteristics.

Data Collection Process

The panelists gathered to participate in the alignment study from July 14 to July 16, 2015 in Arlington, VA (see Appendix D: Item Alignment Study Agenda). On the morning of the first day, all panelists received a detailed training on the purpose of the study, the rubrics, and the rating process.

After the training, the panelists separated into their panels to begin the rating process. Each panel included a researcher who acted as a facilitator and answered questions about the application of the rubrics for DOK, content centrality, and performance centrality. This study employed a consensus model in which individual panelists first made independent ratings and then discussed their ratings in groups to reach agreement. Researchers captured both individual and consensus ratings. The facilitators observed the group discussions, recorded the consensus ratings, and took notes describing the panel's decision-making process and the rationale for their ratings.

Panelists rated content centrality, performance centrality, and DOK for the items, academic grade-level targets, focal KSAs, and Essential Understandings first as individuals and then as a group to achieve consensus. Table leaders documented consensus ratings, and panelists turned in their individual ratings to table leaders.

Once panelists finished the content centrality, performance centrality, and DOK ratings, groups completed the Differentiation Form together and the table leader captured the group's consensus. Finally, all special education experts individually completed the Minimizing Barriers Form.

Inter-rater Agreement

During the study, the facilitators for each panel monitored the independent and group ratings and retrained panelists or helped panelists understand the rating rubrics, as appropriate. The consensus process used in this study required each panel to come to 100% agreement on each consensus rating that researchers used for analysis. To ensure the reliability of the panelists' ratings, researchers calculated inter-rater agreement for panelists' initial independent ratings (see Exhibit 9). Overall, panelists across grade spans and content areas reported high levels of agreement. Panelists at ELA grades 7–8 exhibited the lowest levels of agreement, though the lack of agreement primarily occurred at grade 7 in rating the DOK of items.

Exhibit 9. Item Level Interrater Agreement, by Panel

Content	Rating	Grades 3–4	Grades 5–6	Grades 7–8	Grade 11
ELA	Item DOK	0.82	0.93	0.64	0.82
	Item to target content centrality	0.73	0.91	0.71	0.74
	Item to target performance centrality	0.77	0.78	0.74	0.89
Mathematics	Item DOK	0.84	0.78	0.78	0.89
	Item to target content centrality	0.84	0.83	0.81	0.82
	Item to target performance centrality	0.70	0.92	0.87	0.90

Participant Evaluation

At the conclusion of the study, panelists completed an evaluation of the alignment study's process (see Appendix E: Panelist Evaluation Form). Researchers asked panelists to rate their agreement—*strongly agree* (4), *agree* (3), *disagree* (2), or *strongly disagree* (1)—with a series of statements about the study. The results of these

evaluations reflect high levels of satisfaction with the process and the outcomes of the study (see Exhibit 10). Of the 26 panelists who completed a questionnaire, all strongly agreed that their performance and content centrality ratings were reasonable and that the DOK ratings for the items were reasonable. All 26 panelists also agreed that they contributed to the sessions.

Exhibit 10. Panelist Evaluation Results

Statement	Average Rating
1. Facilities were appropriate to the work.	3.5
2. Introductory presentation and training sessions were clear.	3.9
3. Introductory and training materials were clear.	3.8
4. The process used was appropriate to the work.	3.9
5. The process used resulted in good information about the items, standards, and assessments.	3.9
6. I feel like I was able to make a contribution to the sessions.	4.0
7. I am satisfied with the group-consensus DOK ratings for each standard.	3.9
8. I am confident that my performance and content centrality ratings are reasonable.	4.0
9. I am confident that my DOK ratings for the items are reasonable.	4.0

Considerations for Analysis and Interpretation of Results

Readers of this report must consider the characteristics and the intended purpose of the items and academic grade-level content targets when interpreting the study's results. The relationship between the academic grade-level content targets and the focal KSAs differs from the relationship between the academic grade-level content targets and the Essential Understandings. Test developers designed the Essential Understandings to measure introductory skills that students need to access a particular academic grade-level content target; thus, the Essential Understandings are purposefully less closely related to the academic grade-level content target than the focal KSAs are.

Similarly, the relationship between the items and the academic grade-level content targets differs from the relationship between the items and the corresponding focal KSAs or Essential Understandings. Because test developers designed the focal KSAs and Essential Understandings to be narrower in scope than the academic grade-level content targets, researchers expected to see closer relationships between the items and focal KSAs or Essential Understandings; however, researchers still expected to see close relationships between the items and the academic grade-level content targets.

Results

Academic Content of the Operational Items

Panelists linked each of the operational items to a national standard defined by the NCTE or NCTM; if they could not link an item to a national standard, panelists indicated whether the item was foundational. In both ELA and mathematics, panelists linked 100% of the items to a national standard, thus establishing that the items were academic. Panelists often selected multiple national standards for each item, so researchers did not calculate frequencies for each standard, as doing so would not have represented the number of items.

Alignment of Operational Items to the Academic Grade-Level Content Targets

The following section contains information about the relationship of the items to the grade-level academic content targets. Researchers describe results by content area and grade.

Exhibit 11 displays the number of grade-level academic content targets and items that panelists reviewed for each grade. Note that the number of mathematics items reviewed was the same across all grades, but the number of ELA items differed depending on the number of items that accompanied each passage.

Exhibit 11. Number of Academic Grade-Level Content Targets and Items Reviewed by Grade

	ELA		Mathematics	
	Targets	Items	Targets	Items
Grade	*N*	*N*	*N*	*N*
3	12	95	10	100
4	12	96	10	100
5	10	77	10	100
6	12	76	10	100
7	11	87	10	100
8	11	87	10	100
11	11	79	10	99
Total	79	597	70	699

Alignment of ELA Items and Academic Grade-Level Content Targets

The following section contains information about the alignment between the items and the academic grade-level content targets.

ELA Content Centrality

To what extent does the focus of achievement in the operational items maintain fidelity with the content of the associated academic grade-level content targets?

Panelists rated content centrality on a three-point scale measuring the degree of alignment (*no link*, *far link*, and *near link*) between items and academic grade-level content targets.

Panelists rated 98.3% of items as linked to the academic grade-level content target and rated most of the items as having a *far* (50%) or *near* (48%) *link* (see Exhibit 12). Ten items were rated as having *no link*—nine of these were items at complexity level 1 and were aligned to Essential Understandings (see Appendix F: Alignment of Items to Focal KSA or Essential Understanding). Panelists cited overreach or mismatch as reasons for a *no link* rating. Overall, however, these results suggest that the items have a strong relationship to the academic grade-level content targets.

Exhibit 12. ELA Content Centrality of Items to Relation to Academic Grade-Level Content Targets, by Grade

	Items	**No Link**		**Far Link**		**Near Link**		**Linked to Target**
Grade	*N*	n	%	n	%	n	%	**%**
3	95	2	2.1	41	43.2	52	54.7	97.9
4	96	0	.0	37	38.5	59	61.5	100.0
5	77	1	1.3	64	83.1	12	15.6	98.7
6	76	2	2.6	67	88.2	7	9.2	97.4
7	87	2	2.3	38	43.7	47	54.0	97.7
8	87	1	1.2	16	18.4	70	80.5	98.9
11	79	2	2.5	38	48.1	39	49.4	97.5
Total	597	10*	1.7	301	50.4	286	47.9	98.3

Note. Nine of the 10 *no link* items were items at complexity level 1 and were aligned to Essential Understandings.

Performance Centrality

To what extent does the focus of achievement in the operational item maintain fidelity with the specified performance of the associated academic grade-level content target?

Panelists rated performance centrality on a three-point scale (*none, some, all*) based on the degree of match between the items and the academic grade-level content targets.

Panelists' ratings indicated that 93% of the items had *some* or *all* of the performance expectations found in the corresponding academic grade-level content targets (see Exhibit 13). Of the 42 items that panelists rated as having *none* of the performance expectations, 37 were items at complexity level 1 and corresponded to Essential Understandings. Panelists at grades 7 and 8 rated more items as having *none* of the

performance expectations of academic grade-level content targets than did panelists at other grades.

Exhibit 13. ELA Performance Centrality of Items to Relation to Academic Grade-Level Content Targets, by Grade

	Items	None		Some		All		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	95	2	2.1	26	27.4	67	70.5	97.9
4	96	0	0	29	30.2	67	69.8	100.0
5	77	4	5.2	61	79.2	12	15.6	94.8
6	76	2	2.6	67	88.2	7	9.2	97.4
7	87	19	21.8	47	54.0	21	24.1	78.1
8	87	12	13.8	48	55.2	27	31.0	86.2
11	79	3	3.8	15	19.0	61	77.2	96.2
Total	597	42	7.0	293	49.1	262	43.9	93.0

Note. Thirty-seven of the 42 *No Link* items were items at complexity level 1 and were aligned to Essential Understandings.

Depth of Knowledge

Do the DOK levels of the item reflect fidelity with the DOK levels of the associated academic grade level content targets?

Panelists also rated the DOK for the ELA academic grade-level content targets. Across all grades, panelists rated the majority of the content targets at the *performance* (49%) and *comprehension* (35%) levels (see Exhibit 14). Panelists did not rate any academic grade-level content targets at the *attention* level.

Exhibit 14. ELA DOK Ratings: Academic Grade-Level Content Targets

		Attention		Memorize		Performance		Comprehension		Application	
Grade	***N***	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
3	12	0	0	2	16.7	7	58.3	3	25.0	0	0.0
4	12	0	0	1	8.3	6	50.0	5	41.7	0	0.0
5	10	0	0	2	20.0	5	50.0	2	20.0	1	10.0
6	12	0	0	3	25.0	2	16.7	7	58.3	0	0.0
7	11	0	0	0	0.0	8	72.7	3	27.3	0	0.0
8	11	0	0	1	9.1	6	54.6	4	36.4	0	0.0
11	11	0	0	1	9.1	5	45.5	4	36.4	1	9.1
Total	79	0	0	10	12.7	39	49.4	28	35.4	2	2.5

Panelists also rated the DOK for the operational assessment items (see Exhibit 15). Panelists rated most of the items at the *memorize* (41%) or *performance* (41%) levels. Panelists did not rate any items at the *attention* or *application* levels.

Exhibit 15. ELA DOK Ratings: Items

		Attention		Memorize		Performance		Comprehension		Application	
Grade	***N***	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
3	95	0	0.0	47	49.5	42	44.2	6	6.3	0	0.0
4	96	0	0.0	33	34.4	40	41.7	23	24.0	0	0.0
5	77	0	0.0	64	83.1	13	16.9	0	0.0	0	0.0
6	76	0	0.0	67	88.2	9	11.8	0	0.0	0	0.0
7	87	0	0.0	36	41.4	47	54.0	4	4.6	0	0.0
8	87	0	0.0	37	42.5	48	55.2	2	2.3	0	0.0
11	79	0	0.0	15	19.0	46	58.2	18	22.8	0	0.0
Total	597	0	0.0	245	41.0	245	41.0	53	8.9	0	0.0

Researchers compared the panelists' items DOK ratings with the panelists' academic grade-level content target DOK ratings (see Exhibit 16). Specifically, researchers compared the percentages of item DOKs that were above, at, or below the academic grade-level content target DOKs. In grades 3, 4, and 11, at least 50% of the item DOK ratings were at or above the academic grade-level content target DOK ratings. When researchers eliminated items at complexity level 1, at least 50% of the item DOK ratings in grades 5, 6, 7, and 8 were also at or above the academic grade-level content target DOK ratings.

Exhibit 16. Comparison of the DOK Ratings of ELA Items and the DOK Ratings of Academic Grade-Level Content Targets

	Items	Above		At		Below		At/Above
Grade	*N*	n	%	n	%	n	%	**%**
3	95	2	2.1	64	67.4	29	30.5	69.5
4	96	4	4.2	66	68.8	26	27.1	72.9
5	77	0	0.0	24	31.2	53	68.8	31.2
6	76	0	0.0	25	32.9	51	67.1	32.9
7	87	2	2.3	39	44.8	46	52.9	47.1
8	87	6	6.9	36	41.4	45	51.7	48.3
11	79	6	7.6	37	46.8	36	45.6	54.4
Total	597	20	3.4	291	48.7	286	47.9	52.1

Mathematics Items and Academic Grade-Level Content

Targets Alignment

Content Centrality

To what extent does the focus of achievement in the operational item maintain fidelity with the content of the associated academic grade-level content target?

Researchers examined the content centrality of the items by grade (see Exhibit 17). Across grades, panelists rated 40 items (6%) as having *no link.* Thirty-one of the 40 items rated *no link* were complexity level 1 items, and 30 of the 40 items rated *no link* were aligned to focal KSAs or Essential Understandings.

Exhibit 17. Mathematics Content Centrality of Items in Relation to Academic Grade-Level Content Targets, by Grade

	Items	No Link		Far Link		Near Link		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	100	13	12.1	17	17.2	70	70.7	87.9
4	100	6	6.0	28	28.0	66	66.0	94.0
5	100	4	4.0	34	34.0	62	62.0	96.0
6	100	0	0.0	36	36.0	64	64.0	100.0
7	100	9	9.0	42	42.0	49	49.0	91.0
8	100	2	2.0	59	59.0	39	39.0	98.0
11	99	6	6.1	47	47.5	46	46.5	94.0
Total	699	40	5.7	263	37.6	396	56.7	94.3

Performance Centrality

To what extent does the focus of achievement in the operational item maintain fidelity with the specified performance of the associated academic grade-level content target?

Researchers also examined the performance centrality ratings for the items (see Exhibit 18). Approximately 89% of all items had *some* or *all* of the performance expectations found in the academic grade-level content targets. Of the 75 items rated as having no performance centrality, 45 were complexity level 1 items and 67 were aligned in performance to the focal KSAs or Essential Understandings. Panelists at grades 3 and 4 rated more items as having no performance centrality than did panelists at other grades.

Exhibit 18. Mathematics Performance Centrality of Items in Relation to Academic Grade-Level Content Targets, by Grade

	Items	None		Some		All		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	100	26	26.0	22	22.0	52	52.0	74.0
4	100	25	25.0	29	29.0	46	46.0	75.0
5	100	7	7.0	17	17.0	76	76.0	93.0
6	100	0	0.0	0	0.0	100	100.0	100.0
7	100	7	7.0	42	42.0	51	51.0	93.0
8	100	2	2.0	54	54.0	44	44.0	98.0
11	99	8	8.1	14	14.1	77	77.8	91.9
Total	699	75	10.7	178	25.5	446	63.8	89.3

Depth of Knowledge

Do the DOK levels of the item reflect fidelity with the DOK levels of the associated academic grade level content targets?

Panelists first rated the DOK for the academic grade-level content targets, ranking the majority of items at the *memorize* (31%)*, performance* (46%)*,* or *comprehension* (20%) levels (see Exhibit 19). Panelists did not rate the DOK of any academic grade-level content targets at the *attention* level.

Exhibit 19. Mathematics DOK Ratings: Academic Grade-Level Content Targets

		Attention		Memorize		Performance		Comprehension		Application	
Grade	***N***	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
3	10	0	0.0	6	60.0	3	30.0	1	10.0	0	0.0
4	10	0	0.0	2	20.0	8	80.0	0	0.0	0	0.0
5	10	0	0.0	5	50.0	4	40.0	1	10.0	0	0.0
6	10	0	0.0	3	30.0	4	40.0	3	30.0	0	0.0
7	10	0	0.0	3	30.0	3	30.0	3	30.0	1	10.0
8	10	0	0.0	3	30.0	4	40.0	3	30.0	0	0.0
11	10	0	0.0	0	0.0	6	60.0	3	30.0	1	10.0
Total	70	0	0.0	22	31.4	32	45.7	14	20.0	2	2.9

Panelists also rated the DOK of the operational assessment items, ranking the majority of items at the *memorize* (52%) or *performance* (42%) levels (see Exhibit 20). Panelists did not rate the DOK of any items at the *attention* or *application* levels.

Exhibit 20. Mathematics DOK Ratings: Items

		Attention		Memorize		Performance		Comprehension		Application	
Grade	***N***	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
3	100	0	0	86	86.0	14	14.0	0	0.0	0	0
4	100	0	0	82	82.0	18	18.0	0	0.0	0	0
5	100	0	0	45	45.0	51	51.0	4	4.0	0	0
6	100	0	0	67	67.0	31	31.0	2	2.0	0	0
7	100	0	0	41	41.0	48	48.0	11	11.0	0	0
8	100	0	0	25	25.0	58	58.0	17	17.0	0	0
11	99	0	0	20	20.2	75	75.8	4	4.0	0	0
Total	699	0	0	366	52.4	295	42.3	38	5.4	0	0

Researchers also investigated whether panelists rated the item DOK levels at or below the academic grade-level content target DOK levels (see Exhibit 21). In all grades except grades 4 and 6, at least 50% of the items had DOK ratings at or above the DOK ratings of the academic grade-level content targets. When researchers eliminated complexity level 1 items at grades 4 and 6, at least 50% of the items had DOK ratings at or above the academic grade-level content target DOK ratings.

Exhibit 21. Comparison of DOK Ratings of Mathematics Items and DOK Ratings of Academic Grade-Level Content Targets

	Items	Above		At		Below		At/Above
Grade	*N*	*N*	%	n	%	n	%	**%**
3	100	2	2.0	70	70.0	28	28.0	72.0
4	100	4	4.0	30	30.0	66	66.0	34.0
5	100	23	23.0	55	55.0	22	22.0	78.0
6	100	7	7.0	38	38.0	55	55.0	45.0
7	100	6	6.0	53	53.0	41	41.0	59.0
8	100	29	29.0	36	36.0	35	35.0	65.0
11	99	2	2.0	51	51.5	46	46.5	53.5
Total	699	73	10.4	333	47.6	293	41.9	58.1

Summary of Alignment of Operational Items to the Academic Grade-Level Content Targets

NCSC designed the AA-AAS to assess academic grade-level content for a wide variety of students with significant cognitive disabilities. All items assessed grade-level content and over 90% of the items had a *far* or *near* link to the content targets. A majority of the items (93% in ELA and 89% in math) met the performance expectations found in the academic grade-level content targets. Most of the items' DOK ratings were in the middle of the DOK distribution, and panelists did not rate any items in the lowest DOK level. When compared to the DOK levels of the academic grade-level content targets, the DOK levels of items were very similar, with over half rated *at* or *above* the level found in the academic grade-level content targets.

Alignment of Academic Grade-Level Content Targets, Focal KSAs, and Essential Understandings

Panelists also compared the alignment of the focal KSAs and Essential Understandings to the academic grade-level content targets. These components are part of the evidence-centered design-based process (ECD) used to create items. Part of the ECD-based process required the creation of focal KSAs and Essential Understandings. This section describes the relationship between the academic grade-level content targets and the focal KSAs and Essential Understandings.

ELA Results

Content Centrality

Panelists rated the content centrality of the focal KSAs in relation to the academic grade-level content targets (see Exhibit 22). Panelists rated all focal KSAs as having a *far* or *near* link to the academic grade-level content targets. Across all grades except grade 11, panelists rated 50% or more of the focal KSAs as having a *near* link to the academic grade-level content targets.

Exhibit 22. ELA Content Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade

	Focal KSAs	No Link		Far Link		Near Link		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	12	0	0.0	3	25.0	9	75.0	100.0
4	12	0	0.0	6	50.0	6	50.0	100.0
5	10	0	0.0	3	30.0	7	70.0	100.0
6	12	0	0.0	5	41.7	7	58.3	100.0
7	11	0	0.0	5	45.5	6	54.6	100.0
8	11	0	0.0	5	45.5	6	54.6	100.0
11	11	0	0.0	7	63.6	4	36.4	100.0
Total	79	0	0.0	34	43.0	45	57.0	100.0

Panelists also rated the content centrality of the Essential Understandings in relation to the academic grade-level content targets (see Exhibit 23). Panelists at all grades rated most of the Essential Understandings (82% to 100%) as having a *far* or *near* link to the academic grade level content targets, though panelists most often identified the Essential Understandings as having a *far* link. Panelists across all grades rated four Essential Understandings as having *no link*. Because the Essential Understandings target the introductory content required for an academic grade-level content target, NCSC expected that the Essential Understandings would have a *far link* connection to the academic grade-level content targets.

Exhibit 23. ELA Content Centrality of Essential Understandings in Relation to Academic Grade-Level Content Targets, by Grade

	Essential Understandings	No Link		Far Link		Near Link		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	12	0	0.0	11	91.7	1	8.3	100.0
4	12	0	0.0	12	100.0	0	0.0	100.0
5	10	1	10.0	9	90.0	0	0.0	90.0
6	12	0	0.0	12	100.0	0	0.0	100.0
7	11	1	9.1	8	72.7	2	18.2	90.9
8	11	0	0.0	9	81.8	2	18.2	100.0
11	11	2	18.2	9	81.8	0	0.0	81.8
Total	79	4	5.1	70	88.6	5	6.3	94.9

Performance Centrality

Panelists rated the performance centrality of the focal KSAs in relation to the academic grade-level content targets (see Exhibit 24). Panelists rated 100% of the focal KSAs across all grades as having *some* or *all* of the performance expectations found in the corresponding academic grade-level content targets.

Exhibit 24. ELA Performance Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade

	Focal KSA	None		Some		All		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	12	0	0.0	2	16.7	10	83.3	100.0
4	12	0	0.0	1	8.3	11	91.7	100.0
5	10	0	0.0	1	10.0	9	90.0	100.0
6	12	0	0.0	1	8.3	11	91.7	100.0
7	11	0	0.0	4	36.4	7	63.6	100.0
8	11	0	0.0	0	0.0	11	100.0	100.0
11	11	0	0.0	0	0.0	11	100.0	100.0
Total	79	0	0.0	9	11.4	70	88.6	100.0

Panelists also rated the performance centrality of the Essential Understandings in relation to the academic grade-level content targets (see Exhibit 25). Panelists rated most of the Essential Understandings as having *some* or *all* of the performance expectations found in the academic grade-level content targets (82% to 100% across all grades). Panelists across all grades identified the majority of the Essential Understandings as exhibiting *some* of the performance expectations found in the academic grade-level content targets. NCSC expected this outcome because Essential Understandings were designed to capture only the introductory skills and content present in the academic grade-level content targets.

Exhibit 25. ELA Performance Centrality of Essential Understandings in Relation to the Academic Grade-Level Content Targets, by Grade

	Essential Understanding	None		Some		All		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	12	0	0.0	11	91.7	1	8.3	100.0
4	12	0	0.0	12	100.0	0	0.0	100.0
5	10	1	10.0	9	90.0	0	0.0	90.0
6	12	0	0.0	12	100.0	0	0.0	100.0
7	11	1	9.1	7	63.6	3	27.3	90.9
8	11	0	0.0	9	81.8	2	18.2	100.0
11	11	2	18.2	7	63.6	2	18.2	81.8
Total	79	4	5.1	67	84.8	8	10.1	94.9

Depth of Knowledge

Researchers analyzed the focal KSA DOK ratings and found that panelists rated almost all focal KSAs at the *memorize* (15.2%)*, performance* (50.6%), or *comprehension* (32.9%) levels (see Exhibit 26). Panelists did not rate any focal KSAs at the *attention* level and rated only one focal KSA at the *application* level.

Exhibit 26. ELA DOK Ratings: Focal KSAs

		Attention		Memorize		Performance		Comprehension		Application	
Grade	***N***	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
3	12	0	0.0	2	16.7	8	66.7	2	16.7	0	0.0
4	12	0	0.0	1	8.3	6	50.0	5	41.7	0	0.0
5	10	0	0.0	2	20.0	5	50.0	2	20.0	1	10.0
6	12	0	0.0	4	33.3	2	16.7	6	50.0	0	0.0
7	11	0	0.0	1	9.1	8	72.7	2	18.2	0	0.0
8	11	0	0.0	1	9.1	6	54.6	4	36.4	0	0.0
11	11	0	0.0	1	9.1	5	45.5	5	45.5	0	0.0
Total	79	0	0.0	12	15.2	40	50.6	26	32.9	1	1.3

Panelists also provided DOK ratings for the Essential Understandings (see Exhibit 27). Panelists rated the DOK of 100% of the Essential Understandings at the *memorize* (67.1%) or *performance* (32.9%) levels. Panelists did not rate the DOK of any Essential Understandings at the *attention* or *application* levels.

Exhibit 27. ELA DOK Ratings: Essential Understandings

		Attention		***Memorize***		***Performance***		**Comprehension**		**Application**	
Grade	***N***	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
3	12	0	0.0	11	91.7	1	8.3	0	0.0	0	0.0
4	12	0	0.0	11	91.7	1	8.3	0	0.0	0	0.0
5	10	0	0.0	9	90.0	1	10.0	0	0.0	0	0.0
6	12	0	0.0	11	91.7	1	8.3	0	0.0	0	0.0
7	11	0	0.0	3	27.3	8	72.7	0	0.0	0	0.0
8	11	0	0.0	4	36.4	7	63.6	0	0.0	0	0.0
11	11	0	0.0	4	36.4	7	63.6	0	0.0	0	0.0
Total	79	0	0.0	53	67.1	26	32.9	0	0.0	0	0.0

Researchers investigated whether panelists rated the DOK levels of focal KSAs at or below the DOK levels of academic grade-level content target. Panelists rated 93.7% of the DOK levels of the focal KSAs at or above the DOK levels of the academic grade-level content target (see Exhibit 28).

Exhibit 28. Comparison of DOK Ratings of Focal KSAs and DOK Ratings of Academic Grade-Level Content Targets

	Focal KSA	Above		At		Below		At/Above
Grade	*N*	n	%	n	%	n	%	**%**
3	12	0	0.0	11	91.7	1	8.3	91.7
4	12	0	0.0	12	100.0	0	0.0	100.0
5	10	0	0.0	10	100.0	0	0.0	100.0
6	12	0	0.0	11	91.7	1	8.3	91.7
7	11	0	0.0	9	81.8	2	18.2	81.8
8	11	0	0.0	11	100.0	0	0.0	100.0
11	11	0	0.0	10	90.9	1	9.1	90.9
Total	79	0	0.0	74	93.7	5	6.3	93.7

Researchers also compared the DOK levels of Essential Understandings with the DOK levels of the academic grade-level content targets (see Exhibit 29). The majority (70.9%) of the Essential Understandings had a lower DOK level when compared to the academic grade-level content target DOK levels. Researchers expected this outcome given the targeted ability level of the Essential Understandings. Only at grade 7 did panelists give more than half of the Essential Understandings the same DOK ratings as those of academic grade-level content targets.

Exhibit 29. Comparison of DOK Ratings of the Essential Understandings and DOK Ratings of Academic Grade-Level Content Targets

	EUs	Above		At		Below		At/Above
Grade	*N*	n	%	n	%	n	%	**%**
3	12	0	0.0	2	16.7	10	83.3	16.7
4	12	0	0.0	1	8.3	11	91.7	8.3
5	10	0	0.0	3	30.0	7	70.0	30.0
6	12	0	0.0	3	25.0	9	75.0	25.0
7	11	0	0.0	6	54.5	5	45.5	54.5
8	11	0	0.0	4	36.4	7	63.6	36.4
11	11	0	0.0	4	36.4	7	63.6	36.4
Total	79	0	0.0	23	29.1	56	70.9	29.1

Mathematics Results

Content Centrality

Panelists rated the content centrality of the mathematics focal KSAs in relation to the academic grade-level content targets (see Exhibit 30). Panelists rated 100% of focal KSAs as having a *far* or *near* link to the academic grade-level content targets. Panelists rated 77.1% of focal KSAs as having a *near* link to the academic grade-level content targets.

Exhibit 30. Mathematics Content Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade

	Focal KSA	No Link		Far Link		Near Link		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	10	0	0.0	3	30.0	7	70.0	100.0
4	10	0	0.0	3	30.0	7	70.0	100.0
5	10	0	0.0	0	0.0	10	100.0	100.0
6	10	0	0.0	0	0.0	10	100.0	100.0
7	10	0	0.0	3	30.0	7	70.0	100.0
8	10	0	0.0	4	40.0	6	60.0	100.0
11	10	0	0.0	3	30.0	7	70.0	100.0
Total	70	0	0.0	16	22.9	54	77.1	100.0

Panelists also rated the content centrality of the Essential Understandings in relation to the academic grade-level content targets (see Exhibit 31). Panelists rated the majority of the Essential Understandings as having a *far* link to the corresponding academic grade-level content target (78.6%), and panelists identified five Essential Understandings as having *no* link to the academic grade-level content targets. However, NCSC designed the Essential Understandings to capture the introductory skills and content present in the academic grade-level content targets; NCSC did not intend for them to align perfectly.

Exhibit 31. Mathematics Content Centrality of Essential Understandings in Relation to Academic Grade-Level Content Targets, by Grade

	Essential Understanding	No Link		Far Link		Near Link		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	10	0	0.0	9	90.0	1	10.0	100.0
4	10	1	10.0	7	70.0	2	20.0	90.0
5	10	1	10.0	8	80.0	1	10.0	90.0
6	10	0	0.0	7	70.0	3	30.0	100.0
7	10	1	10.0	9	90.0	0	0.0	90.0
8	10	0	0.0	9	90.0	1	10.0	100.0
11	10	2	20.0	6	60.0	2	20.0	80.0
Total	70	5	7.1	55	78.6	10	14.3	92.9

Performance Centrality

Panelist rated the performance centrality of the focal KSAs in relation to the academic grade-level content targets (see Exhibit 32). All of the focal KSAs had *some* or *all* of the performance expectations found in the academic grade-level content targets. Panelists rated 78.6% of the focal KSAs as having *all* of the performance expectations and 21.4% of the focal KSAs as having *some* of the performance expectations.

Exhibit 32. Mathematics Performance Centrality of Focal KSAs in Relation to Academic Grade-Level Content Targets, by Grade

	Focal KSA	None		Some		All		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	10	0	0.0	5	50.0	5	50.0	100.0
4	10	0	0.0	2	20.0	8	80.0	100.0
5	10	0	0.0	1	10.0	9	90.0	100.0
6	10	0	0.0	0	0.0	10	100.0	100.0
7	10	0	0.0	3	30.0	7	70.0	100.0
8	10	0	0.0	4	40.0	6	60.0	100.0
11	10	0	0.0	0	0.0	10	100.0	100.0
Total	70	0	0.0	15	21.4	55	78.6	100.0

Panelists also rated the performance centrality of the Essential Understandings in relation to the academic grade-level content targets (see Exhibit 33). Most of the Essential Understandings had *some* or *all* of the performance expectations found in the academic grade-level content targets (80% to 100% across all grades). Across all grades, panelists identified eight Essential Understandings (11.4%) as possessing *none* of the performance expectations found in the academic grade-level content standards. Researchers expected this for the Essential Understandings.

Exhibit 33. Mathematics Performance Centrality of Essential Understandings in Relation to Academic Grade-Level Content Targets, by Grade

	Essential Understanding	None		Some		All		Linked to Target
Grade	*N*	n	%	n	%	n	%	**%**
3	10	2	20.0	8	80.0	0	0.0	80.0
4	10	1	10.0	9	90.0	0	0.0	90.0
5	10	2	20.0	8	80.0	0	0.0	80.0
6	10	0	0.0	4	40.0	6	60.0	100.0
7	10	1	10.0	9	90.0	0	0.0	90.0
8	10	0	0.0	9	90.0	1	10.0	100.0
11	10	2	20.0	7	70.0	1	10.0	80.0
Total	70	8	11.4	54	77.1	8	11.4	88.6

Depth of Knowledge

Panelists also rated the DOK for the focal KSAs and Essential Understandings. Panelists rated almost all focal KSAs at the *memorize* (24.3%), *performance* (48.6%), or *comprehension* (22.9%) levels (see Exhibit 34). Across all grades, panelists did not rate any of the focal KSAs at the *attention* level and rated only three focal KSAs (4.3%) at the *application* level.

Exhibit 34. Mathematics DOK Ratings: Focal KSAs

		Attention		Memorize		Performance		Comprehension		Application	
Grade	***N***	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
3	10	0	0.0	6	60.0	3	30.0	1	10.0	0	0.0
4	10	0	0.0	2	20.0	8	80.0	0	0.0	0	0.0
5	10	0	0.0	2	20.0	6	60.0	2	20.0	0	0.0
6	10	0	0.0	3	30.0	4	40.0	3	30.0	0	0.0
7	10	0	0.0	3	30.0	3	30.0	3	30.0	1	10.0
8	10	0	0.0	1	10.0	6	60.0	3	30.0	0	0.0
11	10	0	0.0	0	0.0	4	40.0	4	40.0	2	20.0
Total	70	0	0.0	17	24.3	34	48.6	16	22.9	3	4.3

Researchers examined the DOK ratings of the Essential Understandings and found that panelists rated almost all items at the *memorize* (67.1%) or *performance* (27.1%) levels (see Exhibit 35). Panelist did not rate any Essential Understandings at the *attention* level and rated only one (1.4%) grade 11 Essential Understanding at the *application* level.

Exhibit 35. Mathematics DOK Ratings: Essential Understandings

		Attention		Memorize		Performance		Comprehension		Application	
Grade	***N***	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
3	10	0	0.0	8	80.0	2	20.0	0	0.0	0	0.0
4	10	0	0.0	10	100.0	0	0.0	0	0.0	0	0.0
5	10	0	0.0	8	80.0	2	20.0	0	0.0	0	0.0
6	10	0	0.0	4	40.0	4	40.0	2	20.0	0	0.0
7	10	0	0.0	5	50.0	5	50.0	0	0.0	0	0.0
8	10	0	0.0	8	80.0	1	10.0	1	10.0	0	0.0
11	10	0	0.0	4	40.0	5	50.0	0	0.0	1	10.0
Total	70	0	0.0	47	67.1	19	27.1	3	4.3	1	1.4

Researchers examined the fidelity of the DOK levels of the focal KSAs to the DOK levels of the academic grade-level content targets (see Exhibit 36). The majority (74.3%) of the focal KSAs had the same DOK levels as their academic grade-level content targets. The exception occurred in grade 5 ratings: 50% of the focal KSAs had a higher DOK rating than those of their academic grade-level content targets.

Exhibit 36. Comparison of DOK Ratings of Focal KSAs and DOK Ratings of Academic Grade-Level Content Targets

	Items	Above		At		Below		At/Above
Grade	*N*	n	%	n	%	n	%	**%**
3	10	1	10.0	8	80.0	1	10.0	90.0
4	10	0	0.0	10	100.0	0	0.0	100.0
5	10	5	50.0	2	20.0	3	30.0	70.0
6	10	0	0.0	10	100.0	0	0.0	100.0
7	10	2	20.0	6	60.0	2	20.0	80.0
8	10	2	20.0	8	80.0	0	0.0	80.0
11	10	2	20.0	8	80.0	0	0.0	80.0
Total	70	12	17.1	52	74.3	6	8.6	91.4

Researchers also compared the DOK ratings of the Essential Understandings and the DOK ratings of the academic grade-level content targets (see Exhibit 37). Nearly half (47.1%) of the Essential Understandings' DOK levels were at or above those found in their academic grade-level content targets. A little over half (52.9%) of Essential Understandings' DOK levels were below those of the academic grade-level content targets, as expected.

Exhibit 37. Comparison of DOK Ratings of Essential Understandings and DOK Ratings of Academic Grade-Level Content Targets

	Items	Above		At		Below		At/Above
Grade	*N*	n	%	n	%	n	%	**%**
3	10	0	0.0	7	70.0	3	30.0	70.0
4	10	0	0.0	2	20.0	8	80.0	20.0
5	10	1	10.0	4	40.0	5	50.0	50.0
6	10	0	0.0	8	80.0	2	20.0	80.0
7	10	2	20.0	1	10.0	7	70.0	30.0
8	10	0	0.0	4	40.0	6	60.0	60.0
11	10	0	0.0	4	40.0	6	60.0	60.0
Total	70	3	4.3	30	42.9	37	52.9	47.1

Summary of the Alignment of Content Targets

Test developers designed the focal KSAs and Essential Understandings to help item writers target different levels of understanding. NCSC anticipated that the focal KSAs would have a strong relationship to the academic grade-level content targets but that

the Essential Understandings would have a weaker relationship. The results of the analyses support the NCSC design. Specifically, the focal KSAs had the same content, performance, and DOK levels as their academic grade-level content targets. The Essential Understandings had content and performance expectations that were similar to those of their academic grade-level content targets, but they had lower DOK levels, as expected given the design of the AA-AAS.

Differentiation Across Grades

Panelists examined the vertical relationship of the items across grades using the Differentiation Form. Specifically, panelists provided their holistic judgments on the ELA and mathematics items using Webb's (2005) definitions of change (see Exhibit 8 for definitions). Panelists rated the degree of change using the following system: (a) 0%: no evidence of change; (b) 25%: limited evidence of change; (c) 50%: partial evidence of change; and (d) 75%: clear evidence of change. Panelists also provided examples of evidence for each of their ratings.

Across ELA grades 3 through 11, panelists indicated that the items changed in terms of all five definitions and provided examples (see Exhibit 38). Panelists at all grades who provided ratings reported at least limited or partial evidence of change in the *broader, deeper, prerequisite*, and *identical* relationships. Only panelists in grades 3–4 and grades 8–11 panels indicated that the next higher grade did not introduce new skills.

Across mathematics grades 3–11, panelists indicated that the items provided limited or partial evidence of change in terms of all five definitions and provided examples (see Exhibit 39). In the grades 3–4 panel, panelists indicated that the items did not measure *broader* skills. Panelists in grades 8 and 11 indicated that no skills at grade 11 were identical to skills at grades 8. This result is reasonable given the gap between the two grades.

Exhibit 38. ELA Differentiation Across Grade Levels

Vertical Relationship	Degree of Relationship						Examples
	Grade 3–4	Grade 4–5	Grade 5–6	Grade 6–7	Grade 7–8	Grade 8–11	
Broader	25%	50%	50%	25%	25%	50%	Grade 6 items compared multiple sources whereas grade 5 items compared elements within a single source; organization of ideas (sequencing) evolved in grade 6; expository writing in grade 4 versus comparisons in grade 5; point of view varied at grade 11 (first and third), while it remained third at grade 8.
Deeper	50%	50%	50%	25%	25%	25%	Topic, main idea, and theme went deeper in grade 4 than in grade 3; passages were more complex in grade 5 than in grade 4; in grade 5, when students were asked to support details in the text, answer choices are verbatim from text, but in grade 6, students must infer answer choices; grade 11 took away more of the visual supports that appeared in grade 8.
Prerequisite	25%	25%	25%	25%	25%	75%	Compare and contrast two or more concepts within the same text in grade 5 was a prerequisite in grade 6; grade 8 asked word meaning; grade 11 asked why the author used that word; grade 8 asked to compare two texts; grade 11 asked how to use text to address a question.
New	0%	50%	25%	25%	25%	0%	Identify settings as a skill was not present in

Vertical Relationship	Degree of Relationship						Examples
	Grade 3–4	Grade 4–5	Grade 5–6	Grade 6–7	Grade 7–8	Grade 8–11	
							grade 4 but was present in grade 5; writing at grade 7 focused on different writing skills than at grade 6; grade 8 asked for the accurate use of words, but grade 7 did not.
Identical	--*	50%	50%	50%	50%	25%	Context clues, multiple meanings, finding conclusions, and main idea all present at grades 4 and 5; multiple meaning words; beginning, middle, and end at grades 5 and 6; at grades 6 and 7, producing a product, context clues, identifying evidence to support a claim (the standards look different, but the items assessing this skill look the same) are the same; at grades 7 and 8, one or more pieces of evidence to support something was the same.

*Missing.

Exhibit 39. Mathematics Differentiation Across Grade Levels

Vertical Relationship	Degree of Relationship						Examples
	Grades 3–4	Grades 4–5	Grades 5–6	Grades 6–7	Grades 7–8	Grades 8–11	
Broader	0%	75%	75%	75%	50%	25%	Rounding to tenths, place value charts, fraction operations, coordinates, conversion, and rounding decimals were broader at grade 5 than at grade 4; grade 5 worked on converting measurement, which was applied to ratios in grade 6; in grades 6 and 7, area of rectangles led to surface area, ratios led to proportional relationships.
Deeper	25%	25%	50%	75%	50%	50%	Fractional equivalency and reduction of models were deeper at grade 4 than at grade 3; grade 5 used equations without variables, but grade 6 used equations with variables; divide and multiply with integers, area of circles, linear equations with +/-, operations with percent, and ratios were all deeper at grade 7 than at grade 6; graphing, making predictions, solving equations and surface area and volume were deeper at grade 11 than at grade 8.
Prerequisite	25%	75%	50%	75%	75%	75%	Place value, equivalency, comparisons, degrees of difference, measuring perimeter/area , arrays, classifying shapes, rounding were prerequisite skills needed for grade 5; operations and ratios to percent were prerequisite skills for grade 7; from grades 7 to 8, going from noting

Vertical Relationship	Degree of Relationship						Examples
	Grades 3–4	Grades 4–5	Grades 5–6	Grades 6–7	Grades 7–8	Grades 8–11	
							relationships to describing relationships, and ratios to similarities; number line problems, calculating volume, plotting points, graphing equations, and solving 1–step equations were prerequisite skills needed for grade 11.
New	50%	50%	75%	25%	75%	50%	Numerical fractions, geometry, and grouping of objects were new at grade 4 compared with grade 3; multiplicative, plotting, measurement tools, conversions, line graphs, 2–step problems were new at grade 5; finding area, mean, median, mode, positive and negative integers, and ratios were new at grade 6; percent, scatter plots, liner functions, irrational numbers, congruency and similarities, and volume were new at grade 8.
Identical	50%	25%	25%	25%	25%	0%	Graphing expectations, comparison of fractions, word problems, and rounding were identical at grades 3 and 4; operations with adding and subtracting decimals was the same at grades 5 and 6; addition problems with variables were similar at grades 6 and 7.

Barriers

Panelists who were experts in special education provided feedback about whether the assessment items and materials allowed students who use a wide range of communication modes, accommodations, or modifications to access the assessment. In total, four special education experts in ELA and three special education experts in mathematics (N=7) completed the Minimizing Barriers Form. The form instructed special education experts to assess whether the assessment was accessible to students characterized by any of the following:

- visual impairment/legally blind;
- hearing impaired;
- deaf/blind;
- nonverbal——responds using printed words;
- nonverbal—responds using pictures;
- nonverba—responds using manual signs;
- nonverbal—responds using eye-gaze;
- verbal but does not use hands; or
- communicates with objects or by indicating “yes” or “no.”

For each type of student, the special education experts indicated whether the student could demonstrate his or her knowledge on the assessment (a) as designed with flexibility built into the tasks/items; (b) with accommodations; (c) with modifications; or (d) not at all. Panelists provided their response by indicating “yes” or “no.” Across grade and content area, all special education experts indicated that all types of students could demonstrate their knowledge and skills. Furthermore, all special education experts responded that students could do the alternate assessment as designed, with flexibility built into tasks and with accommodations. No special education experts indicated that students would need modifications.

Overall Analysis of Alignment

NCSC designed the operational assessment items to assess academic grade-level content for a wide variety of students with significant cognitive disabilities. All items assessed grade-level content, and panelists rated over 90% of the items as having a *far* or *near link* to the content targets. A majority of the items (93% in ELA and 89% in math) met the performance expectations found in the academic grade-level content targets. Most of the items’ DOK ratings were in the middle of the DOK distribution, with no item’s DOK rated in the lower levels. When compared to the DOK levels of the academic grade-level content targets, the DOK levels of items were very similar, with over half rated *at* or *above* the level found in the academic grade-level content targets. The focal KSAs and Essential Understandings had a strong link, both in content and

performance, to the academic grade-level content targets. The DOK levels for the focal KSAs were almost identical to those of the academic grade-level content targets, whereas the DOK levels for the Essential Understandings were below those of the academic grade-level content targets. Overall, researchers found strong coherence among the operational items and the content targets for both ELA and mathematics, and there was strong vertical coherence in skills assessed by the items across the grade levels.

The study provided strong evidence that the assessment provided access to students who use various communication modes. Panelists indicated that the assessment items provided provisions to these students, and panelists indicated that no modifications were necessary to enable student access to the assessment items.

Limitations and Recommendations

Researchers conducted the study using ELA and mathematics content experts and special education experts in the field. Researchers intended to recruit at least one special education expert per grade span, but due to availability issues, could recruit only seven special education experts. Though stakeholders may want to consider the impact on results of a smaller sample of special education experts, researchers noted that the impact was minimal.

The assessment items exhibited strong alignment with the academic grade-level content targets and the focal KSAs and Essential Understandings. The focal KSAs and Essential Understandings also exhibited strong alignment with the academic grade-level content targets. Prior to the next administration, NCSC state partners might consider increasing the DOK levels at a few of the ELA and mathematics grades; however, they should note that in this study, NCSC developers intentionally designed items at complexity level 1 to have lower DOK levels.

References

Flowers, C., Wakeman, S., Browder, D., & Karvonen, M. (2007). *Links for academic learning: An alignment protocol for alternate assessments based on alternate achievement standards.* Charlotte, NC: University of North Carolina.

Title I—Improving the Academic Achievement of the Disadvantaged; Final Rule, 68 Fed. Reg. 68,699 (Dec. 9, 2003), 34 C.F.R. § 200 *et seq.* (2003).

Webb, N. L. (2005, November). *Alignment, depth of knowledge, and change*. Paper presented at the annual meeting of the Florida Educational Research Association, Miami, Florida.

Appendix A: Alignment Panelist Background Questionnaire

Alignment Panelist Background Questionnaire

In reporting the results of this alignment study to those who must interpret and use them, it will be important to describe the background of the panelists who conducted the ratings; the report describing this alignment study will include summary information about the panelists who participated in this process. Researchers will use your responses to generate this summary information, but will not be used to identify you or be connected to your ratings. Please do not indicate your name on this form. You may use the back of this form to provide additional information, as needed.

1. Please indicate your current position. Include content area(s) and grade level(s).
   _____
   _____

2. Please indicate other teaching or administrative positions you have held.
   _____
   _____

3. For how many years have you been a professional educator?
   _____

4. How many years of experience do you have in your area of expertise (ELA/mathematics, special education, etc.)?
   _____
   _____

5. Please indicate the teaching and/or administrative certificates you currently hold.
   _____
   _____

6. Please indicate your degree(s) and other qualifications.
   _____
   _____

7. Do you have specialized training or experience in working with students with disabilities and/or students with the most significant cognitive disabilities? If yes, please explain.
   _____
   _____

8. Please check all of the following statements that apply to you:

- ❑ I conduct/have conducted professional development for teachers in my content area.
- ❑ I have/had a leadership role in curriculum planning in my school or district.
- ❑ I have taught future teachers in a higher education setting.

9. If there is other information that would help us understand your expertise in your content area, please describe it here (like experience as an assessment item writer or reviewer, range finder, etc.).

   _____

   _____

   _____

10. What is your gender? ❑ Male ❑ Female
11. What is your ethnicity? _____

Appendix B: Differentiation Form

Differentiation of Content Across Grades/Grade Bands in the NCSC 2015 Operational Assessment Items

Please provide a holistic judgment about the differences found across the grade levels or grade bands for the assessment items. Using the following definitions (Webb, 2005), please indicate the degree to which the change is evident. When there is some level of differentiation, provide at least one example. Use the notes of interest column to capture anything of importance. Remember that only those assessment items which were rated as academic and having at least some content centrality should be included in the current analysis.

(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 Area:_____

Vertical Relationship	**Ratings Per Grade Level/Grade Band**				**Examples**	**Notes of Interest**
	@75% (clear)	@50% (partial)	@25% (limited)	@0% (no)		
Broader						
Deeper						
Prerequisite						
New						
Identical						

Appendix C: Minimizing Barriers Form

Minimizing Barriers for Students Checklist

Subject: ELA Math Rater ID:_____

Type of student	**No provision** for students with these characteristics	Can do alternate assessment as designed, with **flexibility built into tasks**	Can do with **accommodations** available/stated (no change in construct measured)	Can do with **modifications** or supports stated (may alter construct being measured)
Visual impairment/ legally blind	Y	Y N	Y N	Y N
Hearing impaired	Y	Y N	Y N	Y N
Deaf/ blind	Y	Y N	Y N	Y N
Nonverbal; responds using printed words	Y	Y N	Y N	Y N
Nonverbal; responds using pictures	Y	Y N	Y N	Y N
Nonverbal; responds using manual signs	Y	Y N	Y N	Y N
Nonverbal; responds using eye gaze	Y	Y N	Y N	Y N
Verbal but no use of hands	Y	Y N	Y N	Y N
Communicates with objects or by indicating yes/no	Y	Y N	Y N	Y N

Appendix D: Item Alignment Study Agenda

Day 1: Tuesday, July 14, 2015

Time	Activity
8:00 to 8:30 am	Registration, welcome, and introductions
8:30 to 10:00 am	Full group training
10:00 to 10:15 am	Break
10:15 am to 12:00 pm	Academic grade-level content target and focal KSA/Essential Understanding ratings for all grades
12:00 to 12:45 pm	Lunch
12:45 to 2:45 pm	Item ratings: Standard, DOK (grades 3, 5, 7, and 11)
2:45 to 3:00 pm	Break
3:00 to 4:45 pm	Item ratings: DOK, centrality (grades 3, 5, 7, and 11)
4:45 to 5:00 pm	Wrap-up, sign is materials, and dismiss

Day 2: Wednesday, July 15, 2015

Time	Activity
8:00 to 8:15 am	Registration, welcome and logistics
8:15 to 10:00 am	Item ratings: DOK, centrality (grades 3, 5, 7, 11)
10:00 to 10:15 am	Break
10:15 am to 12:00 pm	Item ratings: DOK, centrality (grades 3, 5, 7, and 11); transition to grades 4, 6, 8, and 11
12:00 to 12:45 pm	Lunch
12:45 to 2:45 pm	Item ratings: Standard, DOK (grades 4, 6, 8, and 11)
2:45 to 3:00 pm	Break
3:00 to 4:45 pm	Item ratings: DOK, centrality (grades 4, 6, 8, and 11)
4:45 to 5:00 pm	Wrap-up, sign in materials, and dismiss

Day 3: Thursday, July 16, 2015

Time	Activity
8:00 to 8:15 am	Registration, welcome and logistics
8:30 to 10:00 am	Item ratings: DOK, centrality (grades 4, 6, 8, and 11)
10:00 to 10:15 am	Break
10:15 am to 12:00 pm	Finish item ratings; brief discussion regarding differentiation and barriers
12:00 to 12:45 pm	Lunch
12:45 to 2:45 pm	Differentiation; barriers
2:45 to 3:00 pm	Break
3:00 to 4:45 pm	Differentiation; barriers
4:45 to 5:00 pm	Wrap-up, sign in materials, and dismiss

Appendix E: Panelist Evaluation Form

EVALUATION OF SESSION PROCESSES AND OUTCOMES
Study of the Quality of Alignment
Among the Items, Core Content Connectors, focal Knowledge, Skills, and Abilities, and Essential Understandings for the NCSC AA-AAS

With which panel did you work?

- ☐ ELA Grades 3–4
- ☐ ELA Grades 5–6
- ☐ ELA Grades 7–8
- ☐ ELA Grade 11
- ☐ Mathematics Grades 3–4
- ☐ Mathematics Grades 5–6
- ☐ Mathematics Grades 7–8
- ☐ Mathematics Grade 11

Please check the box that that most closely reflects your opinion.

	Strongly agree	Agree somewhat	Disagree somewhat	Strongly disagree
1. The facilities were appropriate to the work.	☐	☐	☐	☐
2. The introductory presentation and training sessions were clear.	☐	☐	☐	☐
3. The introductory and training materials were clear.	☐	☐	☐	☐
4. The process used was appropriate to the work.	☐	☐	☐	☐
5. The process used resulted in good information about the items, standards, and assessments.	☐	☐	☐	☐
6. I feel like I was able to make a contribution to the sessions.	☐	☐	☐	☐
7. I am satisfied with the group-consensus DOK ratings for each standard.	☐	☐	☐	☐
8. I am confident that my performance and content centrality ratings are reasonable.	☐	☐	☐	☐
9. I am confident that my DOK ratings for the items are reasonable.	☐	☐	☐	☐

THANK YOU FOR YOUR TIME AND EFFORT!

If there are comments you would like to make about the study process or the ratings that resulted, or about what you learned from participating, or anything else, please make them here. Feel free to use the back of the paper if you need more space.

Appendix F: Alignment of Items to Focal KSAs or Essential Understandings

1. ELA Content Centrality of Items in Relation to Focal KSAs or Essential Understandings, by Grade

	Items	No Link		Far Link		Near Link		Linked to Focal KSA/EU
Grade	*N*	*N*	%	n	%	n	%	**%**
3	95	1	1.1	26	27.4	68	71.6	98.9
4	96	0	0.0	16	16.7	80	83.3	100.0
5	77	4	5.2	53	68.8	20	26.0	94.8
6	76	0	0.0	44	57.9	32	42.1	100.0
7	87	0	0.0	5	5.7	82	94.3	100.0
8	87	0	0.0	5	5.7	82	94.3	100.0
11	79	0	0.0	15	19.0	64	81.0	100.0
Total	597	5	0.8	164	27.5	428	71.7	99.2

2. ELA Performance Centrality of Items in Relation to Focal KSAs or Essential Understandings, by Grade

	Items	No Link		Far Link		Near Link		Linked to Focal KSA/EU
Grade	*N*	n	%	n	%	n	%	**%**
3	95	1	1.1	8	8.4	86	90.5	98.9
4	96	0	0.0	13	13.5	83	86.5	100.0
5	77	4	5.2	53	68.8	20	26.0	94.8
6	76	0	0.0	44	57.9	32	42.1	100.0
7	87	2	2.3	43	49.4	42	48.3	97.7
8	87	3	3.4	41	47.1	43	49.4	96.6
11	79	0	0.0	13	16.5	66	83.5	100.0
Total	597	10	1.7	215	36.0	372	62.3	98.3

3. Mathematics Content Centrality of Items to Focal KSAs or Essential Understandings, by Grade

	Items	No Link		Far Link		Near Link		Linked to Focal KSA/EU
Grade	*N*	n	%	n	%	n	%	**%**
3	100	5	5.0	16	16.0	79	79.0	95.0
4	100	0	0.0	20	20.0	80	80.0	100.0
5	100	0	0.0	29	29.0	71	71.0	100.0
6	100	0	0.0	26	26.0	74	74.0	100.0
7	100	6	6.0	30	30.0	64	64.0	94.0
8	100	4	4.0	35	35.0	61	61.0	96.0
11	99	1	1.0	16	16.2	82	82.8	99.0
Total	699	16	2.3	172	24.6	511	73.1	97.7

4. Mathematics Performance Centrality of Items to Focal KSAs or Essential Understandings, by Grade

	Items	None		Some		All		Linked to Focal KSA/EU
Grade	*N*	*N*	%	n	%	n	%	**%**
3	100	6	6.0	28	28.0	66	66.0	94.0
4	100	13	13.0	20	20.0	67	67.0	87.0
5	100	3	3.0	0	0.0	97	97.0	97.0
6	100	0	0.0	0	0.0	100	100.0	100.0
7	100	4	4.0	32	32.0	64	64.0	96.0
8	100	2	2.0	37	37.0	61	61.0	98.0
11	99	0	0.0	8	8.1	91	91.9	100.0
Total	699	28	4.0	125	17.9	546	78.1	96.0