



2023 No. 128

# Independent ELA Alignment Study for the Nebraska Department of Education

## Final Report

**Prepared for:** Nebraska Department of Education  
Trudy Clark, Director of Statewide Assessment  
Becky Michael, ELA Content Specialist

**NWEA**

Laura Ann Kump, Senior Operations Manager  
Ricky Foust, Director of Program Management  
Karen Davies, Program Manager  
Kellie Schmidt, Senior Manager, ELA Content

**Authors:** Emily A. Borawski  
Yvette M. Nemeth  
Yachen Luo  
Dannele C. Ferreras

**Prepared under:** Purchase Order Number: 4148  
(HumRRO No. S22-28)

**Date:** November 15, 2023

# Independent ELA Alignment Study for the Nebraska Department of Education

## Table of Contents

Executive Summary .....	1
Nebraska Student-Centered Assessment System .....	1
Nebraska’s College and Career Ready Standards .....	1
Alignment Criteria.....	1
Test Events .....	2
Alignment Workshop .....	2
Overview of Findings.....	3
Recommendations .....	5
Criterion 1: Items Represent Intended Content.....	5
Criterion 2: Items Represent Intended Categories .....	5
Criterion 3: Depth of Knowledge .....	5
Criterion 4: Achievement Level Descriptors .....	6
Chapter 1: Introduction .....	7
Chapter 2: Methods .....	9
Nebraska Student-Centered Assessment System .....	9
Nebraska’s College and Career Ready Standards .....	9
Alignment Criteria .....	9
Test Events .....	11
Panelists.....	11
Facilitator Training.....	13
Alignment Workshop .....	13
Test Security .....	14
Non-Secure Materials .....	15
Training .....	15
Chapter 3: Alignment Results .....	16
Items Assigned to a Nebraska Standard.....	17
Criterion 1: Items Represent Intended Content.....	18
Criterion 2: Items Represent Intended Categories.....	30
Criterion 3: Items Reflect Levels of Cognitive Complexity .....	44
Criterion 4: Items Reflect Achievement Level Descriptors .....	51
Process Evaluation Results .....	58

## Table of Contents (Continued)

Chapter 4: Summary and Recommendations .....	60
Criterion 1: Items Represent Intended Content.....	61
Criterion 2: Items Represent Intended Categories .....	62
Criterion 3: Depth of Knowledge .....	63
Criterion 4: Achievement Level Descriptors .....	63
References.....	64
Appendix A. Agenda .....	65
Appendix B. Panelist Requirements .....	67
Appendix C. Panelist Instructions .....	68
Appendix D. Panelist Training Slides .....	74
Appendix E. Standards (Grade 3 Example).....	79
Appendix F. Cognitive Complexity (DOK Wheel).....	87
Appendix G. Achievement Level Descriptors (Grade 3 Example) .....	88
Appendix H. Correlation Analysis .....	89
Appendix I. Number of Unique and Shared Items by Grade, Test Event, and Strand .....	90
Appendix J. DOK by Grade, Test Event, and Strand .....	94
Appendix K. ALD by Grade, Test Event, and Strand .....	106
Appendix L. Process Evaluation Tables by Grade.....	118

## Table of Contents (Continued)

### List of Tables

Table 1. Overall Alignment Benchmark Criteria .....	3
Table 2. Summary of Results by Criterion and Strand by Grade .....	4
Table 3. Chapter Descriptions .....	7
Table 4. Appendix Descriptions .....	8
Table 5. NSCAS-to-NE Standards Alignment Criteria by Test Event .....	10
Table 6. Number of Panelists by Grade .....	11
Table 7. Panelist Demographic Characteristics .....	12
Table 8. Benchmark Evaluation Criteria by Test Event .....	16
Table 9. Key Evaluation Documents .....	17
Table 10. Items Assigned to a Nebraska Standard – All Grades .....	17
Table 11. Items NOT Assigned to a Nebraska Standard – All Grades .....	17
Table 12. Criterion 1 Evaluative Benchmark .....	18
Table 13. Number of Standards Assessed by Test Event and Strand – Grade 3 .....	19
Table 14. Number of Standards Assessed by Test Event and Strand – Grade 4 .....	21
Table 15. Number of Standards Assessed by Test Event and Strand – Grade 5 .....	23
Table 16. Number of Standards Assessed by Test Event and Strand – Grade 6 .....	25
Table 17. Number of Standards Assessed by Test Event and Strand – Grade 7 .....	27
Table 18. Number of Standards Assessed by Test Event and Strand – Grade 8 .....	29
Table 19. Criterion 2 Evaluative Benchmark .....	30
Table 20. Category Representation Target Percentage Ranges for Test Blueprints and Study Criterion .....	31
Table 21. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 3 .....	33
Table 22. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 4 .....	35
Table 23. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 5 .....	37
Table 24. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 6 .....	39
Table 25. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 7 .....	41
Table 26. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 8 .....	43
Table 27. Criterion 3 Evaluative Benchmark .....	44
Table 28. Depth of Knowledge Levels and Definitions .....	44
Table 29. Distribution of Depth of Knowledge Levels – All Grades .....	44
Table 30. DOK Assessed by Test Event – Grade 3 .....	45
Table 31. DOK Assessed by Test Event – Grade 4 .....	46
Table 32. DOK Assessed by Test Event – Grade 5 .....	47



## Table of Contents (Continued)

Table 33. DOK Assessed by Test Event – Grade 6 .....	48
Table 34. DOK Assessed by Test Event – Grade 7 .....	49
Table 35. DOK Assessed by Test Event – Grade 8 .....	50
Table 36. Achievement Level Descriptor Definitions .....	51
Table 37. Distribution of Achievement Level Descriptors – All Grades .....	51
Table 38. ALD Assessed by Test Event – Grade 3 .....	52
Table 39. ALD Assessed by Test Event – Grade 4 .....	53
Table 40. ALD Assessed by Test Event – Grade 5 .....	54
Table 41. ALD Assessed by Test Event – Grade 6 .....	55
Table 42. ALD Assessed by Test Event – Grade 7 .....	56
Table 43. ALD Assessed by Test Event – Grade 8 .....	57
Table 45. Overall Alignment Benchmark Criteria .....	60
Table 46. Summary of Results by Criterion and Strand by Grade Level .....	61
Table 47. Summary Across Achievement Levels .....	62
Table H1. Correlation between DOK and ALD by Grade .....	89
Table I1. Number of Unique and Shared Items by Test Event and Strand – Grade 3 .....	91
Table I2. Number of Unique and Shared Items by Test Event and Strand – Grade 4 .....	91
Table I3. Number of Unique and Shared Items by Test Event and Strand – Grade 5 .....	92
Table I4. Number of Unique and Shared Items by Test Event and Strand – Grade 6 .....	92
Table I5. Number of Unique and Shared Items by Test Event and Strand – Grade 7 .....	93
Table I6. Number of Unique and Shared Items by Test Event and Strand – Grade 8 .....	93
Table J1. DOK by Grade, Test Event, and Strand – Grade 3 .....	94
Table J2. DOK by Grade, Test Event, and Strand – Grade 4 .....	96
Table J3. DOK by Grade, Test Event, and Strand – Grade 5 .....	98
Table J4. DOK by Grade, Test Event, and Strand – Grade 6 .....	100
Table J5. DOK by Grade, Test Event, and Strand – Grade 7 .....	102
Table J6. DOK by Grade, Test Event, and Strand – Grade 8 .....	104
Table K1. ALD by Grade, Test Event, and Strand – Grade 3 .....	106
Table K2. ALD by Grade, Test Event, and Strand – Grade 4 .....	108
Table K3. ALD by Grade, Test Event, and Strand – Grade 5 .....	110
Table K4. ALD by Grade, Test Event, and Strand – Grade 6 .....	112
Table K5. ALD by Grade, Test Event, and Strand – Grade 7 .....	114
Table K6. ALD by Grade, Test Event, and Strand – Grade 8 .....	116
Table L1. Panelist Evaluation Survey Results – Grade 3 .....	118
Table L2. Panelist Evaluation Survey Results – Grade 4 .....	119
Table L3. Panelist Evaluation Survey Results – Grade 5 .....	120
Table L4. Panelist Evaluation Survey Results – Grade 6 .....	121
Table L5. Panelist Evaluation Survey Results – Grade 7 .....	122
Table L6. Panelist Evaluation Survey Results – Grade 8 .....	123

## Table of Contents (Continued)

### List of Figures

Figure 1. Panelists by County .....	13
Figure 2. Alignment Workshop Data Collection Steps .....	14

# Independent ELA Alignment Study for the Nebraska Department of Education

## Executive Summary

The Northwest Evaluation Association (NWEA), on behalf of the Nebraska Department of Education (NDE), contracted the Human Resources Research Organization (HumRRO) to evaluate the degree of alignment between the Nebraska Student-Centered Assessment System (NSCAS) in English Language Arts (ELA) and Nebraska's College and Career Ready Standards (NE Standards) in ELA. Alignment studies are required as part of the federal assessment peer review process, provide validity evidence that the assessment measures the intended content, and inform future assessment item development. This alignment study gathered critical evidence to support inferences made about students' scores on the NSCAS in ELA.

### Nebraska Student-Centered Assessment System

Nebraska's Student-Centered Assessment System (NSCAS) is a "statewide assessment system that embodies Nebraska's holistic view of students. And helps them prepare for success in postsecondary education, career, and civic life" (*NSCAS – Nebraska Department of Education*, n.d.).

The NSCAS Growth, administered annually in the spring, is the component of NSCAS that assesses whether students have learned what they are expected to learn at their grade level. The test is administered online to all students in Grades 3-8 through Computer Adaptive Testing (CAT). However, a paper-pencil option is available for students with accommodations. The NSCAS in ELA includes approximately 45 test questions and is estimated to take 90 minutes to complete.

### Nebraska's College and Career Ready Standards

"Nebraska Revised Statute 79-760.01 requires the Nebraska State Board of Education to 'adopt measurable academic content standards for at least the grade levels required for statewide assessment.' Those standards shall cover the subject areas of reading, writing, mathematics, science, and social studies, and the State Board of Education shall develop a plan to review and update standards for those subject areas every seven years" (*Content Area Standards – Nebraska Department of Education*, n.d.).

In September 2021, the Nebraska State Board of Education approved Nebraska's College and Career Ready Standards for English Language Arts. The 2021 NE Standards in ELA require students to gain mastery of content in Reading Prose and Poetry (RP), Reading Informational Text (RI), Vocabulary (V), and Writing (W). These content categories will be referred to as "strands" in this report.

### Alignment Criteria

Alignment criteria were developed by HumRRO and approved by the Nebraska Department of Education (NDE). The inability to meet all criteria does *not* indicate that the test is invalid, only that a particular assessment aspect may need to be addressed through future item development and modifications to the test specifications.

### **This alignment study intended to address the following research questions:**

1. To what extent does the NSCAS in ELA reflect the breadth of the NE Standards in ELA?
2. To what extent does the NSCAS in ELA reflect the intended distributions of the strands outlined in the test blueprints?
3. To what extent does the NSCAS in ELA reflect a range and distribution of depth of knowledge (DOK)?
4. How well do the Achievement Level Descriptors (ALDs) capture the knowledge and skills expressed in the items?

HumRRO used an alignment methodology based on Webb’s original static form alignment criteria (Webb 1997, 1999, 2002, 2005; Wise, et al., 2015). Using this as our base, we tailored the methods to address Nebraska’s specific assessment system design for their standards and assessments and current alignment practice. We also applied an aspect of the Achieve model (2018), which incorporates the test blueprints into the alignment evaluation. We collected evidence from the NE Standards, test blueprints, and items from Grades 3-8. The purpose was to gather evidence to support the claims that the assessments align with the test blueprints and that the items are connected to an appropriate NE standard.

### **Test Events**

For the CAT test events, we requested that NWEA randomly select four CAT test events from each of the three achievement levels—Developing, On Track, and Advanced. In each achievement level, the test event was randomly selected from students obtaining the median score within the achievement level score range. Therefore, for each grade level, there were a total of 12 test events (four in Developing, four in On Track, and four in Advanced).

### **Alignment Workshop**

The virtual alignment workshop took place July 24–28, 2023. Based on qualification criteria developed in collaboration with NDE, NWEA, and HumRRO, NWEA recruited educators to serve on grade-level panels in Grades 3-8. Educators participated in a general training session led by HumRRO, which provided background on alignment, an overview of the study’s methodology, and item ratings to be collected during the workshop. Panelists received additional training on workshop materials, accessing and navigating the item viewing platform, and data collection processes.

Panelists then performed iterative steps for each item their panels reviewed. These steps included 1.) viewing secure test items, 2.) entering independent ratings into a spreadsheet, 3.) discussing independent ratings with other alignment workshop participants, and 4.) determining final ratings for each item as a group. For final ratings, panelists were instructed to reach a majority agreement (because reaching 100% consensus across all panelists for all items would be too time-consuming for this workshop) on any item in which all panelists disagree with the selected NE Standards, DOK, or ALD. The majority agreement rating for each item on the NE Standard, DOK, and/or ALD was determined through a group discussion by panelists.

## Data generated during this study included:

- *Ratings of standard identification, DOK, and ALD.* First, panelists independently identified the NE standard that best captured the assessed item content. Second, panelists were shown the standard to which the item was written. If the independent standard and the intended standard aligned, panelists moved on to their independent DOK rating. If the independent and intended standards did *not* align, panelists identified which standard was a better fit and then moved on to their independent DOK rating. Following the DOK rating, panelists moved on to their independent Achievement Level Descriptor (ALD) rating.
- *Majority ratings of standard identification, DOK, and ALDs.* A majority rating discussion was held for any item that all panelists did not assign the same standard, DOK, or ALD. A customized rating sheet was developed to allow the HumRRO facilitator to record the final majority ratings.
- *Demographic and process evaluation surveys.* At the end of the workshop, panelists completed a process evaluation survey in which they provided feedback about the quality of the workshop. The results of the process evaluation survey are outlined in Appendix L.

## Overview of Findings

Table 1 outlines the evaluative guidelines for the overall benchmark criteria, which involves a two-step process. First, test events are evaluated *within* each of the three achievement levels (Developing, On Track, and Advanced). Meeting at least three out of four test event benchmarks results in a "Met" rating while meeting or partially meeting at least two benchmarks leads to a "Partially Met" rating. If fewer than two benchmarks are met or partially met, the criterion is considered "Not Met."

Next, we assess results across the three achievement levels. If all three achievement levels are met, the final criterion is "Met." Meeting or partially meeting two achievement levels leads to a "Partially Met" rating while meeting or partially meeting less than two achievement levels results in a "Not Met" rating. These guidelines offer a structured approach to evaluating and interpreting the overall performance of Criterion 1, 2, and 3 across test events and achievement levels.

**Table 1. Overall Alignment Benchmark Criteria**

Criteria	Step 1 : Within Achievement Level	Step 2: Across Achievement Levels (Final Rating)
Criterion 1, 2, and 3	<p><b>Met:</b> At least three out of four test event benchmarks are met within each achievement level.</p> <p><b>Partially Met:</b> At least two of four test event benchmarks are met or partially met within each achievement level.</p> <p><b>Not Met:</b> Less than two of four test event benchmarks are met or partially met within each achievement level.</p>	<p><b>Met:</b> All three achievement levels are met.</p> <p><b>Partially Met:</b> Two achievement levels are met or partially met.</p> <p><b>Not Met:</b> Less than two achievement levels are met or partially met.</p>



Table 2 summarizes the alignment criteria results for Grades 3-8.

Criterion 1 measures whether items represent the intended content. Specifically, this criterion measures that alignment between the NE Standards and test items on each test event. For Criterion 1, Reading Prose and Poetry in Grades 3, 5, and 6 partially met the benchmark, while Grades 4, 7, and 8 did not meet the benchmark. Specifically, eight of 12 test events in Grade 4, 8 of 12 test events in Grade 7, and 11 of 12 test events in Grade 8 had less than half of the standards measured by items. For Reading Informational Text and Vocabulary, all grades partially met the benchmark. However, for the Writing strand, the evaluative benchmark was not met across grades. Specifically, 12 of 12 grade test events had less than half of the standards measured by items.

Criterion 2 measures whether items represent intended categories. Specifically, this criterion compares the expected distribution of items by content strand, as presented in the test blueprints, to the distribution of items on each test event. In Criterion 2, all grades partially met or met the benchmark. For Reading Prose and Poetry, Grades 3, 5, 7, and 8 partially met the benchmark, while Grades 4 and 6 met the benchmark. For Reading Informational Text, Grades 3, 4, and 6 partially met, and Grades 5, 7, and 8 met the benchmark. Vocabulary was partially met for Grades 3, 5, and 6 and met for Grades 4, 7, and 8. Lastly, Writing was met for Grades 3, 4, 5, 6, and 8 and partially met for Grade 7.

Criterion 3 measures whether items reflect levels of cognitive complexity. Specifically, the purpose of this criterion is to evaluate the type of cognitive processing required by items to examine the items' breadth of cognitive complexity using Webb's DOK. In Criterion 3, all grades met the benchmark except for Grade 4, which partially met the benchmark.

In summary, while there were variations in performance across different criteria and grade levels, most grades met or partially met the evaluative benchmarks. The results in the body of the report further detail the benchmark criteria by each test event.

**Table 2. Summary of Results by Criterion and Strand by Grade**

Grade	Criterion 1: Items Represent Intended Content	Criterion 2: Items Represent Intended Categories	Criterion 3: Items Reflect Levels of Cognitive Complexity
Grade 3	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Met</b>	<b>Met</b>
Grade 4	RP: <b>Not Met</b> RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: <b>Met</b> RI: Partially Met V: <b>Met</b> W: <b>Met</b>	Partially Met
Grade 5	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: Partially Met RI: <b>Met</b> V: Partially Met W: <b>Met</b>	<b>Met</b>
Grade 6	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: <b>Met</b> RI: Partially Met V: Partially Met W: <b>Met</b>	<b>Met</b>

Grade	Criterion 1: Items Represent Intended Content	Criterion 2: Items Represent Intended Categories	Criterion 3: Items Reflect Levels of Cognitive Complexity
Grade 7	RP: <b>Not Met</b> RI: <b>Partially Met</b> V: <b>Partially Met</b> W: <b>Not Met</b>	RP: <b>Partially Met</b> RI: <b>Met</b> V: <b>Met</b> W: <b>Partially Met</b>	<b>Met</b>
Grade 8	RP: <b>Not Met</b> RI: <b>Partially Met</b> V: <b>Partially Met</b> W: <b>Not Met</b>	RP: <b>Partially Met</b> RI: <b>Met</b> V: <b>Met</b> W: <b>Met</b>	<b>Met</b>

## Recommendations

### *Criterion 1: Items Represent Intended Content*

Based on the results, there is partial support that items represent the intended content. Examination of the blueprint NE Standards to be assessed by items indicates that there are more standards than items allowed, especially with the Writing strand. Based on these findings, we present the following recommendation for NDE's consideration:

- Revise the test specifications to align with the standard level for the Vocabulary and Writing strands rather than the sub-standard level. This is particularly relevant because the Writing strand included 20 or more sub-standards in numerous cases across various grade levels.

### *Criterion 2: Items Represent Intended Categories*

Most benchmarks across grades and content strands were either "Met" or "Partially Met." To strengthen the content strand blueprint target, for any strand that was "Partially Met," we present the following recommendation for NDE's consideration:

- Conduct a review of the NE Standards assigned to items in ELA to ensure Reading Prose and Poetry, Reading Informational Text, and Vocabulary are appropriately associated with the test items. NDE or NWEA can complete this review. Outcomes of this review may include but are not limited to re-assigning an NE Standard to an item.
- Review, across grade-level assessments, the ELA item banks for coverage of content strands. Where necessary, develop more items to ensure an adequate pool for CAT assessments.
- Examine the CAT algorithm to help ensure that the items represent the intended categories specified in the test blueprint.

### *Criterion 3: Depth of Knowledge*

The findings indicate that most items aligned with the DOK level 2. Across all grades, 70% or more of the items were aligned with a DOK level 2 or higher, except for Grade 4. However, there were a handful of grade-levels where no DOK 3 items were administered on one or more test events, while the other test events had at least one DOK 3 item. Based on these findings, we present the following recommendation for NDE's consideration:

- Evaluate the number of DOK 3 items available to determine whether a greater development effort should be made to increase the number of DOK 3 items.
- Continue to ensure balanced and effective item development by focusing on item writing efforts that maintain an appropriate distribution of DOK levels across grade levels.

#### ***Criterion 4: Achievement Level Descriptors***

The findings indicate that most items aligned with ALD level 2. Across all grades, 70% or more of the items were aligned with an ALD level 2 or higher. However, there were several grade levels where no items were aligned with an ALD level 3. Based on these findings, we present the following recommendation for NDE's consideration:

- Evaluate the number of ALD level 3 items to determine whether a greater development effort should be made to increase the number of ALD level 3 items.
- Continue to ensure balanced and effective item development by focusing on item writing efforts that maintain an appropriate distribution of ALD levels across grade levels.

## Chapter 1: Introduction

The Standards for Educational and Psychological Testing (AERA, APA, & NCME, 2014) identify alignment as a key component of validity evidence that should be collected for an assessment. Similarly, the federal Assessment Peer Review Guidance specifies that assessments must be aligned to a state’s academic content standards (U.S. Department of Education, 2018). Independent alignment studies demonstrate the validity of the assessments based on content. These evaluations document the breadth of knowledge and the level of cognitive processing expected of students during test performance. Alignment results can inform ongoing item development and test form assembly by identifying gaps in content coverage or areas in which the complexity of the test items does not match what is expected of students during instruction. In other words, an alignment study can provide validity evidence about a state assessment system by demonstrating that an assessment (a) represents the full range of the content standards intended to be assessed and (b) measures student knowledge in the same manner and at the same level of complexity as expected in the content standards.

To meet state and Federal requirements, the Nebraska Department of Education (NDE) requested an independent review of the alignment between the Nebraska Standards in English Language Arts (ELA) and Nebraska’s Student-Centered Assessment System (NSCAS) in ELA.<sup>1</sup> The Human Resources Research Organization (HumRRO) conducted the requested alignment study in July 2023.

The remaining chapters of this report present detailed information about the methods we used to examine the alignment of the NSCAS with the Nebraska Standards and our analysis of the data we collected.

**The chapters are presented as follows:**

**Table 3. Chapter Descriptions**

Chapter	Description
Chapter 1	Chapter 1 provides an <b>introduction</b> to the study and explains the importance of alignment in educational assessments, particularly in relation to validity evidence.
Chapter 2	Chapter 2 explains our <b>alignment method</b> , including the activities we completed to evaluate the alignment of the NSCAS assessment with the Nebraska Standards.
Chapter 3	Chapter 3 presents <b>results</b> describing the alignment of the NSCAS ELA assessment items to standards.
Chapter 4	Chapter 4 provides <b>recommendations</b> for the NDE to strengthen the alignment of the NSCAS assessments over time.

<sup>1</sup> NWEA serves as the vendor for the NSCAS.

Additionally, the appendices are presented as follows:

**Table 4. Appendix Descriptions**

Appendix	Description
Appendix A	Appendix A contains the Nebraska Alignment Workshop Agenda
Appendix B	Appendix B contains the panelists recruitment requirements
Appendix C	Appendix C contains the panelist rating instructions
Appendix D	Appendix D contains the panelist training slides
Appendix E	Appendix E contains an example of the Grade 3 NE Standards
Appendix F	Appendix F contains the Depth of Knowledge (DOK) Wheel
Appendix G	Appendix G contains an example of the Grade 3 Achievement Level Descriptors (ALD)
Appendix H	Appendix H contains a correlation analysis between DOK and ALD by grade
Appendix I	Appendix I contains the number of unique and shared Items by grade, test event, and strand
Appendix J	Appendix J contains DOK ratings by grade, test event, and strand
Appendix K	Appendix K contains ALD ratings by grade, test event, and strand
Appendix L	Appendix L contains the process evaluation results by grade



## Chapter 2: Methods

This chapter presents an overview of Nebraska’s Student-Centered Assessment System (NSCAS) and Nebraska’s College and Career Ready Standards (NE Standards). We also explain our alignment methodology, including the activities we completed to evaluate the alignment of the NSCAS assessment with the NE Standards in ELA.

### Nebraska Student-Centered Assessment System

Nebraska’s Student-Centered Assessment System (NSCAS) is a “statewide assessment system that embodies Nebraska’s holistic view of students. And helps them prepare for success in postsecondary education, career, and civic life” (*NSCAS – Nebraska Department of Education*, n.d.).

The NSCAS Growth, administered annually in the spring, is the component of NSCAS that assesses whether students have learned what they are expected to learn at their grade level. The test is administered online to all students in Grades 3-8 through Computer Adaptive Testing (CAT); however, a paper-pencil option is available for students with accommodations. The NSCAS in ELA includes approximately 45 test questions and is estimated to take 90 minutes to complete.

### Nebraska’s College and Career Ready Standards

“Nebraska Revised Statute 79-760.01 requires the Nebraska State Board of Education to ‘adopt measurable academic content standards for at least the grade levels required for statewide assessment.’ Those standards shall cover the subject areas of reading, writing, mathematics, science, and social studies, and the State Board of Education shall develop a plan to review and update standards for those subject areas every seven years” (*Content Area Standards – Nebraska Department of Education*, n.d.).

In September 2021, the Nebraska State Board of Education approved Nebraska’s College and Career Ready Standards for English Language Arts. The 2021 NE Standards in ELA require students to gain mastery of content in Reading Prose and Poetry (RP), Reading Informational Text (RI), Vocabulary (V), and Writing (W). These content categories will be referred to as “strands” in this report.

### Alignment Criteria

Alignment studies provide evidence to support the claim that assessments measure the content they are intended to measure. In this case, the content, or the measurement construct, is described for the NSCAS by the 2021 NE Standards in ELA. The alignment workshop was designed to evaluate how well the test items represent (align with) the 2021 NE Standards in ELA. The results presented in this report provide initial evidence of whether the NSCAS ELA assessment measures the content of the NE Standards.

#### **This alignment study intended to address the following research questions:**

1. To what extent does the NSCAS in ELA reflect the breadth of the NE Standards in ELA?
2. To what extent does the NSCAS in ELA reflect the intended distributions of the strands outlined in the test blueprints?

3. To what extent does the NSCAS in ELA reflect a range and distribution of depth of knowledge?
4. How well do the Achievement Level Descriptors (ALDs) capture the knowledge and skills expressed in the items?

Our methodology used four criteria to evaluate the alignment of the NSCAS in ELA with the NE Standards in ELA. Table 5 provides a brief description and evaluative benchmark associated with the criteria for each test event.

**Table 5. NSCAS-to-NE Standards Alignment Criteria by Test Event**

Criteria	Description	Benchmark
<b>Criterion 1:</b> Items Represent Intended Content	This criterion measured the alignment between the NE standards and test items on each test event.	<p><b>Met:</b> At least 75% of the NE Standards are assessed by items.</p> <p><b>Partially Met:</b> 50% - 74% of the NE Standards are assessed by items.</p> <p><b>Not Met:</b> Less than 50% of the NE Standards are assessed by items.</p>
<b>Criterion 2:</b> Items Represent Intended Categories	This criterion compared the expected distribution of items by content strand, as presented in the test blueprints, to the distribution of items on each test event.	<p><b>Met:</b> Nebraska content strands are +/- 5% from the minimum and maximum target values outlined in the blueprint.</p> <p><b>Not Met:</b> Nebraska content strands are <b>not</b> within +/- 5% of the minimum and maximum target values outlined in the blueprint.</p>
<b>Criterion 3:</b> Items Reflect Levels of Cognitive Complexity	This criterion focused on the cognitive complexity of items. The purpose of this criterion is to evaluate the type of cognitive processing required by items to examine the items' breadth of cognitive complexity using Webb's DOK.	<p><b>Met:</b> At least 70% of items are rated at cognitive complexity level 2 or above.</p> <p><b>Not Met:</b> Less than 70% of items are rated at cognitive complexity level 2 or above.</p>
<b>Criterion 4:</b> Items Reflect Levels of Achievement Level Descriptors	This criterion focused on the range of achievement level descriptors (ALDs). Some states include this as additional complexity information in their Peer Review submission. Using well-defined ALDs is consistent with the principles of assessment design.	Items on each test event will reflect a range of ALDs.

## Test Events

For the CAT test events, we requested that NWEA, the testing vendor, randomly select four CAT test events from each of the three achievement levels—Developing, On Track, and Advanced. In each achievement level, the test event was randomly selected from students obtaining the median score within the achievement level score range. Therefore, for each grade level, there were a total of 12 test events (four in Developing, four in On Track, and four in Advanced).

## Panelists

HumRRO, NWEA, and NDE developed qualification criteria for educators who applied to participate in grade-level review panels for the alignment study. The qualification criteria are presented in Appendix B. Participation requirements were focused on teaching experience and knowledge of the 2021 NE Standards in ELA. NWEA used these qualification criteria to recruit panelists, which were approved by NDE. The number of panelists varied by grade and ranged from three to seven panelists (Table 6).

**Table 6. Number of Panelists by Grade**

Grade	Number of Panelists
ELA 3	4
ELA 4	3
ELA 5	5
ELA 6	6
ELA 7	5
ELA 8	7

Panelists represented various demographic subgroups and regions across the state of Nebraska. Across all panels, women comprised 100% of the panelists. Most panelists identified as White/Non-Hispanic (90%). Panelists also represented a range of ages, with most panelists between the ages of 26 years old to 55 years old (84%). Additionally, 83% of panelists earned an advanced degree, with 70% having earned a master's degree and 13% having earned a doctoral degree or equivalent. Moreover, panelists were experienced educators, with 57% reporting more than fifteen years of classroom teaching experience. Panelists were also experienced in teaching students from various diverse backgrounds, including but not limited to students from low socioeconomic households (97%), students with disabilities (93%), and English language learners (90%). Table 7 summarizes the demographics of panelists participating in this study.

**Table 7. Panelist Demographic Characteristics**

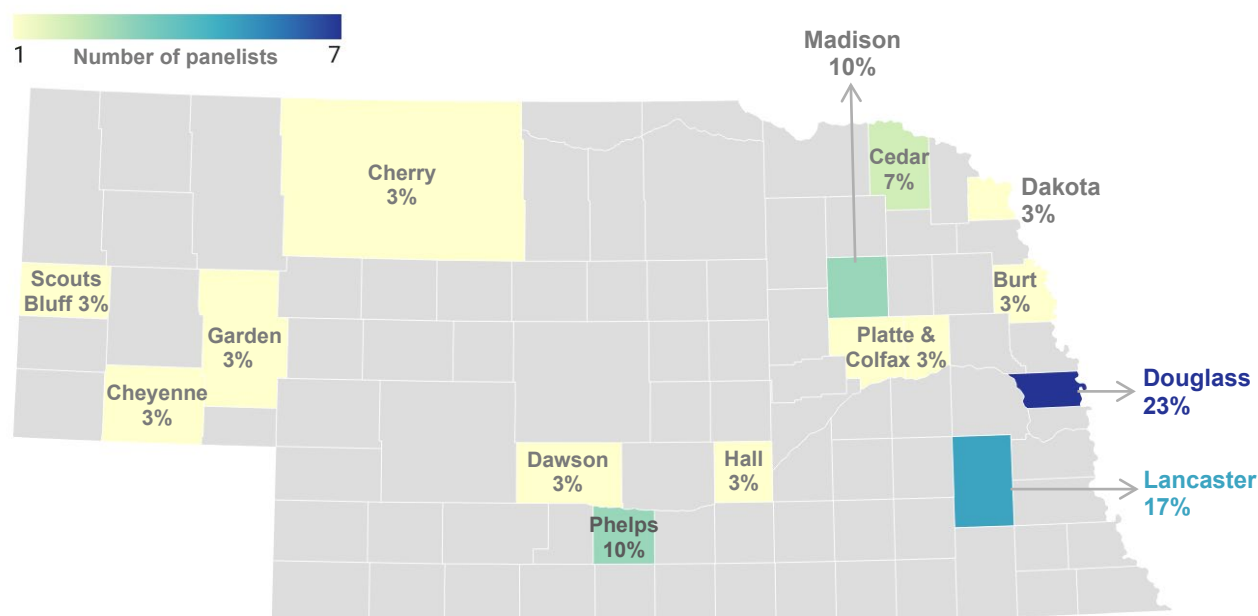
Category	Description	Count	Percentage
<b>Gender</b>	Woman	30	100%
	Man	0	0%
	Non-Binary	0	0%
<b>Race/Ethnicity</b>	White, Non-Hispanic/Latino	27	90%
	Hispanic/Latino	2	7%
	Prefer not to disclose	1	3%
<b>Age</b>	25 or under	1	3%
	26-35	6	20%
	36-45	8	27%
	46-55	11	37%
	56-65	4	13%
<b>Education</b>	Associate degree	1	3%
	Baccalaureate Degree	4	13%
	Master's Degree	21	70%
	Ph.D. or equivalent (e.g., EdD, JD)	4	13%
<b>Years of Teaching Experience</b>	Under 10	7	23%
	10-14	6	20%
	15 or more	17	57%
<b>Teaching experience with diverse backgrounds*</b>	Yes—Students from low socioeconomic households	29	97%
	Yes—Students receiving free and/or reduced lunch	29	97%
	Yes—Students with disabilities	28	93%
	Yes—English language learners	27	90%
	Yes—Students of color	26	87%
	Other (e.g., students with medical dietary concerns and refugees)	2	3%
	No	1	3%

\* Teaching experience with diverse backgrounds is a “select all that apply” response option.

Percentages will sum to greater than 100%.

Panelists represented a variety of counties across the state of Nebraska, with the majority of panelists representing the Douglass (23%) and Lancaster (17%) counties. Figure 1 below provides a visual representation of the counties across the state that were represented by panelists who participated in this study.

**Figure 1. Panelists by County**



## Facilitator Training

In preparation for the alignment workshop, HumRRO led a virtual facilitator training on July 18, 2023, with NDE staff in attendance. The facilitator training focused on providing HumRRO facilitators with an overview of the study background and purpose, workshop materials, alignment basics, data collection process, and facilitator responsibilities.

## Alignment Workshop

The virtual alignment workshop took place July 24–28, 2023. Educators participated in a general training session led by HumRRO, which provided background on alignment, an overview of the study’s methodology, and the item ratings to be collected during the workshop. Panelists received additional training on workshop materials, accessing and navigating the item viewing platform, and data collection processes.

Panelists then performed iterative steps for each item they reviewed. These steps included 1.) viewing secure test items, 2.) entering independent ratings into a spreadsheet, 3.) discussing independent ratings with other alignment workshop participants, and 4.) determining final ratings for each item as a group. For final ratings, panelists were instructed to reach a majority agreement (because reaching 100% consensus across all panelists for all items would be too time-consuming for this workshop) on any item in which all panelists disagree with the selected Nebraska Standards, DOK, or ALD. The majority agreement rating for each item on the Nebraska Standard, DOK, and/or ALD was determined through a group discussion with all panelists. An overview of these steps is outlined in the graphic below.



**Figure 2. Alignment Workshop Data Collection Steps**



**Data generated during this study included:**

- *Ratings of standard identification, DOK, and ALD.* First, panelists independently identified the NE standard that best captured the item content being assessed. Second, panelists were shown the standard to which the item was written. If the independent and intended standards align, panelists moved on to their independent DOK rating. If the independent and intended standards did *not* align, panelists identified which standard was a better fit and then moved on to their independent DOK rating. Following the DOK rating, panelists moved on to their independent Achievement Level Descriptor (ALD) rating.
- *Majority ratings of standard identification, DOK, and ALDs.* A majority rating discussion was held for any item that all panelists did not assign the same standard, DOK, or ALD. A customized rating sheet was developed to allow the HumRRO facilitator to record the final majority ratings.
- *Demographic and process evaluation surveys.* At the end of the workshop, panelists completed demographic and process evaluation surveys in which they provided feedback about the quality of the workshop.

The Grades 4, 5, and 6 panels completed their item ratings early, adjourning on Thursday, July 27, 2023. The Grades 3, 7, and 8 panels completed their item ratings on time and adjourned on Friday, July 28, 2023.

## Test Security

Test security was ensured in several ways. First and foremost, all panelists had to sign a non-disclosure agreement (NDA) stating they understood they were responsible for test security of the items being reviewed and would not share any test content with outside individuals. Before the workshop, HumRRO staff were given secure access to the NSCAS ELA items through the on-line Content Review Tool. Accounts for HumRRO facilitators and panelists to log into the Content Review Tool each day during the workshop were created. To further maintain the security of the items, panelist access to the items was turned on each morning and turned off at the conclusion of each workshop day.

## Non-Secure Materials

NWEA and NDE also provided HumRRO with several reference materials to help inform panelists' item ratings. These materials included the 2021 NE Standards in ELA (separated by grade level), the Depth of Knowledge Wheel, and the Achievement Level Descriptors (separated by grade level).

In addition to the references provided by NWEA and NDE, HumRRO developed electronic spreadsheets (i.e., Google Sheets) that panelists used to enter item ratings. Facilitators monitored each panelist's ratings in a main spreadsheet. HumRRO also provided training materials, including the panelist and facilitator instructions and training slides for panelists and facilitators. Additionally, HumRRO developed demographic and process evaluation surveys that were used to collect feedback from panelists on demographics, alignment training, and panel facilitation. HumRRO provided all workshop materials to panelists in electronic form through Google Drive.

## Training

Panelists participating in the alignment workshop received training before they began rating items. All panelists participated in a general training session led by HumRRO, which provided background on alignment, an overview of the study's alignment methodology and the item ratings to be collected during the workshop.

After the general training session, panelists were released to their grade-level panels for additional in-depth training conducted by their HumRRO facilitator. This training focused on the rating process and the procedures for accessing and using the reference materials to inform their ratings for each item. Panelists then calibrated their ratings with at least the first three items to ensure they shared a common understanding of each rating and used the same approach when evaluating items in the context of the ratings.

NWEA and NDE staff did not engage with panelists beyond the general training session to ensure independence of ratings. However, NWEA and NDE were available to answer panelist questions related to the 2021 NE Standards that HumRRO facilitators communicated via a Microsoft Teams chat and/or video call.

## Chapter 3: Alignment Results

This chapter summarizes the data and information collected during the Nebraska ELA alignment workshop. The majority agreement rating for each item was determined through a group discussion. Results are presented for each grade level panel on the following criteria:

**Table 8. Benchmark Evaluation Criteria by Test Event**

Criteria	Description	Benchmark
<b>Criterion 1:</b> Items Represent Intended Content	This criterion measured the alignment between the NE standards and test items on each test event.	<p><b>Met:</b> At least 75% of the NE Standards are assessed by items.</p> <p><b>Partially Met:</b> 50% - 74% of the NE Standards are assessed by items.</p> <p><b>Not Met:</b> Less than 50% of the NE Standards are assessed by items.</p>
<b>Criterion 2:</b> Items Represent Intended Categories	This criterion compared the expected distribution of items by content strand, as presented in the test blueprints, to the distribution of items on each test event.	<p><b>Met:</b> Nebraska content strands are +/- 5% from the minimum and maximum target values outlined in the blueprint.</p> <p><b>Not Met:</b> Nebraska content strands are <b>not</b> within +/- 5% of the minimum and maximum target values outlined in the blueprint.</p>
<b>Criterion 3:</b> Items Reflect Levels of Cognitive Complexity	This criterion focused on the cognitive complexity of items. The purpose of this criterion is to evaluate the type of cognitive processing required by items to examine the items' breadth of cognitive complexity using Webb's DOK.	<p><b>Met:</b> At least 70% of items are rated at cognitive complexity level 2 or above</p> <p><b>Not Met:</b> Less than 70% of items are rated at cognitive complexity level 2 or above.</p>
<b>Criterion 4:</b> Items Reflect Levels of Achievement Level Descriptors	This criterion focused on the range of achievement level descriptors (ALDs). Some states include this as additional complexity information in their Peer Review submission. Using well-defined ALDs is consistent with the principles of assessment design.	Items on each test event will reflect a range of ALDs. No specific evaluation benchmark was utilized.

We used four key documents to evaluate the alignment of the NSCAS in ELA with the respective NE Standards:

**Table 9. Key Evaluation Documents**

Key Documents	Description
Nebraska Standards in ELA*	This document lists all standards per grade level in ELA.
Test Blueprints*	This document lists the Nebraska Standards by content strand and the target item percentage by strand.
Cognitive Complexity Definitions	This document provides the cognitive complexity definitions, as defined by Webb.
Achievement Level Descriptors	This document provides the achievement level descriptors (ALDs), which describe the knowledge, skills, and processes that students demonstrate on state tests at pre-determined levels of achievement for each tested grade level.

\* It is important to note that the Reading Prose and Poetry (RP) and Reading Informational Text (RI) strands drill down to the standard level and Vocabulary (V) and Writing (W) drill down to the sub-standard level in both the NE Standards and Test Blueprints.

## Items Assigned to a Nebraska Standard

Tables 10-11 below describe the number and percentage of items assigned to an NE Standard. The data is disaggregated by grade level, representing how the alignment between items and standards varies across grades. It's worth highlighting that nearly all items assessed content found in the NE Standards across all grade levels, with percentages ranging from 97% to 99%.

**Table 10. Items Assigned to a Nebraska Standard – All Grades**

Grade	Number of Unique Items	Items Assigned to an NE Standard	
		#	%
ELA 3	246	243	99%
ELA 4	241	234	97%
ELA 5	230	228	99%
ELA 6	215	213	99%
ELA 7	226	223	99%
ELA 8	235	228	97%

**Table 11. Items NOT Assigned to a Nebraska Standard – All Grades**

Grade	Number of Unique Items	Items NOT Assigned to an NE Standard	
		#	%
ELA 3	246	3	1%
ELA 4	241	7	3%
ELA 5	230	2	1%
ELA 6	215	2	1%
ELA 7	226	3	1%
ELA 8	235	7	3%

## Criterion 1: Items Represent Intended Content

Criterion 1 examined the content alignment between items and NE Standards. We reviewed the extent to which items on each of the 12 CAT test events covered the intended NE Standards. For this criterion, we present results evaluating the breadth of NE Standards by grade, test event, and content strand.

**Table 12. Criterion 1 Evaluative Benchmark**

Benchmark	Description
<b>Met</b>	At least 75% of the NE Standards are assessed by items for each test event.
<b>Partially Met</b>	50%-74% of the NE Standards are assessed by items for each test event.
<b>Not Met</b>	Less than 50% of the NE Standards are assessed by items for each test event.

The NE Standards and test blueprint were the key documents used to evaluate this criterion. First and foremost, the test blueprint lists the content strands of the associated NE Standard that items should measure. The NE Standards are designed and written as discrete statements of the knowledge and skills a student should be taught in each subject and grade level. Locally assessed standards, Foundations of Reading standards, and Speaking and Listening standards were *not* included in the denominators. Additionally, it is important to note that the test blueprint details the Reading Prose and Poetry and Reading Informational Text strands at the *standard level* and the Vocabulary and Writing strands at the *sub-standard level*.

For each grade and each test event, we evaluated the alignment between the items and NE Standards by comparing the number of majority agreement final NE Standards to the number of NE Standards based on the content strands in the test blueprint. Some items were assigned more than one NE Standard by reviewers as the final majority agreement. All assigned NE Standards were included in the counts. As these analyses were based on the majority agreement, items for which reviewers could not identify a NE Standard or a majority agreement could not be reached were excluded from all counts. A detailed breakdown of these data by grade, test event, and blueprint content strands is provided in Tables 13-18. In general, we expected to find that the number of majority agreement NE Standards identified covered the overall range or breadth of blueprint NE Standards listed for each content strand.

In Grade 3, all test events contained items measuring at least half or more of the standards for Reading Prose and Poetry, except for one test event in the Developing achievement level where only three of seven standards were assessed by items. This was also the case for Reading Informational Text except for one test event in the Advanced achievement level where only two of seven standards were assessed by items. For the Vocabulary strand, two test events in the Developing achievement level and two test events in the Advanced achievement level had less than half of the standards measured by items. For the Writing strand, all test events had less than half of the standards measured by items. Across all test events, at least one content strand fell into the Not Met category.



**Table 13. Number of Standards Assessed by Test Event and Strand – Grade 3**

Achievement Level	Test Event	RP (7 Standards)	RI (7 Standards)	V (6 Standards)	W (20 Standards)	Summary Across Strands
Developing	1	5 (71%) - Partially Met	5 (71%) - Partially Met	2 (33%) - Not Met	4 (20%) - Not Met	2 of 4 – Not Met
	2	3 (43%) - Not Met	5 (71%) - Partially Met	2 (33%) - Not Met	4 (20%) - Not Met	3 of 4 – Not Met
	3	4 (57%) - Partially Met	6 (86%) - Met	4 (67%) - Partially Met	3 (15%) - Not Met	1 of 4 – Not Met
	4	5 (71%) - Partially Met	6 (86%) - Met	3 (50%) - Partially Met	3 (15%) - Not Met	1 of 4 – Not Met
On Track	1	5 (71%) - Partially Met	5 (71%) - Partially Met	2 (33%) - Partially Met	4 (20%) - Not Met	1 of 4 – Not Met
	2	5 (71%) - Partially Met	4 (57%) - Partially Met	3 (50%) - Partially Met	4 (20%) - Not Met	1 of 4 – Not Met
	3	5 (71%) - Partially Met	4 (57%) - Partially Met	3 (50%) - Partially Met	5 (25%) - Not Met	1 of 4 – Not Met
	4	5 (71%) - Partially Met	4 (57%) - Partially Met	3 (50%) - Partially Met	4 (20%) - Not Met	1 of 4 – Not Met
Advanced	1	4 (57%) - Partially Met	5 (71%) - Partially Met	4 (67%) - Partially Met	4 (20%) - Not Met	1 of 4 – Not Met
	2	6 (86%) - Met	5 (71%) - Partially Met	2 (33%) - Not Met	2 (10%) - Not Met	2 of 4 – Not Met
	3	4 (57%) - Partially Met	5 (71%) - Partially Met	2 (33%) - Not Met	4 (20%) - Not Met	2 of 4 – Not Met
	4	4 (57%) - Partially Met	2 (29%) - Not Met	3 (50%) - Partially Met	4 (20%) - Not Met	2 of 4 – Not Met
Summary Across Achievement Levels		1 of 12 – Not Met	1 of 12 – Not Met	4 of 12 – Not Met	12 of 12 – Not Met	–

In Grade 4, three test events in the Developing achievement level, two in the On Track achievement level, and three in the Advanced achievement level had less than half of the standards measured by items for Reading Prose and Poetry. For Reading Informational Text, two test events in the Developing achievement level and one test event in the On Track achievement level had less than half of the standards measured by items. For the Vocabulary strand, one test event from the Developing achievement level, two from the On Track achievement level, and two from the Advanced achievement level had less than half of the standards measured by items. For the Writing strand, all test events had less than half of the standards measured by items. Across all test events, at least one content strand fell into the Not Met category.

**Table 14. Number of Standards Assessed by Test Event and Strand – Grade 4**

Achievement Level	Test Event	RP (7 Standards)	RI (7 Standards)	V (5 Standards)	W (20 Standards)	Summary Across Strands
Developing	1	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	5 (25%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	2	3 (43%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	4 (80%) - <b>Met</b>	6 (30%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	3	5 (71%) - <b>Partially Met</b>	5 (71%) - <b>Partially Met</b>	4 (80%) - <b>Met</b>	3 (15%) - <b>Not Met</b>	<b>1 of 4</b> – Not Met
	4	2 (29%) - <b>Not Met</b>	2 (29%) - <b>Not Met</b>	2 (40%) - <b>Not Met</b>	5 (25%) - <b>Not Met</b>	<b>4 of 4</b> – Not Met
On Track	1	5 (71%) - <b>Partially Met</b>	5 (71%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	4 (20%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	2	4 (57%) - <b>Partially Met</b>	5 (71%) - <b>Partially Met</b>	4 (80%) - <b>Met</b>	6 (30%) - <b>Not Met</b>	<b>1 of 4</b> – Not Met
	3	2 (29%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - <b>Partially Met</b>	6 (30%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	4	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	5 (25%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
Advanced	1	3 (43%) - <b>Not Met</b>	5 (71%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	6 (30%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	2	4 (57%) - <b>Partially Met</b>	6 (86%) - <b>Met</b>	2 (40%) - <b>Not Met</b>	4 (20%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	3	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	5 (25%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	4	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	4 (80%) - <b>Met</b>	5 (25%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>		<b>8 of 12</b> – Not Met	<b>3 of 12</b> – Not Met	<b>5 of 12</b> – Not Met	<b>12 of 12</b> – Not Met	<b>–</b>

In Grade 5, one test event in the v achievement level and two in the Advanced achievement level had less than half of the standards measured by items for Reading Prose and Poetry. For the Reading Informational Text strand, only one test event in the Developing achievement level and one in the Advanced achievement level had less than half of the standards measured by items. For the Vocabulary strand, only one test event from the On Track achievement level had less than half of the standards measured by items. For the Writing strand, all test events had less than half of the standards measured by items. Across all test events, at least one content strand fell into the Not Met category.

**Table 15. Number of Standards Assessed by Test Event and Strand – Grade 5**

Achievement Level	Test Event	RP (7 Standards)	RI (7 Standards)	V (5 Standards)	W (19 Standards)	Summary Across Strands
Developing	1	5 (71%) - Partially Met	5 (71%) - Partially Met	3 (60%) - Partially Met	3 (16%) - <b>Not Met</b>	1 of 4 – Not Met
	2	5 (71%) - Partially Met	2 (29%) - <b>Not Met</b>	3 (60%) - Partially Met	4 (21%) - <b>Not Met</b>	2 of 4 – Not Met
	3	5 (71%) - Partially Met	6 (86%) - <b>Met</b>	3 (60%) - Partially Met	3 (16%) - <b>Not Met</b>	1 of 4 – Not Met
	4	6 (86%) - <b>Met</b>	5 (71%) - Partially Met	5 (100%) - <b>Met</b>	4 (21%) - <b>Not Met</b>	1 of 4 – Not Met
On Track	1	4 (57%) - Partially Met	4 (57%) - Partially Met	3 (60%) - Partially Met	3 (16%) - <b>Not Met</b>	1 of 4 – Not Met
	2	4 (57%) - Partially Met	5 (71%) - Partially Met	3 (60%) - Partially Met	4 (21%) - <b>Not Met</b>	1 of 4 – Not Met
	3	4 (57%) - Partially Met	6 (86%) - <b>Met</b>	5 (100%) - <b>Met</b>	3 (16%) - <b>Not Met</b>	1 of 4 – Not Met
	4	3 (43%) - <b>Not Met</b>	4 (57%) - Partially Met	2 (40%) - <b>Not Met</b>	3 (16%) - <b>Not Met</b>	3 of 4 – Not Met
Advanced	1	6 (86%) - <b>Met</b>	5 (71%) - Partially Met	4 (80%) - <b>Met</b>	4 (21%) - <b>Not Met</b>	1 of 4 – Not Met
	2	5 (71%) - Partially Met	6 (86%) - <b>Met</b>	4 (80%) - <b>Met</b>	5 (26%) - <b>Not Met</b>	1 of 4 – Not Met
	3	3 (43%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - Partially Met	5 (26%) - <b>Not Met</b>	3 of 4 – Not Met
	4	1 (14%) - <b>Not Met</b>	4 (57%) - Partially Met	4 (80%) - <b>Met</b>	5 (26%) - <b>Not Met</b>	2 of 4 – Not Met
<b>Summary Across Achievement Levels</b>		<b>3 of 12 – Not Met</b>	<b>2 of 12 – Not Met</b>	<b>1 of 12 – Not Met</b>	<b>12 of 12 – Not Met</b>	<b>–</b>

In Grade 6, one test event in the Developing achievement level and one in the On Track achievement level had less than half of the standards measured by items for Reading Prose and Poetry. For the Reading Informational Text strand, only one test event in the Developing achievement level and one in the On Track achievement level had less than half of the standards measured by items. For the Vocabulary strand, only one test event from the Developing achievement level had less than half of the standards measured by items. For the Writing strand, all test events had less than half of the standards measured by items. Across all test events, at least one content strand fell into the Not Met category.

**Table 16. Number of Standards Assessed by Test Event and Strand – Grade 6**

Achievement Level	Test Event	RP (7 Standards)	RI (7 Standards)	V (5 Standards)	W (22 Standards)	Summary Across Strands
Developing	1	4 (57%) - Partially Met	2 (29%) - <b>Not Met</b>	2 (40%) - <b>Not Met</b>	5 (23%) - <b>Not Met</b>	3 of 4 – Not Met
	2	4 (57%) - Partially Met	5 (71%) - Partially Met	3 (60%) - Partially Met	5 (23%) - <b>Not Met</b>	1 of 4 – Not Met
	3	3 (43%) - <b>Not Met</b>	4 (57%) - Partially Met	4 (80%) - <b>Met</b>	3 (14%) - <b>Not Met</b>	2 of 4 – Not Met
	4	5 (71%) - Partially Met	5 (71%) - Partially Met	4 (80%) - <b>Met</b>	5 (23%) - <b>Not Met</b>	1 of 4 – Not Met
On Track	1	3 (43%) - <b>Not Met</b>	4 (57%) - Partially Met	4 (80%) - <b>Met</b>	5 (23%) - <b>Not Met</b>	2 of 4 – Not Met
	2	5 (71%) - Partially Met	5 (71%) - Partially Met	4 (80%) - <b>Met</b>	4 (18%) - <b>Not Met</b>	1 of 4 – Not Met
	3	6 (86%) – <b>Met</b>	3 (43%) - <b>Not Met</b>	4 (80%) - <b>Met</b>	6 (27%) - <b>Not Met</b>	2 of 4 – Not Met
	4	4 (57%) - Partially Met	5 (71%) - Partially Met	4 (80%) - <b>Met</b>	4 (18%) - <b>Not Met</b>	1 of 4 – Not Met
Advanced	1	4 (57%) - Partially Met	4 (57%) - Partially Met	3 (60%) - Partially Met	4 (18%) - <b>Not Met</b>	1 of 4 – Not Met
	2	5 (71%) - Partially Met	5 (71%) - Partially Met	2 (40%) - Partially Met	4 (18%) - <b>Not Met</b>	1 of 4 – Not Met
	3	6 (86%) - <b>Met</b>	4 (57%) - Partially Met	3 (60%) - Partially Met	4 (18%) - <b>Not Met</b>	1 of 4 – Not Met
	4	4 (57%) - Partially Met	5 (71%) - Partially Met	3 (60%) - Partially Met	3 (14%) - <b>Not Met</b>	1 of 4 – Not Met
<b>Summary Across Achievement Levels</b>		<b>2 of 12 – Not Met</b>	<b>2 of 12 – Not Met</b>	<b>1 of 12 – Not Met</b>	<b>12 of 12 – Not Met</b>	<b>–</b>



In Grade 7, three test events in the Developing achievement level, two in the On Track achievement level, and three in the Advanced achievement level had less than half of the standards measured by items for Reading Prose and Poetry. For the Reading Informational Text strand, only one test event in the On Track achievement level, and one in the Advanced achievement level had less than half of the standards measured by items. For the Vocabulary strand, two test events from the Developing achievement level and all four test events from the On Track achievement level had less than half of the standards measured by items. For the Writing strand, all test events had less than half of the standards measured by items. Across all test events, at least one content strand fell into the Not Met category.

**Table 17. Number of Standards Assessed by Test Event and Strand – Grade 7**

Achievement Level	Test Event	RP (7 Standards)	RI (7 Standards)	V (5 Standards)	W (20 Standards)	Summary Across Strands
Developing	1	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	5 (25%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	2	4 (57%) - <b>Partially Met</b>	4 (57%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	5 (25%) - <b>Not Met</b>	<b>1 of 4</b> – Not Met
	3	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	4 (20%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	4	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	4 (20%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
On Track	1	2 (29%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	4 (20%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	2	4 (57%) - <b>Partially Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	5 (25%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	3	3 (43%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	2 (40%) - <b>Not Met</b>	5 (25%) - <b>Not Met</b>	<b>4 of 4</b> – Not Met
	4	4 (57%) - <b>Partially Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	5 (25%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
Advanced	1	4 (57%) - <b>Partially Met</b>	5 (71%) - <b>Partially Met</b>	4 (80%) - <b>Met</b>	4 (20%) - <b>Not Met</b>	<b>1 of 4</b> – Not Met
	2	2 (29%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - <b>Partially Met</b>	4 (20%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	3	2 (29%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	5 (25%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	4	3 (43%) - <b>Not Met</b>	6 (86%) - <b>Met</b>	4 (80%) - <b>Met</b>	4 (20%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>		<b>8 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>6 of 12</b> – Not Met	<b>12 of 12</b> – Not Met	<b>–</b>

In Grade 8, all test events had less than half of the standards measured by item except for one test event in the On Track achievement level. For the Reading Informational Text strand, two test events in the Developing achievement level, three in the On Track achievement level, and two in the Advanced achievement level had less than half of the standards measured by items. For the Vocabulary strand, one test event from the Developing achievement level and two from the Advanced achievement level had less than half of the standards measured by items. For the Writing strand, all test events had less than half of the standards measured by items. Across all test events, at least two content strands fell into the Not Met category, primarily in the Reading Prose and Poetry and Writing strands.

**Table 18. Number of Standards Assessed by Test Event and Strand – Grade 8**

Achievement Level	Test Event	RP (7 Standards)	RI (7 Standards)	V (5 Standards)	W (23 Standards)	Summary Across Strands
Developing	1	2 (29%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - <b>Partially Met</b>	5 (22%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	2	1 (14%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	4 (17%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	3	1 (14%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - <b>Partially Met</b>	4 (17%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	4	2 (29%) - <b>Not Met</b>	5 (71%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	6 (26%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
On Track	1	4 (57%) - <b>Partially Met</b>	3 (43%) - <b>Not Met</b>	4 (80%) - <b>Met</b>	5 (22%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	2	3 (43%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - <b>Partially Met</b>	5 (22%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	3	3 (43%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - <b>Partially Met</b>	6 (26%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	4	2 (29%) - <b>Not Met</b>	5 (71%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	5 (22%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
Advanced	1	2 (29%) - <b>Not Met</b>	3 (43%) - <b>Not Met</b>	3 (60%) - <b>Partially Met</b>	6 (26%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	2	3 (43%) - <b>Not Met</b>	4 (57%) - <b>Partially Met</b>	3 (60%) - <b>Partially Met</b>	6 (26%) - <b>Not Met</b>	<b>2 of 4</b> – Not Met
	3	1 (14%) - <b>Not Met</b>	5 (71%) - <b>Partially Met</b>	2 (40%) - <b>Not Met</b>	4 (17%) - <b>Not Met</b>	<b>3 of 4</b> – Not Met
	4	2 (29%) - <b>Not Met</b>	1 (14%) - <b>Not Met</b>	1 (20%) - <b>Not Met</b>	6 (26%) - <b>Not Met</b>	<b>4 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>		<b>11 of 12</b> – Not Met	<b>7 of 12</b> – Not Met	<b>3 of 12</b> – Not Met	<b>12 of 12</b> – Not Met	<b>–</b>

## Criterion 2: Items Represent Intended Categories

Criterion 2 examined how panelists' majority agreement ratings of items were distributed across content strands. Specifically, we compared the distribution of items using the majority agreement NE Standard compared with the test blueprint target. We generally expected that the majority agreement NE Standard selected for an item would match the content strand targets in the test blueprint.

**Table 19. Criterion 2 Evaluative Benchmark**

Criteria	Benchmark
Category Representation	<p><b>Met:</b> Nebraska content strands are +/- 5% from the minimum and maximum target values outlined in the blueprint for each test event.</p> <p><b>Not Met:</b> Nebraska content strands are <b>not</b> within +/- 5% of the minimum and maximum target values outlined in the blueprint for each test event.</p>

Table 20 presents the target percentage ranges for each strand, based on the test blueprint, and the target percentage ranges for each strand required for this criterion to be met. The target percentage ranges for this study are +/- 5% of the target percentage ranges noted in the test blueprints for each grade.

**Table 20. Category Representation Target Percentage Ranges for Test Blueprints and Study Criterion**

Grade	RP		RI		V		W				
	Blueprint	Study (+/- 5%)	Blueprint	Study (+/- 5%)	Blueprint	Study (+/- 5%)	Blueprint	Study (+/- 5%)			
ELA 3	28%-33%	23%-38%	28%-33%	23%-38%	15%-20%	10%-25%	23%-28%	18%-33%			
ELA 4											
ELA 5											
ELA 6											
ELA 7	25%-30%	20%-35%	30%-35%	25%-40%							
ELA 8											

In Grade 3, the Reading Prose and Poetry content strand had two test events in the Developing achievement level that did not meet the blueprint target. The Reading Informational Text content strand had one test event in the Developing achievement level, one in the On Track achievement level, and two in the Advanced achievement level that did not meet the blueprint target. The Vocabulary strand had two test events in the Developing achievement level and one test event in the On Track achievement level that did not meet the blueprint target. Across all three achievement levels, it is noteworthy that all test events successfully met the blueprint target for the Writing content strand.



**Table 21. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 3**

Achievement Level	Test Event	Number of Items	RP (23% - 38%)	RI (23% - 38%)	V (10% - 25%)	W (18% - 33%)	Summary Across Strands
Developing	1	30	27% - <b>Met</b>	40% - <b>Not Met</b>	17% - <b>Met</b>	20% - <b>Met</b>	<b>1 of 4</b> – Not Met
	2	32	22% - <b>Not Met</b>	31% - <b>Met</b>	28% - <b>Not Met</b>	19% - <b>Met</b>	<b>2 of 4</b> – Not Met
	3	32	22% - <b>Not Met</b>	34% - <b>Met</b>	25% - <b>Met</b>	19% - <b>Met</b>	<b>1 of 4</b> – Not Met
	4	31	26% - <b>Met</b>	29% - <b>Met</b>	26% - <b>Not Met</b>	19% - <b>Met</b>	<b>1 of 4</b> – Not Met
On Track	1	28	32% - <b>Met</b>	21% - <b>Not Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>1 of 4</b> – Not Met
	2	28	32% - <b>Met</b>	25% - <b>Met</b>	18% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	31	26% - <b>Met</b>	32% - <b>Met</b>	26% - <b>Not Met</b>	19% - <b>Met</b>	<b>1 of 4</b> – Not Met
	4	27	26% - <b>Met</b>	30% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
Advanced	1	31	29% - <b>Met</b>	39% - <b>Not Met</b>	13% - <b>Met</b>	19% - <b>Met</b>	<b>1 of 4</b> – Not Met
	2	28	32% - <b>Met</b>	29% - <b>Met</b>	18% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	29	28% - <b>Met</b>	24% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	4	28	36% - <b>Met</b>	18% - <b>Not Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>1 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>			<b>2 of 12</b> – Not Met	<b>4 of 12</b> – Not Met	<b>3 of 12</b> – Not Met	<b>0 of 12</b> – Not Met	<b>–</b>

Note.

- For the On Track achievement level, test events #1, #2, and #4 had one item rated as “None.”
- For the Advanced achievement level, test events #3 and #4 had one item rated as “None.”

In Grade 4, across all three achievement levels, it is noteworthy that all test events successfully met the blueprint target for the Reading Prose and Poetry, Vocabulary, and Writing content strands. The Reading Informational Text strand had three test events in the Developing achievement level and one test event in the On Track achievement level that did not meet the blueprint target.

**Table 22. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 4**

Achievement Level	Test Event	Number of Items	RP (23% - 38%)	RI (23% - 38%)	V (10% - 25%)	W (18% - 33%)	Summary Across Strands
Developing	1	28	29% - <b>Met</b>	29% - <b>Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	24	29% - <b>Met</b>	17% - <b>Not Met</b>	25% - <b>Met</b>	25% - <b>Met</b>	<b>1 of 4</b> – Not Met
	3	29	34% - <b>Met</b>	21% - <b>Not Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>1 of 4</b> – Not Met
	4	27	26% - <b>Met</b>	22% - <b>Not Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	<b>1 of 4</b> – Not Met
On Track	1	23	30% - <b>Met</b>	30% - <b>Met</b>	13% - <b>Met</b>	26% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	28	32% - <b>Met</b>	21% - <b>Not Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>1 of 4</b> – Not Met
	3	27	26% - <b>Met</b>	30% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
	4	27	33% - <b>Met</b>	26% - <b>Met</b>	15% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
Advanced	1	27	30% - <b>Met</b>	30% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	27	26% - <b>Met</b>	37% - <b>Met</b>	15% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	27	26% - <b>Met</b>	33% - <b>Met</b>	22% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
	4	29	28% - <b>Met</b>	24% - <b>Met</b>	28% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>			<b>0 of 12</b> – Not Met	<b>4 of 12</b> – Not Met	<b>0 of 12</b> – Not Met	<b>0 of 12</b> – Not Met	<b>–</b>

**Note.**

- For the Developing achievement level, test event #2 had one item rated as “None.”
- For the Developing achievement level, test event #4 had three items rated as “None.”
- For the On Track achievement level, test events #2, #3, and #4 had one item rated as “None.”
- For the Advanced achievement level, test event #2 had one item rated as “None.”

In Grade 5, the Reading Prose and Poetry strand had two test events in the Advanced achievement level that did not meet the blueprint target. The Reading Informational Text strand had one test event in the Developing achievement level and one in the Advanced achievement level that did not meet the blueprint target. The Vocabulary strand had two test events in the On Track achievement level and one test event in the Advanced achievement level that did not meet the blueprint target. Across all three achievement levels, it is noteworthy that all test events successfully met the blueprint target for the Writing content strand.

**Table 23. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 5**

Achievement Level	Test Event	Number of Items	RP (23% - 38%)	RI (23% - 38%)	V (10% - 25%)	W (18% - 33%)	Summary Across Strands
Developing	1	32	31% - <b>Met</b>	31% - <b>Met</b>	19% - <b>Met</b>	19% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	25	36% - <b>Met</b>	16% - <b>Not Met</b>	24% - <b>Met</b>	24% - <b>Met</b>	<b>1 of 4</b> – Not Met
	3	29	24% - <b>Met</b>	34% - <b>Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	4	32	34% - <b>Met</b>	25% - <b>Met</b>	22% - <b>Met</b>	19% - <b>Met</b>	<b>0 of 4</b> – Not Met
On Track	1	26	27% - <b>Met</b>	31% - <b>Met</b>	19% - <b>Met</b>	23% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	29	28% - <b>Met</b>	31% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	30	23% - <b>Met</b>	27% - <b>Met</b>	30% - <b>Not Met</b>	20% - <b>Met</b>	<b>1 of 4</b> – Not Met
	4	22	32% - <b>Met</b>	27% - <b>Met</b>	9% - <b>Not Met</b>	27% - <b>Met</b>	<b>1 of 4</b> – Not Met
Advanced	1	29	38% - <b>Met</b>	24% - <b>Met</b>	17% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	30	23% - <b>Met</b>	33% - <b>Met</b>	23% - <b>Met</b>	20% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	23	22% - <b>Not Met</b>	30% - <b>Met</b>	17% - <b>Met</b>	26% - <b>Met</b>	<b>1 of 4</b> – Not Met
	4	24	4% - <b>Not Met</b>	42% - <b>Not Met</b>	29% - <b>Not Met</b>	25% - <b>Met</b>	<b>3 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>			<b>2 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>3 of 12</b> – Not Met	<b>0 of 12</b> – Not Met	<b>–</b>

Note.

- For the On Track achievement level, test event #4 had one item rated as “None.”
- For the Advanced achievement level, test event #3 had one item rated as “None.”

In Grade 6, the Reading Prose and Poetry content strand only had one test event in the Developing achievement level that did not meet the blueprint target. The Reading Informational Text content strand had one test event in the Developing achievement level and two in the On Track achievement level that did not meet the blueprint target. The Vocabulary strand had two test events in the On Track achievement level that did not meet the blueprint target. Across all three achievement levels, it is noteworthy that all test events successfully met the blueprint target for the Writing content strand.

**Table 24. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 6**

Achievement Level	Test Event	Number of Items	RP (23% - 38%)	RI (23% - 38%)	V (10% - 25%)	W (18% - 33%)	Summary Across Strands
Developing	1	23	43% - <b>Not Met</b>	13% - <b>Not Met</b>	17% - <b>Met</b>	26% - <b>Met</b>	<b>2 of 4</b> – Not Met
	2	28	25% - <b>Met</b>	32% - <b>Met</b>	18% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	29	24% - <b>Met</b>	31% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	4	29	28% - <b>Met</b>	31% - <b>Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
On Track	1	28	25% - <b>Met</b>	21% - <b>Not Met</b>	29% - <b>Not Met</b>	21% - <b>Met</b>	<b>2 of 4</b> – Not Met
	2	30	33% - <b>Met</b>	23% - <b>Met</b>	23% - <b>Met</b>	20% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	27	30% - <b>Met</b>	22% - <b>Not Met</b>	26% - <b>Not Met</b>	22% - <b>Met</b>	<b>2 of 4</b> – Not Met
	4	28	29% - <b>Met</b>	29% - <b>Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
Advanced	1	29	24% - <b>Met</b>	31% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	29	24% - <b>Met</b>	31% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	26	31% - <b>Met</b>	27% - <b>Met</b>	19% - <b>Met</b>	23% - <b>Met</b>	<b>0 of 4</b> – Not Met
	4	28	29% - <b>Met</b>	32% - <b>Met</b>	18% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>			<b>1 of 12</b> – Not Met	<b>3 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>0 of 12</b> – Not Met	<b>–</b>

Note.

- For the Developing achievement level, test event #2 had one item rated as “None.”
- For the On Track achievement level, test event #1 had one item rated as “None.”



In Grade 7, the Reading Prose and Poetry content strand only had two test events in the On Track achievement level that did not meet the blueprint target. Across all three achievement levels, it is noteworthy that all test events successfully met the blueprint target for the Reading Informational Text content strand. The Vocabulary content strand only had one test event in the Developing achievement level that did not meet the blueprint target. The Writing strand had two test events in the Developing achievement level and one test event in the Advanced achievement level that did not meet the blueprint target.

**Table 25. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 7**

Achievement Level	Test Event	Number of Items	RP (20% - 35%)	RI (25% - 40%)	V (10% - 25%)	W (18% - 33%)	Summary Across Strands
Developing	1	31	23% - <b>Met</b>	35% - <b>Met</b>	23% - <b>Met</b>	16% - <b>Not Met</b>	1 of 4 – Not Met
	2	29	24% - <b>Met</b>	31% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	0 of 4 – Not Met
	3	28	25% - <b>Met</b>	36% - <b>Met</b>	21% - <b>Met</b>	18% - <b>Met</b>	0 of 4 – Not Met
	4	31	26% - <b>Met</b>	26% - <b>Met</b>	29% - <b>Not Met</b>	16% - <b>Not Met</b>	2 of 4 – Not Met
On Track	1	23	17% - <b>Not Met</b>	39% - <b>Met</b>	17% - <b>Met</b>	26% - <b>Met</b>	1 of 4 – Not Met
	2	27	37% - <b>Not Met</b>	26% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	1 of 4 – Not Met
	3	28	25% - <b>Met</b>	36% - <b>Met</b>	18% - <b>Met</b>	21% - <b>Met</b>	0 of 4 – Not Met
	4	27	30% - <b>Met</b>	30% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	0 of 4 – Not Met
Advanced	1	27	26% - <b>Met</b>	33% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	0 of 4 – Not Met
	2	23	22% - <b>Met</b>	30% - <b>Met</b>	22% - <b>Met</b>	17% - <b>Not Met</b>	1 of 4 – Not Met
	3	27	30% - <b>Met</b>	33% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	0 of 4 – Not Met
	4	27	26% - <b>Met</b>	33% - <b>Met</b>	19% - <b>Met</b>	19% - <b>Met</b>	0 of 4 – Not Met
<b>Summary Across Achievement Levels</b>			<b>2 of 12 – Not Met</b>	<b>0 of 12 – Not Met</b>	<b>1 of 12 – Not Met</b>	<b>3 of 12 – Not Met</b>	<b>–</b>

Note.

- For the Developing achievement level, test events #1, #3, and #4 had one item rated as “None.”
- For the Advanced achievement level, test event #2 had two items rated as “None.”
- For the Advanced achievement level, test event #4 had one item rated as “None.”

In Grade 8, the Reading Prose and Poetry content strand had three test events in the Developing achievement level and one in the Advanced achievement level that did not meet the blueprint target. The Reading Informational Text content strand had one test event in the Developing achievement level and one in the On Track achievement level that did not meet the blueprint target. The Vocabulary strand had one test event in the Developing achievement level and one test event in the On Track achievement level that did not meet the blueprint target. Across all three achievement levels, it is noteworthy that all test events successfully met the blueprint target for the Writing content strand.

**Table 26. Category Representation Percentage Ranges for Study Criterion by Strand and Test Event – Grade 8**

Achievement Level	Test Event	Number of Items	RP (20% - 35%)	RI (25% - 40%)	V (10% - 25%)	W (18% - 33%)	Summary Across Strands
Developing	1	29	38% - <b>Not Met</b>	21% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>1 of 4</b> – Not Met
	2	30	37% - <b>Not Met</b>	30% - <b>Met</b>	13% - <b>Met</b>	23% - <b>Met</b>	<b>1 of 4</b> – Not Met
	3	24	17% - <b>Not Met</b>	21% - <b>Not Met</b>	25% - <b>Met</b>	25% - <b>Met</b>	<b>2 of 4</b> – Not Met
	4	32	22% - <b>Met</b>	28% - <b>Met</b>	31% - <b>Not Met</b>	19% - <b>Met</b>	<b>1 of 4</b> – Not Met
On Track	1	28	32% - <b>Met</b>	25% - <b>Met</b>	18% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	22	32% - <b>Met</b>	14% - <b>Not Met</b>	23% - <b>Met</b>	27% - <b>Met</b>	<b>1 of 4</b> – Not Met
	3	29	24% - <b>Met</b>	31% - <b>Met</b>	24% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	4	30	23% - <b>Met</b>	27% - <b>Met</b>	27% - <b>Not Met</b>	20% - <b>Met</b>	<b>1 of 4</b> – Not Met
Advanced	1	28	29% - <b>Met</b>	25% - <b>Met</b>	21% - <b>Met</b>	21% - <b>Met</b>	<b>0 of 4</b> – Not Met
	2	27	26% - <b>Met</b>	33% - <b>Met</b>	19% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
	3	25	16% - <b>Not Met</b>	36% - <b>Met</b>	24% - <b>Met</b>	24% - <b>Met</b>	<b>1 of 4</b> – Not Met
	4	27	30% - <b>Met</b>	26% - <b>Met</b>	15% - <b>Met</b>	22% - <b>Met</b>	<b>0 of 4</b> – Not Met
<b>Summary Across Achievement Levels</b>			<b>4 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>0 of 12</b> – Not Met	<b>–</b>

*Note.*

- For the Developing achievement level, test event #3 had three items rated as “None.”
- For the On Track achievement level, test events #1, #2, and #4 had one item rated as “None.”
- For the Advanced achievement level, test event #1 had one item rated as “None.”
- For the Advanced achievement level, test event #4 had two items rated as “None.”

### Criterion 3: Items Reflect Levels of Cognitive Complexity

Criterion 3 is evaluated based on the percentage of items rated by panelists as reflecting each of the three cognitive complexity levels. The criterion is considered “Met” if 70% of items are rated at cognitive complexity level 2 or above. Table 27 provides a brief definition of Webb’s DOK Levels.

**Table 27. Criterion 3 Evaluative Benchmark**

Criterion	Benchmark
DOK Representation	<p><b>Met:</b> 70% of items are rated at a cognitive complexity level 2 or above.</p> <p><b>Not Met:</b> Less than 70% of items are rated at a cognitive complexity level 2 or above.</p>

**Table 28. Depth of Knowledge Levels and Definitions**

Webb’s DOK Levels	Definition
<b>Level 1:</b> Recall and Reproduction	Requires recall of information, such as a fact, definition, term, simple procedure, or property. Typically, it involves only one step.
<b>Level 2:</b> Skill/Concept	Requires some mental processing beyond recalling or reproducing a response. Typically, it involves more than one step.
<b>Level 3:</b> Strategic Thinking	Requires deep knowledge using reasoning, planning, or using evidence. Typically, has more than one possible answer and requires students to justify their response.
<b>Level 4:</b> Extended Thinking	Requires high cognitive demand and is very complex. Typically includes complex reasoning, experimental design, and planning, and will likely require an extended period of time.

Table 29 summarizes the number of items and their distribution across the Depth of Knowledge (DOK) levels in Grades 3-8. The data shows that items were predominantly aligned with DOK level 2 across all grades, representing 63% to 85% of the items. DOK 1 represented 5% to 26% of the items, and DOK 3 represented 4% to 26%.

**Table 29. Distribution of Depth of Knowledge Levels – All Grades**

Grade	Number of Items	DOK 1		DOK 2		DOK 3	
		#	%	#	%	#	%
ELA 3	246	17	7%	170	69%	59	24%
ELA 4	241	62	26%	152	63%	27	11%
ELA 5	230	23	10%	196	85%	10	4%
ELA 6	215	20	9%	156	73%	37	17%
ELA 7	226	12	5%	153	68%	59	26%
ELA 8	235	27	11%	190	81%	17	7%

In Grade 3, all test events met the evaluative benchmark of 70% or more of the items rated at cognitive complexity level 2 or above.

**Table 30. DOK Assessed by Test Event – Grade 3**

Achievement Level	Test Event	Number of Items	DOK 1	DOK 2	DOK 3	Benchmark
Developing	1	30	1 (3%)	25 (83%)	4 (13%)	<b>Met</b>
	2	32	5 (16%)	21 (66%)	6 (19%)	<b>Met</b>
	3	32	1 (3%)	24 (75%)	7 (22%)	<b>Met</b>
	4	31	3 (10%)	17 (55%)	11 (35%)	<b>Met</b>
On Track	1	28	1 (4%)	18 (64%)	9 (32%)	<b>Met</b>
	2	28	0 (0%)	23 (82%)	5 (18%)	<b>Met</b>
	3	31	2 (6%)	24 (77%)	5 (16%)	<b>Met</b>
	4	27	1 (4%)	22 (81%)	4 (15%)	<b>Met</b>
Advanced	1	31	3 (10%)	21 (68%)	7 (23%)	<b>Met</b>
	2	28	1 (4%)	19 (68%)	8 (29%)	<b>Met</b>
	3	29	0 (0%)	23 (79%)	6 (21%)	<b>Met</b>
	4	28	1 (4%)	23 (82%)	4 (14%)	<b>Met</b>

*Note.* There was one test item in the Developing 1 and On Track 3 test event that was split between RI and RI; it received a DOK of 2.

In Grade 4, all test events for the On Track and Advanced achievement levels met the evaluative benchmark of 70% of items rated at cognitive complexity level 2 or above. However, test event #2 in the On Track achievement level, while meeting the benchmark, had zero DOK 3 items administered compared to one to seven items on all other test events. Only test event #4 met the evaluative benchmark for the Developing achievement level. Additionally, within the Developing achievement level, test events #1 through #3 displayed a considerably higher alignment with DOK level 1 compared to the other test events.

**Table 31. DOK Assessed by Test Event – Grade 4**

Achievement Level	Test Event	Number of Items	DOK 1	DOK 2	DOK 3	Benchmark
Developing	1	28	12 (43%)	15 (54%)	1 (4%)	<b>Not Met</b>
	2	24	10 (42%)	13 (54%)	1 (4%)	<b>Not Met</b>
	3	29	10 (34%)	16 (55%)	3 (10%)	<b>Not Met</b>
	4	27	7 (26%)	15 (56%)	5 (19%)	<b>Met</b>
On Track	1	23	6 (26%)	14 (61%)	3 (13%)	<b>Met</b>
	2	28	6 (21%)	22 (79%)	<b>0 (0%)</b>	<b>Met</b>
	3	27	5 (19%)	18 (67%)	4 (15%)	<b>Met</b>
	4	27	5 (19%)	20 (74%)	2 (7%)	<b>Met</b>
Advanced	1	27	3 (11%)	22 (81%)	2 (7%)	<b>Met</b>
	2	27	4 (15%)	18 (67%)	5 (19%)	<b>Met</b>
	3	27	3 (11%)	17 (63%)	7 (26%)	<b>Met</b>
	4	29	7 (24%)	19 (66%)	3 (10%)	<b>Met</b>

*Note.* There was one test item in the Advanced 2 and Advanced 3 test events that was split between RI and RI; it received a DOK of 3.



In Grade 5, all test events met the evaluative benchmark of 70% of items rated at cognitive complexity level 2 or above. However, there were two test events in the Developing achievement level where zero DOK 3 items were administered compared to one or two items on all the other test events.

**Table 32. DOK Assessed by Test Event – Grade 5**

Achievement Level	Test Event	Number of Items	DOK 1	DOK 2	DOK 3	Benchmark
Developing	1	32	3 (9%)	29 (91%)	0 (0%)	Met
	2	25	2 (8%)	21 (84%)	2 (8%)	Met
	3	29	3 (10%)	24 (83%)	2 (7%)	Met
	4	32	5 (16%)	27 (84%)	0 (0%)	Met
On Track	1	26	2 (8%)	23 (88%)	1 (4%)	Met
	2	29	2 (7%)	26 (90%)	1 (3%)	Met
	3	30	3 (10%)	26 (87%)	1 (3%)	Met
	4	22	1 (5%)	19 (86%)	1 (5%)	Met
Advanced	1	29	2 (7%)	26 (90%)	1 (3%)	Met
	2	30	2 (7%)	26 (87%)	2 (7%)	Met
	3	23	1 (4%)	21 (91%)	1 (4%)	Met
	4	24	2 (8%)	21 (88%)	1 (4%)	Met

*Note.* For the On Track achievement level in test event #4, there was one item that was "None" for DOK, hence 22 test items.

One test item in the On Track 2 test event was split between RP and RP; it received a DOK of 2.

In Grade 6, all test events met the evaluative benchmark of 70% of items rated at cognitive complexity level 2 or above.

**Table 33. DOK Assessed by Test Event – Grade 6**

Achievement Level	Test Event	Number of Items	DOK 1	DOK 2	DOK 3	Benchmark
Developing	1	23	1 (4%)	18 (78%)	4 (17%)	<b>Met</b>
	2	28	7 (25%)	17 (61%)	4 (14%)	<b>Met</b>
	3	29	6 (21%)	18 (62%)	5 (17%)	<b>Met</b>
	4	29	4 (14%)	22 (76%)	3 (10%)	<b>Met</b>
On Track	1	28	0 (0%)	23 (82%)	5 (18%)	<b>Met</b>
	2	30	5 (17%)	22 (73%)	3 (10%)	<b>Met</b>
	3	27	0 (0%)	23 (85%)	4 (15%)	<b>Met</b>
	4	28	1 (4%)	18 (64%)	9 (32%)	<b>Met</b>
Advanced	1	29	0 (0%)	22 (76%)	7 (24%)	<b>Met</b>
	2	29	0 (0%)	21 (72%)	8 (28%)	<b>Met</b>
	3	26	0 (0%)	20 (77%)	6 (23%)	<b>Met</b>
	4	28	5 (18%)	17 (61%)	6 (21%)	<b>Met</b>

In Grade 7, all test events met the evaluative benchmark of 70% of items rated at cognitive complexity level 2 or above.

**Table 34. DOK Assessed by Test Event – Grade 7**

Achievement Level	Test Event	Number of Items	DOK 1	DOK 2	DOK 3	Benchmark
Developing	1	31	4 (13%)	21 (68%)	5 (16%)	Met
	2	29	2 (7%)	21 (72%)	6 (21%)	Met
	3	28	1 (4%)	19 (68%)	7 (25%)	Met
	4	31	2 (6%)	21 (68%)	8 (26%)	Met
On Track	1	23	2 (9%)	11 (48%)	9 (39%)	Met
	2	27	0 (0%)	19 (70%)	8 (30%)	Met
	3	28	2 (7%)	22 (79%)	4 (14%)	Met
	4	27	0 (0%)	18 (67%)	9 (33%)	Met
Advanced	1	27	1 (4%)	18 (67%)	8 (30%)	Met
	2	23	0 (0%)	19 (83%)	4 (17%)	Met
	3	27	0 (0%)	17 (63%)	10 (37%)	Met
	4	27	0 (0%)	20 (74%)	7 (26%)	Met

**Notes.**

- For the Developing achievement level test event #1, one item was “None” for DOK, hence 31 test items.
- For the Developing achievement level test event #3, one item was “None” for DOK, hence 28 test items.
- For the On Track achievement level test event #1, one item was “None” for DOK, hence 23 test items.
- One test item in the Developing 3 test event was split between RI and RI; it did not receive a DOK rating.
- One test item in the On Track 2 test event was split between RP and V; it received a DOK of 2.
- One test item in the Advanced 3 test event was split between RP and RP; it received a DOK of 3.

In Grade 8, all test events met the evaluative benchmark of 70% of items rated at cognitive complexity level 2 or above. However, there was one test event in the Developing achievement level and one in the On Track achievement level where zero DOK 3 items were administered compared to one to four items on all the other test events.

**Table 35. DOK Assessed by Test Event – Grade 8**

Achievement Level	Test Event	Number of Items	DOK 1	DOK 2	DOK 3	Benchmark
Developing	1	29	2 (7%)	26 (90%)	1 (3%)	Met
	2	30	5 (17%)	25 (83%)	0 (0%)	Met
	3	24	2 (8%)	17 (71%)	4 (17%)	Met
	4	32	6 (19%)	23 (72%)	3 (9%)	Met
On Track	1	28	2 (7%)	26 (93%)	0 (0%)	Met
	2	22	0 (0%)	18 (82%)	4 (18%)	Met
	3	29	3 (10%)	23 (79%)	3 (10%)	Met
	4	30	2 (7%)	25 (83%)	3 (10%)	Met
Advanced	1	28	3 (11%)	22 (79%)	3 (11%)	Met
	2	27	2 (7%)	22 (81%)	3 (11%)	Met
	3	25	1 (4%)	22 (88%)	2 (8%)	Met
	4	27	5 (19%)	20 (74%)	1 (4%)	Met

**Notes.**

- For the Developing achievement level test event #3, one item was “None” for DOK, hence 24 test items.
- For the Advanced achievement level test event #4, one item was “None” for DOK, hence 27 test items.
- One test item in the Developing 1 test event was split between RP and RP; it received a DOK of 2.
- One test item in the Developing 2 test event was split between W and W; it received a DOK of 2.

DOK data disaggregated by grade, test event, and strand are located in Appendix J.

## Criterion 4: Items Reflect Achievement Level Descriptors

Achievement level descriptors (ALDs) describe the knowledge, skills, and processes that students demonstrate on state tests at pre-determined levels of achievement for each tested grade level. The Nebraska State Board of Education defines three achievement levels:

1. Developing
2. On Track
3. Advanced

**Table 36. Achievement Level Descriptor Definitions**

ALD Levels	Definition
<b>Level 1:</b> Developing	Developing learners do not yet demonstrate proficiency in the knowledge and skills necessary at this grade level, as specified in the assessed Nebraska College and Career Ready Standards. These results prove that the student may need additional support for academic success at the next grade level.
<b>Level 2:</b> On Track	On Track learners demonstrate proficiency in the knowledge and skills necessary at this grade level, as specified in the assessed Nebraska College and Career Ready Standards. These results prove that the student will likely be ready for academic success at the next grade level.
<b>Level 3:</b> Advanced	Advanced learners demonstrate proficiency in the knowledge and skills necessary at this grade level, as specified in the assessed Nebraska College and Career Ready Standards. These results prove that the student will likely be ready for academic success at the next grade level.

Table 37 describes the number of items and their distribution across the ALDs in Grades 3-8. The data shows that across all grades, items were predominantly aligned with ALD level 2, representing 58% to 80% of the items. ALD 1 represented 8% to 27% of the items, and ALD 3 represented 2% to 26%.

**Table 37. Distribution of Achievement Level Descriptors – All Grades**

Grade	Number of Items	ALD 1		ALD 2		ALD 3	
		#	%	#	%	#	%
ELA 3	246	20	8%	161	65%	64	26%
ELA 4	241	66	27%	139	58%	29	12%
ELA 5	230	42	18%	183	80%	4	2%
ELA 6	215	39	18%	157	73%	16	7%
ELA 7	226	47	21%	138	61%	40	18%
ELA 8	235	39	17%	172	73%	21	9%

In Grade 3, items were classified into the three categories as follows: ALD 1 (ranging from 3% to 19%), ALD 2 (ranging from 46% to 77%), and ALD 3 (ranging from 16% to 43%).

**Table 38. ALD Assessed by Test Event – Grade 3**

Achievement Level	Test Event	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	30	2 (7%)	23 (77%)	5 (17%)
	2	32	6 (19%)	21 (66%)	5 (16%)
	3	32	1 (3%)	22 (69%)	9 (28%)
	4	31	2 (6%)	23 (74%)	6 (19%)
On Track	1	28	3 (11%)	13 (46%)	12 (43%)
	2	28	3 (11%)	20 (71%)	5 (18%)
	3	31	3 (10%)	23 (74%)	5 (16%)
	4	27	1 (4%)	20 (74%)	6 (22%)
Advanced	1	31	1 (3%)	20 (65%)	10 (32%)
	2	28	3 (11%)	15 (54%)	10 (36%)
	3	29	1 (3%)	19 (66%)	9 (31%)
	4	28	2 (7%)	20 (71%)	5 (18%)

*Note.* For the Advanced achievement level test event #4, one item was "None," hence 28 test items.

One test item in the Developing 1 and On Track 3 test events was split between RI and RI; it received an ALD of 2.

In Grade 4, items were classified into the three categories as follows: ALD 1 (ranging from 19% to 39%), ALD 2 (ranging from 52% to 70%), and ALD 3 (ranging from 4% to 19%).

**Table 39. ALD Assessed by Test Event – Grade 4**

Achievement Level	Test Event	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	28	9 (32%)	17 (61%)	2 (7%)
	2	24	8 (33%)	14 (58%)	1 (4%)
	3	29	8 (28%)	18 (62%)	3 (10%)
	4	27	7 (26%)	14 (52%)	3 (11%)
On Track	1	23	9 (39%)	12 (52%)	2 (9%)
	2	28	8 (29%)	16 (57%)	3 (11%)
	3	27	7 (26%)	14 (52%)	5 (19%)
	4	27	8 (30%)	14 (52%)	4 (15%)
Advanced	1	27	9 (33%)	16 (59%)	2 (7%)
	2	27	5 (19%)	19 (70%)	2 (7%)
	3	27	5 (19%)	18 (67%)	4 (15%)
	4	29	7 (24%)	18 (62%)	4 (14%)

**Notes.**

- For the Developing achievement level test event #2, one item was "None," hence 24 test items.
- For the Developing achievement level test event #4, three items were "None," hence 27 test items.
- For the On Track achievement level test event #2, there was one item that was "None," hence 28 test items.
- For the On Track achievement level test event #3, one item was "None," hence 27 test items.
- For the On Track achievement level test event #4, one item was "None," hence 27 test items.
- For the Advanced achievement level test event #2, one item was "None," hence 27 test items.
- One test item in the Advanced 2 and Advanced 3 test events was split between RI and RI; it received an ALD of 2.

In Grade 5, items were classified into the three categories as follows: ALD 1 (ranging from 7% to 28%), ALD 2 (ranging from 72% to 93%), and ALD 3 (ranging from 0% to 5%). Overall, there were few items assigned to ALD 3 across all test events.

**Table 40. ALD Assessed by Test Event – Grade 5**

Achievement Level	Test Event	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	32	4 (13%)	28 (88%)	0 (0%)
	2	25	4 (16%)	21 (84%)	0 (0%)
	3	29	7 (24%)	21 (72%)	1 (3%)
	4	32	7 (22%)	25 (78%)	0 (0%)
On Track	1	26	4 (15%)	22 (85%)	0 (0%)
	2	29	7 (24%)	22 (76%)	0 (0%)
	3	30	2 (7%)	28 (93%)	0 (0%)
	4	22	4 (18%)	16 (73%)	1 (5%)
Advanced	1	29	8 (28%)	21 (72%)	0 (0%)
	2	30	4 (13%)	25 (83%)	1 (3%)
	3	23	6 (26%)	17 (74%)	0 (0%)
	4	24	4 (17%)	19 (79%)	1 (4%)

*Note.* For the On Track achievement level test event #4, there was one item that was “None,” hence 22 test items. One test item in the On Track 2 test event was split between RP and RP; it received an ALD of 1.



In Grade 6, items were classified into the three categories as follows: ALD 1 (ranging from 4% to 36%), ALD 2 (ranging from 61% to 92%), and ALD 3 (ranging from 0% to 21%). Two test events were in the Developing achievement level where no items were assigned an ALD 3.

**Table 41. ALD Assessed by Test Event – Grade 6**

Achievement Level	Test Event	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	23	3 (13%)	20 (87%)	0 (0%)
	2	28	10 (36%)	17 (61%)	0 (0%)
	3	29	8 (28%)	19 (66%)	2 (7%)
	4	29	6 (21%)	21 (72%)	2 (7%)
On Track	1	28	3 (11%)	22 (79%)	3 (11%)
	2	30	4 (13%)	24 (80%)	2 (7%)
	3	27	2 (7%)	24 (89%)	1 (4%)
	4	28	1 (4%)	21 (75%)	6 (21%)
Advanced	1	29	5 (17%)	23 (79%)	1 (3%)
	2	29	4 (14%)	21 (72%)	4 (14%)
	3	26	1 (4%)	24 (92%)	1 (4%)
	4	28	5 (18%)	21 (75%)	2 (7%)

*Note.* For the Developing achievement level test event #2, there was one item that was “None,” hence 28 test items.

In Grade 7, items were classified into the three categories as follows: ALD 1 (ranging from 11% to 42%), ALD 2 (ranging from 45% to 76%), and ALD 3 (ranging from 10% to 30%).

**Table 42. ALD Assessed by Test Event – Grade 7**

Achievement Level	Test Event	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	31	13 (42%)	14 (45%)	4 (13%)
	2	29	4 (14%)	22 (76%)	3 (10%)
	3	28	5 (18%)	15 (54%)	7 (25%)
	4	31	9 (29%)	15 (48%)	7 (23%)
On Track	1	23	4 (17%)	12 (52%)	7 (30%)
	2	27	6 (22%)	17 (63%)	4 (15%)
	3	28	6 (21%)	19 (68%)	3 (11%)
	4	27	6 (22%)	18 (67%)	3 (11%)
Advanced	1	27	5 (19%)	16 (59%)	6 (22%)
	2	23	3 (13%)	14 (61%)	6 (26%)
	3	27	3 (11%)	18 (67%)	6 (22%)
	4	27	4 (15%)	20 (74%)	3 (11%)

**Notes.**

- For the Developing achievement level test event #3, one item was “None,” hence 28 test items.
- One test item in the Developing 3 test event was split between RI and RI; it did not receive an ALD rating.
- One test item in the On Track 2 test event was split between RP and V; it received an ALD of 3.
- One test item in the Advanced 3 test event was split between RP and RP; it received an ALD of 3.

In Grade 8, items were classified into the three categories as follows: ALD 1 (ranging from 4% to 33%), ALD 2 (ranging from 63% to 89%), and ALD 3 (ranging from 0% to 20%). One test event in the On Track achievement level was assigned no ALD 3 items.

**Table 43. ALD Assessed by Test Event – Grade 8**

Achievement Level	Test Event	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	29	5 (17%)	22 (76%)	2 (7%)
	2	30	10 (33%)	19 (63%)	1 (3%)
	3	24	4 (17%)	18 (75%)	2 (8%)
	4	32	7 (22%)	23 (72%)	2 (6%)
On Track	1	28	2 (7%)	25 (89%)	<b>0</b> <b>(0%)</b>
	2	22	2 (9%)	16 (73%)	3 (14%)
	3	29	3 (10%)	24 (83%)	2 (7%)
	4	30	5 (17%)	23 (77%)	2 (7%)
Advanced	1	28	1 (4%)	22 (79%)	4 (14%)
	2	27	1 (4%)	23 (85%)	3 (11%)
	3	25	2 (8%)	18 (72%)	5 (20%)
	4	27	5 (19%)	19 (70%)	3 (11%)

**Notes.**

- For the On Track achievement level test event #1, one item was “None,” hence 28 test items.
- For the On Track achievement level test event #2, one item was “None,” hence 22 test items.
- For the Advanced achievement level, test event #1 had one item that was “None,” hence 28 test items.
- One test item in the Developing 1 test event was split between RP and RP; it received an ALD of 1.
- One test item in the Developing 2 test event was split between W and W; it received an ALD of 2.

ALD data disaggregated by grade, test event, and strand are located in Appendix K.

## Process Evaluation Results

Upon adjourning each panel, HumRRO facilitators administered a process evaluation survey to their panelists.

Overall, panelists evaluated the workshop with high levels of satisfaction (Table 44). On a scale of 1-5, with 1 = “Strongly Disagree” to 5 = “Strongly Agree,” most panelists believed their facilitator did an effective job of facilitating discussion and ensuring all panelists’ perspectives were heard (average = 4.90), the facilitators clearly and promptly addressed questions (average = 4.90), and the facilitator was helpful during the workshop (average = 4.83). Notably, across all panels, 50% of educators reported strong alignment of items with the NE Standards, while the remaining 50% reported partial alignment.

Appendix L provides the complete results of this survey disaggregated by grade level.

**Table 44. Panelist Evaluation Survey Results – All Grades**

Question	Average	
My panel facilitator did an effective job of facilitating discussion and ensuring that all panelists' perspectives were heard	4.90	<div></div>
My panel facilitator clearly and promptly addressed my questions	4.90	<div></div>
The panel facilitator was helpful during the workshop	4.83	<div></div>
My ideas and opinions were listened to and respected by the group	4.77	<div></div>
Everyone had equal opportunity to contribute ideas and opinions	4.77	<div></div>
The materials hosted on Google Drive were useful (e.g., standards)	4.73	<div></div>
Practicing making ratings as a group in my assigned panel helped me better understand the alignment activities	4.73	<div></div>
The hands-on training in my assigned panel was an effective use of time	4.70	<div></div>
The panel-specific hands-on training was well organized	4.63	<div></div>
The other materials shared by my facilitator were useful	4.63	<div></div>
The Google Rating Sheet was useful for recording alignment ratings	4.60	<div></div>
The Google Rating Sheet provided a comprehensive platform for capturing alignment of standards	4.60	<div></div>
It was easy to access the evaluation and demographics forms	4.60	<div></div>
The whole-group training facilitator was helpful during the workshop	4.57	<div></div>
The hands-on training in my assigned panel helped me better understand the alignment activities	4.50	<div></div>
Other support staff were helpful during the workshop	4.47	<div></div>
The group-wide training session was well organized	4.41	<div></div>
The group-wide training session provided a useful overview of the alignment activities for the week	4.37	<div></div>
The group-wide training session effectively outlined the purpose of the alignment workshop	4.37	<div></div>
The group-wide training session clearly described my role as a panelist	4.37	<div></div>
The Content Review Tool allowed me to effectively accomplish my tasks during the workshop	4.37	<div></div>
The group-wide training was an effective use of time	4.33	<div></div>
It was easy to access the items on the Content Review Tool	4.13	<div></div>

Note: Strongly Disagree = 1 to Strongly Agree = 5

## Chapter 4: Summary and Recommendations

In this chapter, we provide an overview of the benchmark criteria, a summary of findings, and recommendations for NDE to consider based on these results. For ease of organization, the summary and recommendations are presented separately for each alignment criterion.

Table 45 outlines the evaluative guidelines for the overall benchmark criteria, which involves a two-step process. First, test events are evaluated *within* the three achievement levels (Developing, On Track, and Advanced). Meeting at least three out of four test event benchmarks results in a "Met" rating while meeting or partially meeting at least two benchmarks leads to a "Partially Met" rating. If fewer than two benchmarks are "Met" or "Partially Met," the criterion is considered "Not Met."

Next, we assess results *across* the three achievement levels. If all three achievement levels are met, the final criterion is "Met." Meeting or partially meeting two achievement levels leads to a "Partially Met" rating while meeting or partially meeting less than two achievement levels results in a "Not Met" rating. These guidelines offer a structured approach to evaluating and interpreting the overall performance of Criterion 1, 2, and 3 across test events and achievement levels by grade.

**Table 45. Overall Alignment Benchmark Criteria**

Criteria	Step 1 : Within Achievement Level	Step 2: Across Achievement Levels (Final Rating)
Criterion 1, 2, and 3	<p><b>Met:</b> At least three out of four test event benchmarks are met within each achievement level.</p> <p><b>Partially Met:</b> At least two of four test event benchmarks are met or partially met within each achievement level.</p> <p><b>Not Met:</b> Less than two of four test event benchmarks are met or partially met within each achievement level.</p>	<p><b>Met:</b> All three achievement levels are met.</p> <p><b>Partially Met:</b> Two achievement levels are met or partially met.</p> <p><b>Not Met:</b> Less than two achievement levels are met or partially met.</p>

Table 46 summarizes the alignment criteria results for the NSCAS ELA assessments for Grades 3-8.

**Table 46. Summary of Results by Criterion and Strand by Grade Level**

Grade	Criterion 1	Criterion 2	Criterion 3
Grade 3	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Met</b>	<b>Met</b>
Grade 4	RP: <b>Not Met</b> RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: <b>Met</b> RI: Partially Met V: <b>Met</b> W: <b>Met</b>	Partially Met
Grade 5	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: Partially Met RI: <b>Met</b> V: Partially Met W: <b>Met</b>	<b>Met</b>
Grade 6	RP: Partially Met RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: <b>Met</b> RI: Partially Met V: Partially Met W: <b>Met</b>	<b>Met</b>
Grade 7	RP: <b>Not Met</b> RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: Partially Met RI: <b>Met</b> V: <b>Met</b> W: Partially Met	<b>Met</b>
Grade 8	RP: <b>Not Met</b> RI: Partially Met V: Partially Met W: <b>Not Met</b>	RP: Partially Met RI: <b>Met</b> V: <b>Met</b> W: <b>Met</b>	<b>Met</b>

### **Criterion 1: Items Represent Intended Content**

This criterion examined the content alignment between 12 test events and the NE Standards. Specifically, we reviewed the majority agreement of the NE Standard identified for each item on the 12 test events.

The results show a diverse range of alignment between the test items and the standards outlined in the test blueprint. Examining the extent to which test events across grades met the Criterion 1 Benchmark, most test events by grade and strand “Partially Met” this criterion (Table 46). However, consistent across all grade levels is the recurring issue of test items not covering the breadth of writing standards outlined in the test blueprint. Many test items intended to evaluate writing proficiency consistently fell short of covering the number of writing standards.

A noteworthy finding is the dynamic nature of alignment with standards as students advance to higher grade levels. For example, Table 47 shows that Grade 3 did not meet benchmark criteria for one of 12 test events for Reading Prose and Poetry, one of 12 for Reading Informational Text, and four of 12 for Vocabulary. However, Grade 4 did not meet benchmark criteria for eight of 12 test events for Reading Prose and Poetry, three of 12 for Reading Informational Text, and five of 12 for Vocabulary—an increase in “Not Met” ratings across all three content strands. Additionally, Table 47 shows that Grade 6 did not meet benchmark criteria for two of 12 test

events for Reading Prose and Poetry, two of 12 for Reading Informational Text, and one of 12 for Vocabulary. However, Grades 7 and 8 did not meet benchmark criteria for eight of 12 and 11 of 12 test events for Reading Prose and Poetry, two of 12 and 7 of 12 for Reading Informational Text, and six of 12 and three of 12 for Vocabulary (respectively)—an increase in “Not Met” ratings across content strands.

**Table 47. Summary Across Achievement Levels**

Grade	RP	RI	V	W
Grade 3	<b>1 of 12</b> – Not Met	<b>1 of 12</b> – Not Met	<b>4 of 12</b> – Not Met	<b>12 of 12</b> – Not Met
Grade 4	<b>8 of 12</b> – Not Met	<b>3 of 12</b> – Not Met	<b>5 of 12</b> – Not Met	<b>12 of 12</b> – Not Met
Grade 5	<b>3 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>1 of 12</b> – Not Met	<b>12 of 12</b> – Not Met
Grade 6	<b>2 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>1 of 12</b> – Not Met	<b>12 of 12</b> – Not Met
Grade 7	<b>8 of 12</b> – Not Met	<b>2 of 12</b> – Not Met	<b>6 of 12</b> – Not Met	<b>12 of 12</b> – Not Met
Grade 8	<b>11 of 12</b> – Not Met	<b>7 of 12</b> – Not Met	<b>3 of 12</b> – Not Met	<b>12 of 12</b> – Not Met

Based on the results, there is partial support that items represent the intended content. Examination of the blueprint NE Standards to be assessed by items indicates that there are more standards than items allowed, especially with the Writing strand. Based on these findings, we present the following recommendation for NDE’s consideration:

- Revise the test specifications to align with the Standard level for the Vocabulary and Writing Strands rather than the sub-standard level. This is particularly relevant because the Writing strand included 20 or more sub-standards in numerous cases across various grade levels.

### **Criterion 2: Items Represent Intended Categories**

This criterion examined how items on each test event met the test blueprint targets for each content strand. Across grades and content strands, most benchmarks were either “Met” or “Partially Met” (Table 46). To strengthen the content strand blueprint target, we recommend the following for any strand that was “Partially Met.”

- Conduct a review of the NE Standards assigned to items in ELA to ensure Reading Prose and Poetry, Reading Informational Text, and Vocabulary are appropriately associated with the test items. This review can be completed by NDE or NWEA. Outcomes of this review may include but are not limited to re-assigning an NE Standard to an item.
- Review, across grade-level assessments, the ELA item banks for coverage of content strands. Where necessary, develop more items to ensure an adequate pool to draw from for CAT assessments.
- Examine the CAT algorithm to help ensure that the items represent the intended categories specified in the test blueprint.



### ***Criterion 3: Depth of Knowledge***

This criterion assessed the depth of knowledge of items. We examined the number of items at each DOK level across items on each test event using majority agreement DOK ratings.

Overall, the findings indicate that most items aligned with the DOK level 2. Across all grades, 70% or more of the items were aligned with a DOK level 2 or higher, except for three Developing test events in Grade 4. However, there were a handful of test events where no DOK 3 items were administered, specifically one test event Grade 4 On Track, two test events Grade 5 Developing, one test event Grade 8 Developing and one test event Grade 8 On Track. All other test events had at least one DOK 3 item. Based on these findings, we present the following recommendation for NDE's consideration:

- Evaluate the number of DOK 3 items available to determine whether a greater development effort should be made to increase the number of DOK 3 items.
- Continue to ensure balanced and effective item development by focusing on item writing efforts that maintain an appropriate distribution of DOK levels across grade levels.

### ***Criterion 4: Achievement Level Descriptors***

This criterion assessed the range of achievement level descriptors of items. We examined the number of items at each ALD level on each test event using majority agreement ALD ratings.

Overall, the findings indicate that most items aligned with ALD level 2. Across all grades, 70% or more of the items were aligned with an ALD level 2 or higher. However, there were several grade levels where no items were aligned with an ALD level 3. Based on these findings, we present the following recommendation for NDE's consideration:

- Evaluate the number of ALD level 3 items to determine whether a greater development effort should be made to increase the number of ALD level 3 items.
- Continue to ensure balanced and effective item development by focusing on item writing efforts that maintain an appropriate distribution of ALD levels across grade levels.

## References

- Achieve, Inc. (2018). *Criteria for procuring and evaluating high-quality and aligned summative science assessments*. Retrieved from <https://www.nextgenscience.org/sites/default/files/Criteria03202018.pdf>
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Content Area Standards* - Nebraska Department of Education. (n.d.). <https://www.education.ne.gov/assessment/nscas-system/>
- NSCAS Overview* - Nebraska Department of Education. (n.d.). <https://www.education.ne.gov/assessment/nscas-system/>
- U.S. Department of Education. (2018). A state's guide to the U.S. Department of Education's Assessment Peer Review Process.
- Webb, N. L. (1997). *Research Monograph No. 6: Criteria for alignment of expectations and assessments in Science and Science education*. Council of Chief State Schools Officers.
- Webb, N. L. (1999). *Research Monograph No. 18: Alignment of Science and Science standards and assessments in four states*. National Institute for Science Education and Council of Chief State School Officers. (ERIC Document Reproduction Service No. ED440852).
- Webb, N. L. (2002). *Alignment study in language arts, mathematics, science, and social studies of state standards and assessments for four states*. Council of Chief State School Officers.
- Webb, N. L. (2005). *Webb alignment tool: Training manual*. Wisconsin Center for Education Research. (n.d.). <http://www.wcer.wisc.edu/WAT/index.aspx>
- Wise, S. L., Kingsbury G. G., & Webb, N. L. (2015). Evaluating content alignment in computerized adaptive testing. *Educational Measurement: Issues and Practice*, 00, 1-8.

## Appendix A. Agenda

### Nebraska Student-Centered Assessment System (NSCAS) in ELA Virtual Alignment Workshop July 24 – July 28, 2023 Agenda

***Note: All times noted on the agenda are Central Time***

Date/Time	Description
<b>Day 1 – Monday, July 24, 2023</b>	
8:30 a.m. – 10:00 a.m.	Join Microsoft Teams meeting with all panelists and HumRRO Facilitators. Welcome, logistics, overview of NSCAS in ELA, general alignment training
10:00 a.m. – 10:15 a.m.	Break
10:15 a.m. – 11:45 a.m.	Join Teams meeting for assigned grade level panel, panelist introductions, confirm access to online documents and Content Review Tool for NSCAS in ELA, review panelist instructions for rating items and calibrate item ratings, and begin iterative alignment rating process
11:45 a.m. – 12:45 p.m.	Lunch Break
12:45 p.m. – 2:15 p.m.	Continue iterative alignment rating process
2:15 p.m. – 2:30 p.m.	Break
2:30 p.m. – 4:00 p.m.	Continue iterative alignment rating process
4:00 p.m.	<i>Adjourn for the day</i>
<b>Day 2 – Tuesday, July 25, 2023</b>	
8:30 a.m. – 10:00 a.m.	If needed: Review and rerate items from Day 1. Continue iterative alignment rating process
10:00 a.m. – 10:15 a.m.	Break
10:15 a.m. – 11:45 a.m.	Continue iterative alignment rating process
11:45 a.m. – 12:45 p.m.	Lunch Break
12:45 p.m. – 2:15 p.m.	Continue iterative alignment rating process
2:15 p.m. – 2:30 p.m.	Break
2:30 p.m. – 4:00 p.m.	Continue iterative alignment rating process
4:00 p.m.	<i>Adjourn</i>

Date/Time	Description
<b>Day 3 – Wednesday, July 26, 2023</b>	
8:30 a.m. – 10:00 a.m.	If needed: Review and rerate items from Day 2. Continue iterative alignment rating process
10:00 a.m. – 10:15 a.m.	Break
10:15 a.m. – 11:45 a.m.	Continue iterative alignment rating process
11:45 a.m. – 12:45 p.m.	Lunch Break
12:45 p.m. – 2:15 p.m.	Continue iterative alignment rating process
2:15 p.m. – 2:30 p.m.	Break
2:30 p.m. – 4:00 p.m.	Continue iterative alignment rating process
4:00 p.m.	<i>Adjourn</i>
<b>Day 4 – Thursday, July 27, 2023</b>	
8:30 a.m. – 10:00 a.m.	If needed: Review and rerate items from Day 3. Continue iterative alignment rating process
10:00 a.m. – 10:15 a.m.	Break
10:15 a.m. – 11:45 a.m.	Continue iterative alignment rating process
11:45 a.m. – 12:45 p.m.	Lunch Break
12:45 p.m. – 2:15 p.m.	Continue iterative alignment rating process
2:15 p.m. – 2:30 p.m.	Break
2:30 p.m. – 4:00 p.m.	Continue iterative alignment rating process
4:00 p.m.	<i>Adjourn</i>
<b>Day 5 – Friday, July 28, 2023</b>	
8:30 a.m. – 10:00 a.m.	If needed: Review and rerate items from Day 3. Continue iterative alignment rating process
10:00 a.m. – 10:15 a.m.	Break
10:15 a.m. – 11:45 a.m.	Continue iterative alignment rating process
11:45 a.m. – 12:45 p.m.	Lunch Break
12:45 p.m. – 2:15 p.m.	Continue iterative alignment rating process
2:15 p.m. – 2:30 p.m.	Break
2:30 p.m. – 3:45 p.m.	Continue iterative alignment rating process
3:45 p.m. – 4:00 p.m.	Complete two short online surveys: <ul style="list-style-type: none"> <li>• Demographic information</li> <li>• Debrief/ Workshop evaluation</li> </ul>
4:00 p.m.	<i>Adjourn</i>

## Appendix B. Panelist Requirements

Nebraska teachers will serve as panelists for the alignment workshop. All reviewers will be confirmed by NDE.

Educators will have the following minimum qualifications for serving as a reviewer:

- Educators are **seasoned, certified, professionals**, who have **strong familiarity** with the Nebraska Standards in ELA. Educators
- Educators are current teachers with at least three years of teaching experience at their respective grade level or teachers who retired after 2021 when the standards were adopted
- Educators have at least read the Nebraska ELA Standards for their grade and related grade span
- Educators have participated in professional development activities related to the Nebraska Standards in ELA, including prior participation with reviewing test items (e.g., prior alignment study experience, prior standard setting study experience)
- Educators have developed a curriculum that incorporates the Nebraska Standards in ELA
- Educators have experience with the NE standards and Range ALDs

## Appendix C. Panelist Instructions

### Nebraska ELA Alignment Workshop Panelist Instructions

#	Title of Material
1	Panelist instructions
2	Panelist rating sheets
3	Panelist training slides
4	Nebraska ELA items - Accessed via the Content Review Tool
5	Nebraska's College and Career Ready Standards for English Language Arts (NE Standards)
6	Depth of Knowledge (DOK) Levels (Cognitive Complexity)
7	Achievement Level Descriptors (ALDs)
8	Demographics form (via MS Forms) – administered at the end of the workshop
9	Process evaluation survey (via MS Forms) – administered at the end of the workshop

#### Terminology:

- **NE Standards:** Nebraska's College and Career Ready Standards for English Language Arts
- **NSCAS:** Nebraska's Student-Centered Assessment System (NSCAS)

#### **Test Security Notice**

Please do **not** use your personal electronic devices while engaged in alignment workshop tasks. If you need to use your phone or other devices for any reason, please step away from the computer or wait to use your devices during a break.

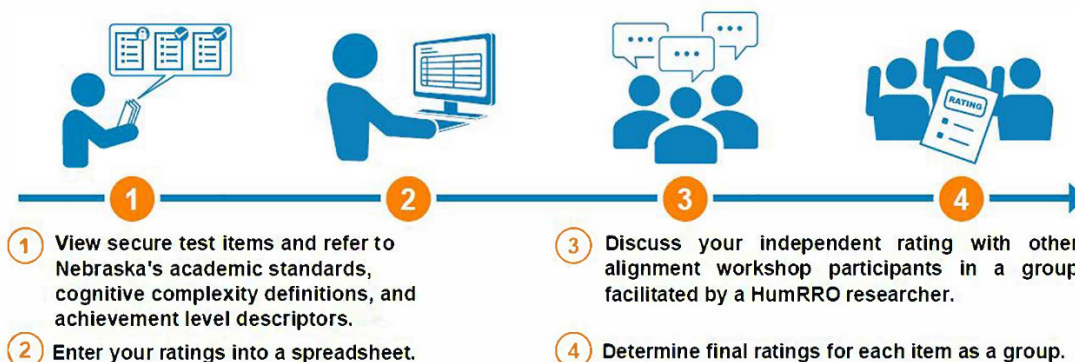
**This rule will be strictly enforced during the workshop.**

### **Task 1: Introductions and Materials Overview**

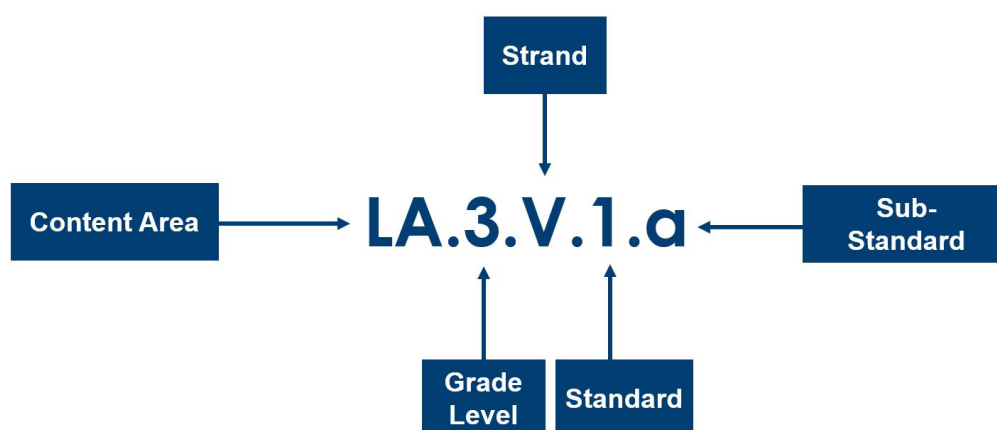
- **Panelist and facilitator introductions**
- **Review the materials in the table above**
  - Google Drive folder with digital materials
  - Facilitator demonstration of how to access Google Sheets. (Follow along on your computer.)
  - Facilitator demonstration of how to log on to the Content Review Tool. (Follow along on your computer.)
  - Materials #8 and #9 (Demographics form and Process Evaluation survey) will be administered at the end of the workshop.

### **Task 2: Training on the NSCAS ELA item alignment**

- **Brief explanation of the process for this task.**
  - You will review NSCAS ELA items administered to students in Nebraska.
  - You will first calibrate your ratings by reviewing a small set of items (typically, the first three to five items). This will be an opportunity for the group to talk through the process and everyone discuss their approach to reviewing each item. This will ensure everyone is thinking about the ratings in the same way. You will then enter your ratings into your rating sheet using the drop-down menus. You will assign a NE Standard(s) that **best** match what the item measures. Your ratings will focus on the alignment of each item to content within the NE Standards, cognitive complexity (DOKs), and achievement level descriptors (ALDs). We will discuss each of these ratings and settle on a final majority rating.
- After calibration, you will independently review a small set of items and enter your ratings into your rating sheet per the instructions above. You will review items in logical sets (e.g., all items in a cluster). Once all items in a set have been reviewed, we will discuss the items as a group and settle on a final majority rating.
- Below is a graphic from the general panelist training that provides a high-level overview of the process.



- Below is a graphic that lays out the structure of the NE Standard codes:



**The NE Standards contain the following components:**

- Content area ("LA" refers to Language Arts)
- Grade Level
- Strand:
  - RP = Reading Prose and Poetry (standard level)
  - RI = Reading for Informational Text (standard level)
  - V = Vocabulary (sub-standard level)
  - W = Writing (sub-standard level)
- Standard (always a number)
- Sub-standard (always a lowercase letter, V & W only)

- Open individual rating sheets (Google Sheets).** Open Google Chrome on your computer and navigate to your individual Google rating sheet. Each sheet has a unique panelist name.
- Review making ratings in the Google Sheet.**



**Facilitator demonstration on how to enter data in the sheet (i.e., using drop-down menus, entering comments).**

- You will need to review only the first sheet. Other sheets are hidden and should not be accessed or modified. If any issue occurs with the drop-down menu options or conditional formatting, notify your facilitator.

Discuss **Columns A and B (Item Sequence and UIN)**

- Columns A and B contain information about each NSCAS ELA item.
- **Column A** indicates the order the item appears in the Content Review Tool. This number will be what you and the panelists use to make sure everyone is talking about the same item.
- **Column B** provides the unique item number (UIN). You will not use this for your ratings, but it is provided in case any items seem to be out of sequence in the Content Review Tool.
- Please ensure that you are viewing the same item in the Content Review Tool that you are rating in your sheet.

Discuss **Column C (Identify the Standard)**

- Column C asks you to identify the Standard code using a drop-down menu.
- You should be very familiar with the 2021 NE Standards document from which the Standard codes are derived. You are permitted to use your own marked-up copies of the 2021 NE Standards if you have their own.
- **If you believe the item does not align with a NE Standard, you should select “None.”** You must enter a comment explaining the reason you entered “None” for this rating in **Column K**.
- If you have a difficult time choosing between two or more Standards, you should select the Standard that **best aligns** with the item. Then, you should enter a comment in **Column K** that includes the other Standard(s) you considered.

Discuss **Column D (Item Writer’s Standard)**

- Once you select a Standard code from **Column C**, the Standard code associated with the item per the item writer will appear in **Column D**. The purpose of this is for you to see what Standard you selected and compare it with the assigned Standard from the item writer.

Discuss **Columns E and F (Standard Text)**

- **Column E** will display the Standard text associated with the selected Standard Code in **Column C**.
- **Column F** will display the Standard text associated with the metadata Standard in **Column D**.

Discuss **Column G (Final Rating)**

- **Column G** will ask you to choose between your selected Standard in **Column C** and the item writer’s Standard in **Column D**.

Discuss **Column H (Final Standard Rating Description)**

- **Column H** will ask you to briefly describe if your final Standard in **Column G** differs from the item writer's Standard in **Column D**, and to please indicate why.

Discuss **Column I (Identify the Depth of Knowledge/ Cognitive Complexity Level)**

- **Column I** is for you to provide the overall cognitive complexity level (1, 2, 3, or 4) that best represents the cognitive demand of the item. Remember you'll need to evaluate the cognitive complexity and **not** the item difficulty (although highly correlated, they are not always the same). Keep in mind that cognitive complexity refers to what the item is asking the student to do and how that task fits into the cognitive complexity framework (reference the DOK document provided).

Discuss **Column J (Identify the Range Achievement Level Descriptor)**

- **Column J** is for you to provide the range achievement level descriptor (1 - Developing, 2 – On Track, or 3 - Advanced). You'll want to reference the Range ALD document provided.

Discuss **Column K (Comments)**

- **Column K** is for you to enter any comments.
- A few simple rules for the comments field:
  - a. If the comments cell is highlighted yellow, it means one or more of the following ratings were selected: "None" for the final Standard **OR** you selected a Standard that differed from the item writer's Standard.
  - b. You may also provide comments or notes regarding the quality of the item or the phenomenon the item references. Panelists should take notes on their own, discuss them, and the facilitator should capture the agreed-upon points in the facilitator spreadsheet.
  - c. The primary purpose of this column is to provide comments related to the alignment of the item to the rating categories. All comments will be **anonymously** provided to the Nebraska Department of Education for review.

### **Task 3: Rating Calibration Task**

- You will rate all indicated fields for the **first item**. Since this is a calibration activity, you should read the item, review the reference materials, then work together to come up with a rating for each rating category. The calibration is a collaborative activity, though you should be reminded that, after calibration, you will rate items independently, then discuss their ratings with the rest of the panel once the rest of the panel has finished rating a set of items.
- During calibration, you should focus on why you agree or disagree and what the most appropriate selections should be. Be sure you spend a little time with cognitive complexity and achievement level descriptors.
- You will repeat calibration for up to two to four additional items.


#### **Task 4: Conduct Independent Item Ratings**

- You should rate all remaining NSCAS ELA items independently in sets before discussing and settling on a majority (items are typically rated in sets based on the corresponding passages). Repeat the process above for each set of items. You will review items in clusters so that any linked items are not broken into separate review and rating sessions.
- You will work independently; however, occasional discussion about any item(s) that is causing someone difficulty is allowed.
- **After discussing an item, you should not change your rating unless you made a coding error.** The facilitator will capture majority ratings among the panelists, but HumRRO wants to be able to gauge the differences between independent panelist ratings and the final majority ratings.

#### **Task 5: Workshop Debrief**

- Once all final majority ratings have been collected, please close all materials (e.g., rating sheet, Content Review Tool, any electronic versions of references) and open the MS Word document with the link to the debriefing surveys.
- You will first take the “Demographic Survey” followed by the “Process Evaluation Survey.”
- Please note that your responses will be anonymous and will only be shared in an aggregate format.

## Appendix D. Panelist Training Slides



**Independent English Language Arts (ELA) Alignment Study for the Nebraska Department of Education**

**Facilitator Training**

July 18, 2023

HumRRO Headquarters: 88 Canal Center Plaza, Suite 700, Alexandria, VA 22314-1570 | Phone: 703.648.3911 | www.humrrro.org

1



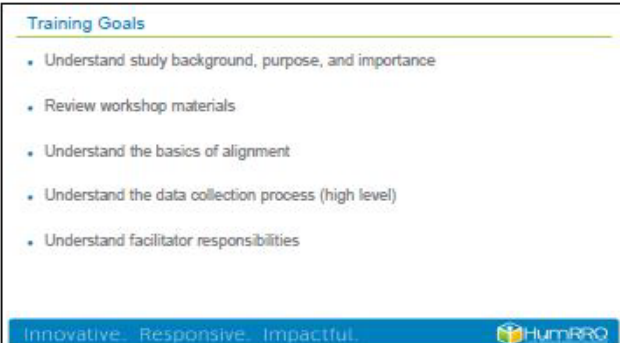
**Introductions**

- Introductions
  - HumRRO staff
  - Nebraska Department of Education staff
  - NWEA staff
- Panel assignments

Panel	HumRRO Facilitator
Grade 3	Soner Dumani
Grade 4	Reese Butterfuss
Grade 5	Lori Schantz
Grade 6	Charlie Wayne
Grade 7	Monica Groben
Grade 8	Emily Dickinson

Innovative. Responsive. Impactful.

2



**Training Goals**

- Understand study background, purpose, and importance
- Review workshop materials
- Understand the basics of alignment
- Understand the data collection process (high level)
- Understand facilitator responsibilities

Innovative. Responsive. Impactful.

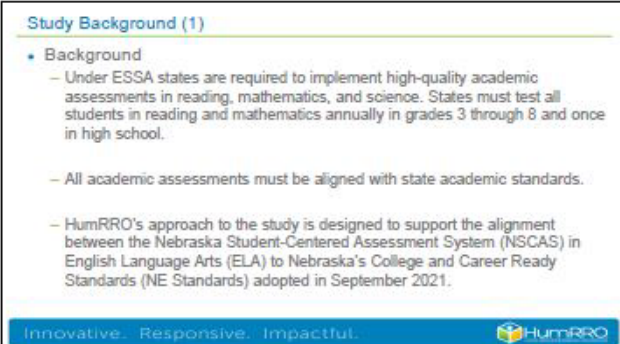
3



**Study Background, Purpose, and the Importance of Alignment**

Innovative. Responsive. Impactful.

4

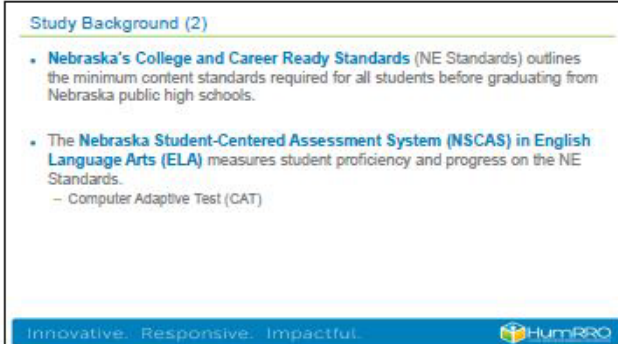


**Study Background (1)**

- Background
  - Under ESSA states are required to implement high-quality academic assessments in reading, mathematics, and science. States must test all students in reading and mathematics annually in grades 3 through 8 and once in high school.
  - All academic assessments must be aligned with state academic standards.
  - HumRRO's approach to the study is designed to support the alignment between the Nebraska Student-Centered Assessment System (NSCAS) in English Language Arts (ELA) to Nebraska's College and Career Ready Standards (NE Standards) adopted in September 2021.

Innovative. Responsive. Impactful.

5



**Study Background (2)**

- Nebraska's College and Career Ready Standards (NE Standards) outlines the minimum content standards required for all students before graduating from Nebraska public high schools.
- The Nebraska Student-Centered Assessment System (NSCAS) in English Language Arts (ELA) measures student proficiency and progress on the NE Standards.
  - Computer Adaptive Test (CAT)

Innovative. Responsive. Impactful.

6

#### Purpose

- Purpose
  - To collect evidence of the alignment of the NSCAS in ELA to the NE Standards (validity evidence)
  - To evaluate the content representation of the test events
  - To evaluate the range and distribution of cognitive complexity or depth of knowledge (DOK)
  - To evaluate how well the achievement level descriptors (rALDs) capture the knowledge and skills expressed in the item

Innovative. Responsive. Impactful.



7

#### Importance of Alignment

- Fairness for all students
  - Consistency in general curriculum
  - Accurate assessment of what students can do and are expected to know from State content standards and the curriculum
  - Improves teacher instruction and student learning
- Federal Peer Review Requirement

Innovative. Responsive. Impactful.



8

## Workshop Materials

Innovative. Responsive. Impactful.



9

#### Workshop Materials – Google Drive (1)

#	Title of Material
1	Facilitator instructions
2	Facilitator rating sheet
3	Facilitator training slides
4	Panelist instructions
5	Panelist rating sheets
6	Panelist training slides

Innovative. Responsive. Impactful.



10

#### Workshop Materials – Google Drive (2)

#	Title of Material
7	NE ELA Items – Accessed via the Content Review Tool
8	NE Standards
9	Cognitive complexity (DOK levels)
10	Achievement level descriptors (ALDs)
11	Demographics form (MS Form)
12	Process evaluation survey (MS Form)

Innovative. Responsive. Impactful.



11

## Standards, Cognitive Complexity (DOK), and Achievement Level Descriptors (ALDs)

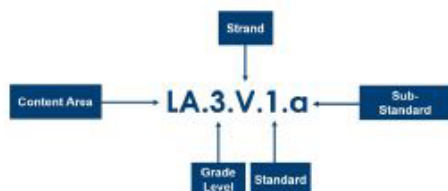
Innovative. Responsive. Impactful.



12

### Standards

- Nebraska's College and Career Ready Standards for English Language Arts (NE Standards)



Innovative. Responsive. Impactful.



13

### Cognitive Complexity

- Refers to the type of **cognitive processing** required to access and respond to the test item
- Frequently measured using Webb's Depth of Knowledge (DOK) definitions
- Cognitive complexity is related to but distinct from difficulty

Innovative. Responsive. Impactful.



14

### Cognitive Complexity: Webb's Depth of Knowledge Levels

Webb's DOK Level	Definition
Level 1: Recall and Reproduction	Requires recall of information, such as a fact, definition, term, simple procedure, or property. Typically involves only one step.
Level 2: Skill/Concept	Requires some mental processing beyond recalling or reproducing a response. Typically involves more than one step.
Level 3: Strategic Thinking	Requires deep knowledge using reasoning, planning, or using evidence. Typically, has more than one possible answer and requires students to justify their response.
Level 4: Extended Thinking (No items in this alignment workshop are DOK 4)	Requires high cognitive demand and is very complex. Typically includes complex reasoning, experimental design and planning, and probably will require an extended period.

Innovative. Responsive. Impactful.



15

### Achievement Level Descriptors

- Achievement level descriptors (ALDs) describe the knowledge, skills, and processes that students demonstrate on state tests at pre-determined levels of achievement for each tested grade level.
- The Nebraska State Board of Education defined three achievement levels for each content area:
  - 1. **Developing**
  - 2. **On Track**
  - 3. **Advanced**

Innovative. Responsive. Impactful.



16

### Achievement Level Descriptors

- 1. Developing:** Developing learners do not yet demonstrate proficiency in the knowledge and skills necessary at this grade level, as specified in the Nebraska College and Career Ready Standards.
- 2. On Track:** On Track learners demonstrate proficiency in the knowledge and skills necessary at this grade level, as specified in the assessed Nebraska College and Career Ready Standards.
- 3. Advanced:** Advanced learners demonstrate proficiency in the knowledge and skills necessary at this grade level, as specified in the assessed Nebraska College and Career Ready Standards.

Innovative. Responsive. Impactful.



17

## Virtual Alignment Workshop

Innovative. Responsive. Impactful.



18



### Virtual Alignment Workshop

- We will collect judgments from Nebraska educators on the following:
  - Items to standards alignment
  - Items to cognitive complexity alignment (DOK)
  - Items to achievement level descriptors (ALDs)

Innovative. Responsive. Impactful.



19

### The Alignment Rule

- For items, "aligned" means the item's content is included in the standard
- **Most items cannot capture the full breadth of a standard**
- Think of the Standards as buckets - If the item fits in any of these buckets, it is aligned. It does not need to fill the bucket.

Innovative. Responsive. Impactful.



20

### Data Collection Process (1)

- Virtual Meeting using multiple platforms/tools
  - Microsoft Teams
  - Google Drive
  - Item Viewing Platform (Content Review Tool)
- Six (6) grade-level panels (Grades 3 through 8)
- Up to 7.5 hours per day, over five (5) days, including scheduled breaks
  - **July 24 – July 28, 8:30 am - 4:00 pm Central Time**
- Start with whole-group training and then break into panels
  - EB will send out the whole group training invite (8:30am – 10:00am CT)
  - Facilitators will send separate meeting links sent (10:15am- 4:00pm CT; 8:30am – 4:00pm CT)

Innovative. Responsive. Impactful.



21

### Data Collection Process (2)

- Prior to the workshop
  - EB to send facilitators an email with a link to the Google Drive folder
- Whole-group training
  - EB will send out the whole-group training invite from HumRRO
  - Study background and purpose
  - Overview of the alignment method
- Small group sessions
  - EB will send facilitators an email with instructions on sending your panelists two MS Teams meeting invites for your individual panels

Innovative. Responsive. Impactful.



22

### Data Collection Process (3)

- What is to be viewed/rated
  - Test items
    - Accessed via the Content Review Tool (CRT)
- Materials to support panelist item ratings
  - Panelist instructions
  - Panelist rating sheet
  - Panelist training slides
  - NE Standards
  - Cognitive complexity levels (DOK levels)
  - Achievement level descriptors (ALDs)

Innovative. Responsive. Impactful.



23

### Data Collection Process (4)



Innovative. Responsive. Impactful.



24

### Facilitation Scenario (2)

- Scenario: The majority of panelists independently select the same rating.
- Is this during calibration (first three items)?
  - **Yes:** Have panelists discuss ratings.
  - **No:** Does this involve a standard/cognitive complexity rating/achievement level descriptor that the group has not discussed yet?
    - **Yes:** Have panelists discuss ratings.
    - **No:** Record majority rating.

Innovative. Responsive. Impactful.



31

### Facilitation Scenario (3)

- Scenario: Panelists are evenly split on a standard, cognitive complexity, or achievement level descriptor.
- Is this a standard rating?
  - **Yes:** if the tie cannot be broken, record "Split" as the final standard rating and note in the facilitator comments the standards the group was split on.
- Is this a cognitive complexity or achievement level descriptor rating?
  - **Yes:** if the ratings are non-adjacent, have group discuss further to ensure common understanding of the levels and possibly get to a majority rating. If still evenly split, record the higher DOK or ALD level and comment in the facilitator rating form that this was an evenly split rating and reference the lower rating.

Innovative. Responsive. Impactful.



32

### HumRRO Support Staff

- Emily Borawski:
  - Whole-Group Training Facilitator
  - Floater
- Yvette Nemeth:
  - Technical Advisor
  - Floater
- Carrie Wiley:
  - Floater

Innovative. Responsive. Impactful.



33

## Thank You!

Charge Code:  
 \$22-28 NE ELA&MathAlign&it

For any questions,  
 please contact:

Emily Borawski  
 ebora@humro.org



Innovative. Responsive. Impactful.

humro.org

34



## Appendix E. Standards (Grade 3 Example)

# NEBRASKA'S COLLEGE AND CAREER READY STANDARDS FOR ENGLISH LANGUAGE ARTS



Approved by the Nebraska State Board of Education on September 2, 2021

## K-12 Comprehensive English Language Arts Standards

Strand	Comprehensive Standard
<b>Foundations of Reading (F)</b>	Students will develop and apply decoding and language comprehension skills and strategies to comprehend and learn from increasingly complex texts.
<b>Reading Prose and Poetry (RP)</b>	Students will learn and apply reading skills and strategies to comprehend grade-level literary texts.
<b>Reading Informational Text (RI)</b>	Students will learn and apply reading skills and strategies to comprehend grade-level informational texts.
<b>Vocabulary (V)</b>	Student will build and use conversational, academic, and discipline-specific, grade-level vocabulary.
<b>Writing (W) and Foundations of Writing (FW)</b>	Students will learn and apply writing skills and strategies to communicate effectively for a variety of purposes.
<b>Speaking and Listening (SL)</b>	Students will learn and apply speaking and listening skills and strategies to communicate effectively for a variety of audiences and purposes.

**Spiraled, Vertical Progressions.** The revised 2021 Nebraska English Language Arts Standards are formatted to support educators in both grade-level and vertical instructional planning. In addition to organization by grade level, the standards and indicators are formatted into spiraled, vertical articulations. This design demonstrates the interrelated nature of skills in the English Language Arts and their progression through the grade levels. The purpose of presenting the standards into vertical charts is to provide educators with a practical tool for the development of a locally-determined, standards-aligned curriculum.

For each standard in the areas of Foundations of Reading, Reading Prose and Poetry, Reading Informational Text, Writing\*, Vocabulary, and Speaking and Listening, the standards and indicators are listed in a table format from the 11-12 grade band and ending at Kindergarten.

# Grade 3 Standards

## ■ READING PROSE AND POETRY

**Central Ideas and Details** | Citing relevant and thorough textual evidence to support ideas, evaluate the development of themes or central ideas in grade-level literary texts.

**LA.3.RP.1** Identify the central message or lesson in a literary text and explain how key details support that idea.

**LA.3.RP.2** Explain how characters respond to major events and challenges in a literary text.

**Author's Craft** | Citing relevant and thorough evidence to support ideas, evaluate the development and interaction of individuals, ideas, and events in grade-level literary and informational text.

**LA.3.RP.3** Determine and explain the point of view in a literary text.

**LA.3.RP.4** Explain how sections of a literary text (e.g., chapters, scenes, stanzas) build on one another and contribute to meaning.

**Knowledge and Ideas** | Citing relevant and thorough textual evidence to support ideas, evaluate how an author's perspective or use of point of view shapes the style and meaning of grade-level literary text.

**LA.3.RP.5** Compare and contrast the themes, settings, and plots of literary texts written by the same author about the same or similar characters (e.g., books from a series).

**LA.3.RP.6** Explain what the text says explicitly and draw inferences when asking and answering questions.

**LA.3.RP.7** Compare and contrast themes, topics, and/or patterns of events in a range of literary texts.

**Do not select the following standard - locally assessed only:**

**Range of Reading and Level of Text Complexity** | Read and comprehend complex, grade-level literary text independently and proficiently.

**LA.3.RP.8** Read and comprehend a wide range of literary texts of appropriate complexity for Grade 3 independently and proficiently.

### ■ Instructional Considerations

- In describing settings or characters, students should explain what in the text the descriptions are based upon.
- Students should be made aware that not all narratives contain a central message or lesson.
- At all grade levels, students should read paired, conceptually-related (by topic, theme, and/or genre) literary and informational texts.
- *Point of view* refers to the vantage point from which a narrative is told.

## READING INFORMATIONAL TEXT

**Central Ideas and Details** | Citing relevant and thorough textual evidence to support ideas, evaluate the development of themes or central ideas in grade-level informational text.

**LA.3.RI.1** Identify the central idea and explain how key details support that idea.

**LA.3.RI.2** Explain the relationship between individuals, historical events, scientific ideas or concepts, or steps in a process.

**Author's Craft** | Citing relevant and thorough evidence to support ideas, evaluate the development and interaction of individuals, ideas, and events in grade-level informational text.

**LA.3.RI.3** Determine and explain the author's purpose in an informational text.

**LA.3.RI.4** Explain how text features (titles, headings, table of contents, glossaries, captions, graphs, maps, and/or other visuals) contribute to meaning.

**Knowledge and Ideas** | Citing relevant and thorough textual evidence to support ideas, evaluate how an author's perspective or use of point of view shapes the style and meaning of grade-level informational text.

**LA.3.RI.5** Compare and contrast the two most important ideas and key details presented by multiple informational texts on the same topic.

**LA.3.RI.6** Identify an author's claim(s) and explain how the author supports the claim(s) in the text.

**LA.3.RI.7** Compare and contrast topics and/or patterns of events in a range of informational texts.

**Do not select the following standard - locally assessed only:**

**Range of Reading and Level of Text Complexity** | Read and comprehend complex, grade-level informational text independently and proficiently.

**LA.3.RI.8** Read and comprehend a wide range of informational texts of appropriate complexity for Grade 3 independently and proficiently.

### Instructional Considerations

- A *claim* refers to an author's primary argument and is supported by textual evidence.
- *Author's craft* refers to the techniques an author uses to develop and support a claim.
- *Point of view* refers to the vantage point from which a story is told, while *perspective* is an author's attitude or belief that is based on personal knowledge and/or experience.

## VOCABULARY

**Acquisition and Use** | Build and use a range of conversational, academic, and discipline-specific grade-level vocabulary and apply to reading, writing, speaking, and listening.

**LA.3.V.1** Acquire and use grade-level academic vocabulary appropriately.

- a. Use sentence-level context clues to determine the meaning of a word or phrase.
- b. Use affixes to determine the meaning of unknown words (e.g., comfortable, uncomfortable).
- c. Use known root words to determine the meaning of unknown words (e.g., company, companion).

**Do not select the following sub-standard - locally assessed only:**

- d. Determine the meanings of key words and phrases using reference materials and classroom resources.

**Context and Connotation** | Determine or clarify the meaning of unknown and multiple-meaning words and phrases, choosing flexibly from a range of strategies.

**LA.3.V.2** Interpret an author's use of figurative, connotative, and technical language in grade-level literary and informational text.

- a. Distinguish between literal and nonliteral meanings of words and phrases in context (e.g., take steps).
- b. Identify real-life connections between words and their use (e.g., describe people who are friendly or helpful).
- c. Distinguish nuances of meaning between related words that describe states of mind or degrees of certainty (e.g., believed, suspected).

### Instructional Considerations

- Academic vocabulary refers to words likely to appear in a variety of content area texts, at or above grade-level, and typically requires explicit instruction. Students should be encouraged to use newly acquired terms frequently in speaking and writing.
- The vast majority of academic vocabulary taught should derive from complex texts—a careful review of texts for challenging words that are central to understanding the meaning of the text, including figurative language, should determine which vocabulary is taught explicitly (sometimes in advance of reading).
- Include a word study component that includes prefixes, root words, and suffixes to accompany text-based methods of vocabulary development.
- Reading aloud to students using texts that are two grade levels higher than their reading level is an evidence-based practice for activating prior knowledge and building vocabulary.

Page 49

Approved by the Nebraska State Board of Education on September 2, 2021



## ■ WRITING

**Production of Writing** | Use a recursive writing process to produce clear and coherent writing appropriate to the discipline, audience, and/or context.

**LA.3.W.1** Write paragraphs using a variety of sentence types.

- a. Capitalize proper nouns (e.g., historic periods, nationalities, languages), proper adjectives (e.g., South American), and appropriate words in titles.
- b. Use commas in addresses and commas and quotation marks in dialogue; use an apostrophe to form and use possessives.
- c. Use frequently occurring nouns (e.g., concrete and abstract), verbs (regular and irregular), and simple verb tenses.
- d. Distinguish between and use coordinating and subordinating conjunctions and independent and dependent clauses.
- e. Explain the function of adjectives and adverbs in simple, compound, and complex sentences.
- f. Use correct subject-verb and pronoun-antecedent agreement in speaking and writing.
- g. Use frequently occurring prepositions and prepositional phrases.

**Do not select the following standard/sub-standards - locally assessed only:**

**LA.3.W.2** Use a recursive writing process to develop, strengthen, and produce writing appropriate to the audience, purpose, and discipline.

- a. Use prewriting activities and resources to plan, organize, and draft writing.
- b. Adapt writing processes to sustain engagement in short and long-term writing tasks of increasing length and complexity.
- c. Improve and clarify the content, structure, and organization of writing by revising, considering feedback from adults and peers.
- d. Improve and clarify writing by editing and proofreading, considering feedback from adults and peers.
- e. Use or decipher multiple formats of print and digital text (e.g., manuscript, cursive, font, graphics, symbols).
- f. Use appropriate print and digital/multimedia tools to produce, enhance, and/or publish writing individually or in collaboration with peers.

**Modes of Writing** | Write in a variety of modes for a variety of purposes and audiences across disciplines.

**LA.3.W.3** Write creative and/or expressive pieces that describe a well-developed event or experience.

- a. Engage and orient the reader by establishing a situation and introducing a narrator and/or character(s).
- b. Include descriptive details about characters, events, or settings.
- c. Use words and phrases to signal a sequence of events.
- d. Provide a closure related to the creative or expressive event or experience.

**LA.3.W.4** Write opinion pieces with supporting reasons and/or evidence.

- a. Introduce a topic or text, state an opinion, and develop a structure that includes reasons and/or evidence.
- b. Use linking words and phrases to connect opinions and reasons.
- c. Provide a concluding statement or section related to the opinion.

**LA.3.W.5** Write informative/explanatory pieces to examine a topic or text and convey ideas and information.

- a. Introduce a topic and group related information together, including illustrations when useful to provide clarity.
- b. Develop the topic with information (e.g., facts, definitions, details) clearly related to the topic.
- c. Use linking words and phrases and key vocabulary to connect ideas and categories of information.
- d. Provide a concluding statement or section related to the topic.

**LA.3.W.6** Locate evidence from literary and/or informational text sources to answer questions about a topic.

- a. Paraphrase information from sources to support ideas while avoiding plagiarism.  
**Do not select the following sub-standard:**
- b. Identify print and digital tools to gather information and ideas to answer questions.
- c. Sort evidence into categories using an appropriate note-taking format to collect and organize information.  
**Do not select the following sub-standards:**
- d. Demonstrate academic integrity by avoiding overreliance on any one source and referencing sources in writing and speaking; provide a list of sources.
- e. Practice safe and ethical behaviors when communicating and interacting with others digitally (e.g., safe information to share, utilize appropriate sites and materials, appropriate language use, respect diverse perspectives).



**Describe Explain Interpret**

**Level One (Recall)**

Draw Identify List Label  
Define Memorize Illustrate  
Calculate Who, What, When, Where, Why Measure  
Arrange State Tabulate Name Report  
Repeat Tell Recognize Use Quote  
Recall Recite Match Infer  
Design Categorize

**Level Two (Skill/Concept)**

Connect Collect and Display  
Synthesize Identify Patterns  
Apply Concepts Organize  
Critique Construct  
Analyze Separate  
Create Cause/Effect Modify  
Prove Estimate Predict  
Interpret

**Level Three (Strategic Thinking)**

Revise Distinguish  
Appraise Use Context Cues  
Develop a Logical Argument Make Observations  
Use Concepts to Solve Non-Routine Problems Summarize  
Explain Phenomena in Terms of Concepts Show  
Formulate Compare  
Draw Conclusions Investigate  
Hypothesize Differentiate  
Cite Evidence

**Level Four (Extended Thinking)**

Design  
Synthesize  
Apply Concepts  
Critique  
Analyze  
Create  
Prove

Webb, Norman L. and others. "Web Alignment Tool" 24 July 2005. Wisconsin Center of Educational Research, University of Wisconsin-Madison. 2 Feb. 2006. <<http://www.wcer.wisc.edu/WAT/index.aspx>>

## Appendix G. Achievement Level Descriptors (Grade 3 Example)

Indicator No.	Indicator Text	Developing	On Track	Advanced
		With a range of texts with text complexity commonly found in Grade 3, a student performing in Developing can likely	With a range of texts with text complexity commonly found in Grade 3, a student performing in On Track can likely	With a range of texts with text complexity commonly found at the intersection of Grade 3 and Grade 4, a student performing in Advanced can likely
<b>Reading Prose and Poetry</b>				
<b>LA.3.RP.1</b>	Identify the central message or lesson in a literary text and explain how key details support that idea.	Identify the central message or lesson in a literary text.	Identify the central message or lesson in a literary text and explain how key details support that idea.	Analyze the central message or lesson in a literary text and explain how key details support that idea.
<b>LA.3.RP.2</b>	Explain how characters respond to major events and challenges in a literary text.	Identify the major events and/or challenges that characters face in a literary text.	Explain how characters respond to major events and challenges in a literary text.	Analyze how characters respond to major events and challenges in a literary text, drawing on specific details such as a character's thoughts, words, or actions.

## Appendix H. Correlation Analysis

The correlation between Depth of Knowledge and Achievement Level Descriptors was examined across grade levels. The results revealed a moderate correlation in Grades 4-8, with values ranging from  $r = .31$  to  $r = .45$ . In Grade 3, a stronger correlation of  $r = .57$  was observed (Table H1).

**Table H1. Correlation between DOK and ALD by Grade**

Grade	Correlation
ELA 3	.57
ELA 4	.45
ELA 5	.39
ELA 6	.37
ELA 7	.31
ELA 8	.41

## **Appendix I. Number of Unique and Shared Items by Grade, Test Event, and Strand**

The number of unique and shared items by grade, test event, and strand were examined. Results showed that many writing items were shared across test events and achievement levels, except for Grade 4, which showed a greater number of unique writing items. These results may indicate that the pool of writing items is shallow and/or the CAT test algorithm is not selecting unique writing items across test events and achievement levels.

**Table 11. Number of Unique and Shared Items by Test Event and Strand – Grade 3**

Achievement Level	Test Event	Number of Items	RP		RI		V		W	
			Unique	Shared	Unique	Shared	Unique	Shared	Unique	Shared
Developing	1	30	4	4	9	3	4	1	0	6
	2	32	5	2	10	0	7	2	1	5
	3	32	5	2	5	6	6	2	0	6
	4	31	8	0	7	2	7	1	0	6
On Track	1	28	3	6	2	4	1	5	0	6
	2	28	5	4	4	3	3	2	0	6
	3	31	3	5	7	3	5	3	0	6
	4	27	2	5	6	2	1	4	0	6
Advanced	1	31	9	0	8	4	4	0	0	6
	2	28	7	2	4	4	3	2	2	4
	3	29	0	8	3	4	4	3	0	6
	4	28	1	9	5	0	5	1	0	6

**Table 12. Number of Unique and Shared Items by Test Event and Strand – Grade 4**

Achievement Level	Test Event	Number of Items	RP		RI		V		W	
			Unique	Shared	Unique	Shared	Unique	Shared	Unique	Shared
Developing	1	28	5	3	4	4	1	5	4	2
	2	24	4	3	0	4	4	2	2	4
	3	29	7	3	4	2	3	4	4	2
	4	27	5	2	6	0	5	0	3	3
On Track	1	23	4	3	1	6	2	1	4	2
	2	28	4	5	1	5	3	3	3	3
	3	27	5	2	5	3	3	2	1	5
	4	27	9	0	4	3	3	1	2	4
Advanced	1	27	5	3	5	3	1	4	2	4
	2	27	1	6	4	6	3	1	2	4
	3	27	7	0	2	7	4	2	0	6
	4	29	6	2	4	3	5	3	0	6

**Table I3. Number of Unique and Shared Items by Test Event and Strand – Grade 5**

Achievement Level	Test Event	Number of Items	RP		RI		V		W	
			Unique	Shared	Unique	Shared	Unique	Shared	Unique	Shared
Developing	1	32	5	5	10	0	5	1	5	1
	2	25	4	5	2	2	1	5	1	5
	3	29	5	2	9	1	6	0	0	6
	4	32	5	6	6	2	6	1	2	4
On Track	1	26	1	6	3	5	0	5	0	6
	2	29	3	5	6	3	4	3	1	5
	3	30	5	2	5	3	6	3	0	6
	4	22	3	4	4	2	1	1	1	5
Advanced	1	29	7	4	2	5	0	5	0	6
	2	30	4	3	7	3	5	2	0	6
	3	23	0	5	7	0	3	1	0	6
	4	24	0	1	5	5	3	4	0	6

**Table I4. Number of Unique and Shared Items by Test Event and Strand – Grade 6**

Achievement Level	Test Event	Number of Items	RP		RI		V		W	
			Unique	Shared	Unique	Shared	Unique	Shared	Unique	Shared
Developing	1	23	4	6	3	0	2	2	3	3
	2	28	7	0	6	3	2	3	4	2
	3	29	7	0	5	4	4	3	4	2
	4	29	6	2	2	7	1	5	1	5
On Track	1	28	6	1	1	5	6	2	1	5
	2	30	3	7	4	3	2	5	0	6
	3	27	1	7	3	3	4	3	1	5
	4	28	4	4	2	6	2	4	1	5
Advanced	1	29	1	6	4	5	1	6	0	6
	2	29	2	5	4	5	1	6	0	6
	3	26	0	8	2	5	0	5	0	6
	4	28	1	7	4	5	3	2	0	6

**Table 15. Number of Unique and Shared Items by Test Event and Strand – Grade 7**

Achievement Level	Test Event	Number of Items	RP		RI		V		W	
			Unique	Shared	Unique	Shared	Unique	Shared	Unique	Shared
Developing	1	31	5	2	10	1	4	3	1	4
	2	29	4	3	9	0	4	3	2	4
	3	28	0	7	8	2	6	0	3	2
	4	31	8	0	6	2	6	3	1	4
On Track	1	23	0	4	5	4	3	1	2	4
	2	27	4	6	2	5	2	3	1	5
	3	28	2	5	4	6	4	1	2	4
	4	27	5	3	5	3	1	4	0	6
Advanced	1	27	1	6	7	2	3	2	0	6
	2	23	0	5	1	6	2	3	0	4
	3	27	2	6	4	5	3	2	0	6
	4	27	0	7	5	4	3	2	1	4

**Table 16. Number of Unique and Shared Items by Test Event and Strand – Grade 8**

Achievement Level	Test Event	Number of Items	RP		RI		V		W	
			Unique	Shared	Unique	Shared	Unique	Shared	Unique	Shared
Developing	1	29	10	1	6	0	6	1	3	3
	2	30	8	3	9	0	2	2	7	0
	3	24	0	4	1	4	2	4	2	4
	4	32	3	4	9	0	8	2	0	6
On Track	1	28	7	2	0	7	5	0	0	6
	2	22	2	5	0	3	1	4	0	6
	3	29	1	6	6	3	3	4	2	4
	4	30	6	1	8	0	7	1	0	6
Advanced	1	28	2	6	3	4	5	1	0	6
	2	27	1	6	3	6	0	5	0	6
	3	25	1	3	9	0	5	1	0	6
	4	27	2	6	4	3	3	1	0	6

## Appendix J. DOK by Grade, Test Event, and Strand

**Table J1. DOK by Grade, Test Event, and Strand – Grade 3**

Achievement Level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Developing	1	RP	8	0 (0%)	6 (75%)	2 (25%)
		RI	11	1 (9%)	9 (82%)*	1 (9%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	4 (57%)	0 (0%)	3 (43%)
		RI	10	0 (0%)	8 (80%)	2 (20%)
		V	9	1 (11%)	7 (78%)	1 (11%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	0 (0%)	4 (57%)	3 (43%)
		RI	11	1 (9%)	7 (64%)	3 (27%)
		V	8	0 (0%)	7 (88%)	1 (13%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	8	1 (13%)	3 (38%)	4 (50%)
		RI	9	1 (11%)	3 (33%)	5 (56%)
		V	8	1 (13%)	5 (63%)	2 (25%)
		W	6	0 (0%)	6 (100%)	0 (0%)
On Track	1	RP	9	1 (11%)	3 (33%)	5 (56%)
		RI	6	0 (0%)	3 (50%)	3 (50%)
		V	6	0 (0%)	5 (83%)	1 (17%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	9	0 (0%)	6 (67%)	3 (33%)
		RI	7	0 (0%)	6 (86%)	1 (14%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	8	1 (13%)	3 (38%)	4 (50%)
		RI	9	1 (11%)	7 (78%)*	1 (11%)
		V	8	0 (0%)	8 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	7	1 (14%)	4 (57%)	2(29%)
		RI	8	0 (0%)	8 (100%)	0 (0%)
		V	5	0 (0%)	3 (60%)	2 (40%)
		W	6	0 (0%)	6 (100%)	0 (0%)



Achievement Level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Advanced	1	RP	9	0 (0%)	6 (67%)	3 (33%)
		RI	12	3 (25%)	7 (58%)	2 (17%)
		V	4	0 (0%)	2 (50%)	2 (50%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	9	1 (11%)	6 (67%)	2 (22%)
		RI	8	0 (0%)	5 (63%)	3 (38%)
		V	5	0 (0%)	3 (60%)	2 (40%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	3	RP	8	0 (0%)	5 (63%)	3 (38%)
		RI	7	0 (0%)	5 (71%)	2 (29%)
		V	7	0 (0%)	6 (86%)	1 (14%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	10	0 (0%)	7 (70%)	3 (30%)
		RI	5	1 (20%)	4 (80%)	0 (0%)
		V	6	0 (0%)	5 (83%)	1 (17%)
		W	6	0 (0%)	6 (100%)	0 (0%)

**Notes.**

- One test item in the Developing 1 and On Track 3 test events was split between RI and RI; it received a DOK of 2.
- For the On Track achievement level, test events #1, #2, and #4 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the Advanced achievement level, test events #3 and #4 had one item rated as “None” for the standard; therefore, this item was not included in this table.

**Table J2. DOK by Grade, Test Event, and Strand – Grade 4**

Achievement Level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Developing	1	RP	8	2 (25%)	5 (63%)	1 (13%)
		RI	8	5 (63%)	3 (38%)	0 (0%)
		V	6	3 (50%)	3 (50%)	0 (0%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	2	RP	7	2 (29%)	5 (71%)	0 (0%)
		RI	4	4 (100%)	0 (0%)	0 (0%)
		V	6	3 (50%)	3 (50%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	3	RP	10	2 (20%)	5 (50%)	3 (30%)
		RI	6	2 (33%)	4 (67%)	0 (0%)
		V	7	4 (57%)	3 (43%)	0 (0%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	4	RP	7	1 (14%)	3 (43%)	3 (43%)
		RI	6	4 (67%)	1 (17%)	1 (17%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
On Track	1	RP	7	2 (29%)	3 (43%)	2 (29%)
		RI	7	3 (43%)	4 (57%)	0 (0%)
		V	3	0 (0%)	3 (100%)	0 (0%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	2	RP	9	0 (0%)	9 (100%)	0 (0%)
		RI	6	1 (17%)	5 (83%)	0 (0%)
		V	6	4 (67%)	2 (33%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	3	RP	7	0 (0%)	5 (71%)	2 (29%)
		RI	8	2 (25%)	5 (63%)	1 (13%)
		V	5	1 (20%)	4 (80%)	0 (0%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	4	RP	9	1 (11%)	7 (78%)	1 (11%)
		RI	7	1 (14%)	5 (71%)	1 (14%)
		V	4	2 (50%)	2 (50%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)

Achievement Level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Advanced	1	RP	8	0 (0%)	7 (88%)	1 (13%)
		RI	8	2 (25%)	5 (63%)	1 (13%)
		V	5	1 (20%)	4 (80%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	0 (0%)	5 (71%)	2 (29%)
		RI	9	1 (11%)	6 (66%)	2 (22%)*
		V	4	2 (50%)	2 (50%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	3	RP	7	0 (0%)	4 (57%)	3 (43%)
		RI	8	2 (25%)	3 (38%)	3 (38%)*
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	4	RP	8	2 (25%)	5 (63%)	1 (13%)
		RI	7	1 (14%)	5 (71%)	1 (14%)
		V	8	4 (50%)	4 (50%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)

**Notes.**

- One test item in the Advanced 2 and Advanced 3 test events was split between RI and RI; it received a DOK of 3.
- For the Developing achievement level, test event #2 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the Developing achievement level, test event #4 had three items rated as “None” for the standard; therefore, this item was not included in this table.
- For the On Track achievement level, test events #2, #3, and #4 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the Advanced achievement level, test event #2 had one item rated as “None” for the standard; therefore, this item was not included in this table.

**Table J3. DOK by Grade, Test Event, and Strand – Grade 5**

Achievement Level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Developing	1	RP	10	0 (0%)	10 (100%)	0 (0%)
		RI	10	1 (10%)	9 (90%)	0 (0%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	2	RP	9	0 (0%)	8 (89%)	1 (11%)
		RI	4	0 (0%)	3 (75%)	1 (25%)
		V	6	2 (33%)	4 (67%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	1 (14%)	6 (86%)	0 (0%)
		RI	10	1 (10%)	7 (70%)	2 (20%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	11	2 (18%)	9 (82%)	0 (0%)
		RI	8	0 (0%)	8 (100%)	0 (0%)
		V	7	3 (43%)	4 (57%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
On Track	1	RP	7	0 (0%)	7 (100%)	0 (0%)
		RI	8	0 (0%)	7 (88%)	1 (13%)
		V	5	2 (40%)	3 (60%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	0 (0%)	6 (86%)	1 (14%)
		RI	9	0 (0%)	9 (100%)	0 (0%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	0 (0%)	6 (86%)	1 (14%)
		RI	8	1 (13%)	7 (88%)	0 (0%)
		V	9	2 (22%)	7 (78%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	7	0 (0%)	6 (86%)	1 (14%)
		RI	6	1 (17%)	5 (83%)	0 (0%)
		V	2	0 (0%)	2 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)

Achievement Level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Advanced	1	RP	11	0 (0%)	11 (100%)	0 (0%)
		RI	7	0 (0%)	6 (86%)	1 (14%)
		V	5	2 (40%)	3 (60%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	0 (0%)	7 (100%)	0 (0%)
		RI	10	1 (10%)	7 (70%)	2 (20%)
		V	7	1 (14%)	6 (87%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	5	0 (0%)	4 (80%)	1 (20%)
		RI	7	1 (14%)	6 (86%)	0 (0%)
		V	4	0 (0%)	4 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	1	0 (0%)	1 (100%)	0 (0%)
		RI	10	0 (0%)	9 (90%)	1 (10%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)

**Notes.**

- One test item in the Developing 1 test event was split between RP and RP; it received a DOK of 2.
- One test item in the Developing 2 test event was split between W and W; it received a DOK of 2.
- One test item in the On Track 2 test event was split between RP and RP; it received a DOK of 2.
- For the On Track achievement level, test event #4 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the Advanced achievement level, test event #3 had one item rated as “None” for the standard; therefore, this item was not included in this table.

**Table J4. DOK by Grade, Test Event, and Strand – Grade 6**

Achievement level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Developing	1	RP	10	1 (10%)	8 (80%)	1 (10%)
		RI	3	0 (0%)	3 (100%)	0 (0%)
		V	4	0 (0%)	3 (75%)	1 (25%)
		W	6	0 (0%)	4 (67%)	2 (33%)
	2	RP	7	2 (29%)	3 (43%)	2 (29%)
		RI	9	3 (33%)	5 (56%)	1 (11%)
		V	5	2 (40%)	3 (60%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	3	RP	7	2 (29%)	5 (71%)	0 (0%)
		RI	9	2 (22%)	5 (56%)	2 (22%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	0 (0%)	3 (50%)	3 (50%)
	4	RP	8	0 (0%)	8 (100%)	0 (0%)
		RI	9	2 (22%)	6 (67%)	1 (11%)
		V	6	2 (33%)	4 (67%)	0 (0%)
		W	6	0 (0%)	4 (67%)	2 (33%)
On Track	1	RP	7	0 (0%)	6 (86%)	1 (14%)
		RI	6	0 (0%)	5 (83%)	1 (17%)
		V	8	0 (0%)	7 (88%)	1 (13%)
		W	6	0 (0%)	4 (67%)	2 (33%)
	2	RP	10	2 (20%)	8 (80%)	0 (0%)
		RI	7	1 (14%)	6 (86%)	0 (0%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	0 (0%)	3 (50%)	3 (50%)
	3	RP	8	0 (0%)	7 (88%)	1 (13%)
		RI	6	0 (0%)	6 (100%)	0 (0%)
		V	7	0 (0%)	7 (100%)	0 (0%)
		W	6	0 (0%)	3 (50%)	3 (50%)
	4	RP	8	0 (0%)	4 (50%)	4 (50%)
		RI	8	1 (13%)	6 (75%)	1 (13%)
		V	6	0 (0%)	5 (83%)	1 (17%)
		W	6	0 (0%)	3 (50%)	3 (50%)

Achievement level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Advanced	1	RP	7	0 (0%)	6 (86%)	1 (14%)
		RI	9	0 (0%)	8 (89%)	1 (11%)
		V	7	0 (0%)	7 (100%)	0 (0%)
		W	6	0 (0%)	1 (17%)	5 (83%)
	2	RP	7	0 (0%)	5 (71%)	2 (29%)
		RI	9	0 (0%)	8 (89%)	1 (11%)
		V	7	0 (0%)	7 (100%)	0 (0%)
		W	6	0 (0%)	1 (17%)	5 (83%)
	3	RP	8	0 (0%)	7 (88%)	1 (13%)
		RI	7	0 (0%)	6 (86%)	1 (14%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	0 (0%)	2 (33%)	4 (67%)
	4	RP	8	1 (13%)	7 (88%)	0 (0%)
		RI	9	3 (33%)	5 (56%)	1 (11%)
		V	5	1 (20%)	4 (80%)	0 (0%)
		W	6	0 (0%)	1 (17%)	5 (83%)

**Notes.**

- For the Developing achievement level, test event #2 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the On Track achievement level, test event #1 had one item rated as “None” for the standard; therefore, this item was not included in this table.

**Table J5. DOK by Grade, Test Event, and Strand – Grade 7**

Achievement level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Developing	1	RP	7	1 (14%)	4 (57%)	2 (29%)
		RI	11	3 (27%)	7 (64%)	1 (9%)
		V	7	0 (0%)	6 (86%)	1 (14%)
		W	5	0 (0%)	3 (60%)	1 (20%)
	2	RP	7	0 (0%)	5 (71%)	2 (29%)
		RI	9	0 (0%)	7 (78%)	2 (22%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	0 (0%)	4 (67%)	2 (33%)
	3	RP	7	0 (0%)	4 (57%)	3 (43%)
		RI	9	1 (11%)	5 (56%)	2 (22%)
		V	6	0 (0%)	5 (83%)	1 (17%)
		W	5	0 (0%)	4 (80%)	1 (20%)
	4	RP	8	1 (13%)	4 (50%)	3 (38%)
		RI	8	1 (13%)	5 (63%)	2 (25%)
		V	9	0 (0%)	8 (89%)	1 (11%)
		W	5	0 (0%)	3 (60%)	2 (40%)
On Track	1	RP	4	0 (0%)	1 (25%)	3 (75%)
		RI	9	2 (22%)	3 (33%)	4 (44%)
		V	4	0 (0%)	4 (100%)	0 (0%)
		W	6	0 (0%)	3 (50%)	2 (33%)
	2	RP	10	0 (0%)	7 (70%)	3 (30%)
		RI	7	0 (0%)	6 (86%)	1 (14%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	6	0 (0%)	3 (50%)	3 (50%)
	3	RP	7	0 (0%)	5 (71%)	2 (29%)
		RI	10	2 (20%)	8 (80%)	0 (0%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	0 (0%)	4 (67%)	2 (33%)
	4	RP	8	0 (0%)	7 (88%)	1 (13%)
		RI	8	0 (0%)	4 (50%)	4 (50%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	6	0 (0%)	3 (50%)	3 (50%)



Achievement level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Advanced	1	RP	7	0 (0%)	5 (71%)	2 (29%)
		RI	9	1 (11%)	5 (56%)	3 (33%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	0 (0%)	3 (50%)	3 (50%)
	2	RP	5	0 (0%)	5 (100%)	0 (0%)
		RI	7	0 (0%)	5 (71%)	2 (29%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	4	0 (0%)	3 (75%)	1 (25%)
	3	RP	7	0 (0%)	4 (57%)	3 (43%)
		RI	9	0 (0%)	8 (89%)	1 (11%)
		V	5	0 (0%)	1 (20%)	4 (80%)
		W	6	0 (0%)	4 (67%)	2 (33%)
	4	RP	7	0 (0%)	7 (100%)	0 (0%)
		RI	9	0 (0%)	5 (56%)	4 (44%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	5	0 (0%)	3 (60%)	2 (40%)

**Notes.**

- One test item in the Developing 3 test event was split between RI and RI; it did not receive a DOK rating.
- One test item in the On Track 2 test event was split between RP and V; it received a DOK of 2.
- One test item in the Advanced 3 test event was split between RP and RP; it received a DOK of 3.
- For the Developing achievement level, test events #1, #3, and #4 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the Advanced achievement level, test event #2 had two items rated as “None” for the standard; therefore, this item was not included in this table.
- For the Advanced achievement level, test event #4 had one item rated as “None” for the standard; therefore, this item was not included in this table.

**Table J6. DOK by Grade, Test Event, and Strand – Grade 8**

Achievement level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Developing	1	RP	10	0 (0%)	10 (100%)	0 (0%)
		RI	6	1 (17%)	5 (83%)	0 (0%)
		V	7	1 (14%)	5 (71%)	1 (14%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	11	2 (18%)	9 (82%)	0 (0%)
		RI	9	2 (22%)	7 (78%)	0 (0%)
		V	4	1 (25%)	3 (75%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	4	0 (0%)	3 (75%)	1 (25%)
		RI	5	1 (20%)	2 (40%)	2 (40%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	4	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	9	4 (44%)	4 (44%)	1 (11%)
		V	10	1 (10%)	9 (90%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
On Track	1	RP	9	0 (0%)	9 (100%)	0 (0%)
		RI	7	1 (14%)	6 (86%)	0 (0%)
		V	5	1 (20%)	4 (80%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	0 (0%)	6 (86%)	1 (14%)
		RI	3	0 (0%)	1 (33%)	2 (67%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	1 (14%)	4 (57%)	2 (29%)
		RI	9	1 (11%)	7 (78%)	1 (11%)
		V	7	1 (14%)	6 (86%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	7	0 (0%)	6 (86%)	1 (14%)
		RI	8	0 (0%)	7 (88%)	1 (13%)
		V	8	2 (25%)	6 (75%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)

Achievement level	Test Event	Strand	Number of Items	DOK 1	DOK 2	DOK 3
Advanced	1	RP	8	1 (13%)	6 (75%)	1 (13%)
		RI	7	2 (29%)	4 (57%)	1 (14%)
		V	6	0 (0%)	6 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	9	0 (0%)	7 (78%)	2 (22%)
		V	5	1 (20%)	4 (80%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	4	0 (0%)	2 (50%)	2 (50%)
		RI	9	0 (0%)	9 (100%)	0 (0%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	8	1 (13%)	6 (75%)	1 (13%)
		RI	7	4 (57%)	3 (43%)	0 (0%)
		V	4	0 (0%)	4 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)

**Notes.**

- For the Developing achievement level, test event #3 had three items rated as “None” for the standard; therefore, this item was not included in this table.
- For the On Track achievement level, test events #1, #2, and #4 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the Advanced achievement level, test event #1 had one item rated as “None” for the standard; therefore, this item was not included in this table.
- For the Advanced achievement level, test event #4 had two items rated as “None” for the standard; therefore, this item was not included in this table.
- One test item in the Developing 1 test event was split between RP and RP; it received a DOK of 2.
- One test item in the Developing 2 test event was split between W and W; it received a DOK of 2.

## Appendix K. ALD by Grade, Test Event, and Strand

**Table K1. ALD by Grade, Test Event, and Strand – Grade 3**

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	RP	8	1 (13%)	5 (63%)	2 (25%)
		RI	10	1 (10%)	7 (70%)	2 (20%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	3 (43%)	2 (29%)	2 (29%)
		RI	10	1 (10%)	8 (80%)	1 (10%)
		V	9	2 (22%)	5 (56%)	2 (22%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	0 (0%)	5 (71%)	2 (29%)
		RI	11	0 (0%)	6 (55%)	5 (45%)
		V	8	0 (0%)	6 (75%)	2 (25%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	4	RP	8	1 (13%)	5 (63%)	2 (25%)
		RI	9	1 (11%)	5 (56%)	3 (33%)
		V	8	0 (0%)	7 (88%)	1 (13%)
		W	6	0 (0%)	6 (100%)	0 (0%)
On Track	1	RP	9	2 (22%)	3 (33%)	4 (44%)
		RI	6	0 (0%)	2 (33%)	4 (67%)
		V	6	0 (0%)	3 (50%)	3 (50%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	2	RP	9	0 (0%)	7 (78%)	2 (22%)
		RI	7	2 (29%)	5 (71%)	0 (0%)
		V	5	0 (0%)	3 (60%)	2 (40%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	3	RP	8	2 (25%)	4 (50%)	2 (25%)
		RI	8	0 (0%)	8 (100%)	0 (0%)
		V	8	0 (0%)	5 (63%)	3 (38%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	4	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	8	0 (0%)	7 (88%)	1 (13%)
		V	5	0 (0%)	1 (20%)	4 (80%)
		W	6	0 (0%)	6 (100%)	0 (0%)

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Advanced	1	RP	9	1 (11%)	6 (67%)	2 (22%)
		RI	12	0 (0%)	9 (75%)	3 (25%)
		V	4	0 (0%)	0 (0%)	4 (100%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	2	RP	9	3 (33%)	3 (33%)	3 (33%)
		RI	8	0 (0%)	4 (50%)	4 (50%)
		V	5	0 (0%)	2 (40%)	3 (60%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	8	0 (0%)	6 (75%)	2 (25%)
		RI	7	1 (14%)	5 (71%)	1 (14%)
		V	7	0 (0%)	1 (14%)	6 (86%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	10	0 (0%)	8 (80%)	2 (20%)
		RI	5	1 (20%)	4 (80%)	0 (0%)
		V	6	0 (0%)	3 (50%)	3 (50%)
		W	6	1 (17%)	5 (83%)	0 (0%)

**Table K2. ALD by Grade, Test Event, and Strand – Grade 4**

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	RP	8	3 (38%)	4 (50%)	1 (13%)
		RI	8	5 (63%)	3 (38%)	0 (0%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	2	RP	7	2 (29%)	5 (71%)	0 (0%)
		RI	4	4 (100%)	0 (0%)	0 (0%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	3	RP	10	2 (20%)	7 (70%)	1 (10%)
		RI	6	3 (50%)	2 (33%)	1 (17%)
		V	7	1 (14%)	5 (71%)	1 (14%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	4	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	6	4 (67%)	2 (33%)	0 (0%)
		V	5	1 (20%)	3 (60%)	1 (20%)
		W	6	1 (17%)	4 (67%)	1 (17%)
On Track	1	RP	7	3 (43%)	4 (57%)	0 (0%)
		RI	7	5 (71%)	2 (29%)	0 (0%)
		V	3	0 (0%)	2 (67%)	1 (33%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	2	RP	9	3 (33%)	5 (56%)	1 (11%)
		RI	6	3 (50%)	3 (50%)	0 (0%)
		V	6	1 (17%)	4 (67%)	1 (17%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	3	RP	7	2 (29%)	4 (57%)	1 (14%)
		RI	8	4 (50%)	4 (50%)	0 (0%)
		V	5	0 (0%)	1 (20%)	4 (80%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	4	RP	9	3 (33%)	3 (33%)	3 (33%)
		RI	7	3 (43%)	4 (57%)	0 (0%)
		V	4	1 (25%)	2 (50%)	1 (25%)
		W	6	1 (17%)	5 (83%)	0 (0%)

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Advanced	1	RP	8	2 (25%)	6 (75%)	0 (0%)
		RI	8	4 (50%)	4 (50%)	0 (0%)
		V	5	1 (20%)	2 (40%)	2 (40%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	2	RP	7	1 (14%)	6 (86%)	0 (0%)
		RI	8	3 (38%)	5 (63%)	0 (0%)
		V	4	0 (0%)	3 (75%)	1 (25%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	3	RP	7	1 (14%)	4 (57%)	2 (29%)
		RI	7	2 (29%)	5 (71%)	0 (0%)
		V	6	1 (17%)	3 (50%)	2 (33%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	4	RP	8	3 (38%)	3 (38%)	2 (25%)
		RI	7	2 (29%)	4 (57%)	1 (14%)
		V	8	1 (13%)	6 (75%)	1 (13%)
		W	6	1 (17%)	5 (83%)	0 (0%)

**Table K3. ALD by Grade, Test Event, and Strand – Grade 5**

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	RP	10	1 (10%)	9 (90%)	0 (0%)
		RI	10	0 (0%)	10 (100%)	0 (0%)
		V	6	2 (33%)	4 (67%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	2	RP	9	1 (11%)	8 (89%)	0 (0%)
		RI	4	0 (0%)	4 (100%)	0 (0%)
		V	6	3 (50%)	3 (50%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	2 (29%)	5 (71%)	0 (0%)
		RI	10	3 (30%)	6 (60%)	1 (10%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	4	RP	11	3 (27%)	8 (73%)	0 (0%)
		RI	8	1 (13%)	7 (88%)	0 (0%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
On Track	1	RP	7	0 (0%)	7 (100%)	0 (0%)
		RI	8	0 (0%)	8 (100%)	0 (0%)
		V	5	4 (80%)	1 (20%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	6	1 (17%)	5 (83%)	0 (0%)
		RI	9	1 (11%)	8 (89%)	0 (0%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	3	RP	7	0 (0%)	7 (100%)	0 (0%)
		RI	8	0 (0%)	8 (100%)	0 (0%)
		V	9	2 (22%)	7 (78%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	7	1 (14%)	6 (86%)	0 (0%)
		RI	6	2 (33%)	4 (67%)	0 (0%)
		V	2	0 (0%)	1 (50%)	1 (50%)
		W	6	1 (17%)	5 (83%)	0 (0%)



Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Advanced	1	RP	11	3 (27%)	8 (73%)	0 (0%)
		RI	7	0 (0%)	7 (100%)	0 (0%)
		V	5	3 (60%)	2 (40%)	0 (0%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	2	RP	7	0 (0%)	7 (100%)	0 (0%)
		RI	10	1 (10%)	8 (80%)	1 (10%)
		V	7	1 (14%)	6 (86%)	0 (0%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	3	RP	5	1 (20%)	4 (80%)	0 (0%)
		RI	7	3 (43%)	4 (57%)	0 (0%)
		V	4	0 (0%)	4 (100%)	0 (0%)
		W	6	2 (33%)	4 (67%)	0 (0%)
	4	RP	1	0 (0%)	1 (100%)	0 (0%)
		RI	10	0 (0%)	9 (90%)	1 (10%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	2 (33%)	4 (67%)	0 (0%)

**Table K4. ALD by Grade, Test Event, and Strand – Grade 6**

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	RP	10	1 (10%)	9 (90%)	0 (0%)
		RI	3	0 (0%)	3 (100%)	0 (0%)
		V	4	1 (25%)	3 (75%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	2	RP	7	4 (57%)	3 (43%)	0 (0%)
		RI	9	4 (44%)	5 (56%)	0 (0%)
		V	5	2 (40%)	3 (60%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	3 (43%)	3 (43%)	1 (14%)
		RI	9	3 (33%)	6 (67%)	0 (0%)
		V	7	2 (29%)	4 (57%)	1 (14%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	8	1 (13%)	6 (75%)	1 (13%)
		RI	9	3 (33%)	6 (67%)	0 (0%)
		V	6	2 (33%)	3 (50%)	1 (17%)
		W	6	0 (0%)	6 (100%)	0 (0%)
On Track	1	RP	7	1 (14%)	6 (86%)	0 (0%)
		RI	6	1 (17%)	4 (67%)	1 (17%)
		V	8	0 (0%)	6 (75%)	2 (25%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	10	1 (10%)	7 (70%)	2 (20%)
		RI	7	1 (14%)	6 (86%)	0 (0%)
		V	7	2 (29%)	5 (71%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	8	0 (0%)	7 (88%)	1 (13%)
		RI	6	2 (33%)	4 (67%)	0 (0%)
		V	7	0 (0%)	7 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	8	0 (0%)	5 (63%)	3 (38%)
		RI	8	1 (13%)	6 (75%)	1 (13%)
		V	6	0 (0%)	4 (67%)	2 (33%)
		W	6	0 (0%)	6 (100%)	0 (0%)

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Advanced	1	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	9	2 (22%)	7 (78%)	0 (0%)
		V	7	1 (14%)	6 (86%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	2	RP	7	1 (14%)	4 (57%)	2 (29%)
		RI	9	1 (11%)	7 (78%)	1 (11%)
		V	7	1 (14%)	5 (71%)	1 (14%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	3	RP	8	0 (0%)	7 (88%)	1 (13%)
		RI	7	0 (0%)	7 (100%)	0 (0%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	4	RP	8	0 (0%)	6 (75%)	2 (25%)
		RI	9	3 (33%)	6 (67%)	0 (0%)
		V	5	2 (40%)	3 (60%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)

**Table K5. ALD by Grade, Test Event, and Strand – Grade 7**

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	11	7 (64%)	4 (36%)	0 (0%)
		V	7	2 (29%)	2 (29%)	3 (43%)
		W	5	2 (40%)	3 (60%)	0 (0%)
	2	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	9	2 (22%)	7 (78%)	0 (0%)
		V	7	0 (0%)	5 (71%)	2 (29%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	3	RP	7	1 (14%)	4 (57%)	2 (29%)
		RI	9	3 (33%)	5 (56%)	0 (0%)
		V	6	0 (0%)	2 (33%)	4 (67%)
		W	5	0 (0%)	4 (80%)	1 (20%)
	4	RP	8	2 (25%)	4 (50%)	2 (25%)
		RI	8	2 (25%)	6 (75%)	0 (0%)
		V	9	2 (22%)	2 (22%)	5 (56%)
		W	5	2 (40%)	3 (60%)	0 (0%)
On Track	1	RP	4	0 (0%)	2 (50%)	2 (50%)
		RI	9	3 (33%)	4 (44%)	2 (22%)
		V	4	0 (0%)	1 (25%)	3 (75%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	2	RP	9	1 (11%)	7 (78%)	1 (11%)
		RI	7	5 (71%)	1 (14%)	1 (14%)
		V	4	0 (0%)	3 (75%)	1 (25%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	10	3 (30%)	7 (70%)	0 (0%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	6	2 (33%)	3 (50%)	1 (17%)
	4	RP	8	2 (25%)	5 (63%)	1 (13%)
		RI	8	2 (25%)	5 (63%)	1 (13%)
		V	5	1 (20%)	3 (60%)	1 (20%)
		W	6	1 (17%)	5 (83%)	0 (0%)

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Advanced	1	RP	7	2 (29%)	4 (57%)	1 (14%)
		RI	9	3 (33%)	5 (56%)	1 (11%)
		V	5	0 (0%)	3 (60%)	2 (40%)
		W	6	0 (0%)	4 (67%)	2 (33%)
	2	RP	5	0 (0%)	4 (80%)	1 (20%)
		RI	7	1 (14%)	4 (57%)	2 (29%)
		V	5	0 (0%)	2 (40%)	3 (60%)
		W	4	0 (0%)	4 (100%)	0 (0%)
	3	RP	6	0 (0%)	3 (50%)	3 (50%)
		RI	9	2 (22%)	7 (78%)	0 (0%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	6	1 (17%)	4 (67%)	1 (17%)
	4	RP	7	2 (29%)	5 (71%)	0 (0%)
		RI	9	1 (11%)	7 (78%)	1 (11%)
		V	5	0 (0%)	4 (80%)	1 (20%)
		W	5	0 (0%)	4 (80%)	1 (20%)

**Table K6. ALD by Grade, Test Event, and Strand – Grade 8**

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Developing	1	RP	9	0 (0%)	8 (89%)	1 (11%)
		RI	6	1 (17%)	5 (83%)	0 (0%)
		V	7	2 (29%)	4 (57%)	1 (14%)
		W	6	1 (17%)	5 (83%)	0 (0%)
	2	RP	11	3 (27%)	7 (64%)	1 (9%)
		RI	9	4 (44%)	5 (56%)	0 (0%)
		V	4	1 (25%)	3 (75%)	0 (0%)
		W	5	2 (40%)	3 (60%)	0 (0%)
	3	RP	4	1 (25%)	2 (50%)	1 (25%)
		RI	5	2 (40%)	3 (60%)	0 (0%)
		V	6	1 (17%)	5 (83%)	0 (0%)
		W	6	0 (0%)	5 (83%)	1 (17%)
	4	RP	7	1 (14%)	5 (71%)	1 (14%)
		RI	9	4 (44%)	5 (56%)	0 (0%)
		V	10	1 (10%)	9 (90%)	0 (0%)
		W	6	1 (17%)	4 (67%)	1 (17%)
On Track	1	RP	9	1 (11%)	8 (89%)	0 (0%)
		RI	7	0 (0%)	7 (100%)	0 (0%)
		V	5	1 (20%)	4 (80%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	0 (0%)	4 (57%)	3 (43%)
		RI	3	2 (67%)	1 (33%)	0 (0%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	7	1 (14%)	6 (86%)	0 (0%)
		RI	9	2 (22%)	5 (56%)	2 (22%)
		V	7	0 (0%)	7 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	7	1 (14%)	6 (86%)	0 (0%)
		RI	8	3 (38%)	4 (50%)	1 (13%)
		V	8	0 (0%)	7 (88%)	1 (13%)
		W	6	1 (17%)	5 (83%)	0 (0%)

Achievement Level	Test Event	Strand	Number of Items	ALD 1	ALD 2	ALD 3
Advanced	1	RP	8	0 (0%)	6 (75%)	2 (25%)
		RI	7	1 (14%)	5 (71%)	1 (14%)
		V	6	0 (0%)	5 (83%)	1 (17%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	2	RP	7	1 (14%)	6 (86%)	0 (0%)
		RI	9	0 (0%)	6 (67%)	3 (33%)
		V	5	0 (0%)	5 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	3	RP	4	0 (0%)	1 (25%)	3 (75%)
		RI	9	1 (11%)	8 (89%)	0 (0%)
		V	6	1 (17%)	3 (50%)	2 (33%)
		W	6	0 (0%)	6 (100%)	0 (0%)
	4	RP	8	2 (25%)	4 (50%)	2 (25%)
		RI	7	3 (43%)	3 (43%)	1 (14%)
		V	4	0 (0%)	4 (100%)	0 (0%)
		W	6	0 (0%)	6 (100%)	0 (0%)

## Appendix L. Process Evaluation Tables by Grade

**Table L1. Panelist Evaluation Survey Results – Grade 3**

Question	Average	
My panel facilitator did an effective job of facilitating discussion and ensuring that all panelists' perspectives were heard	4.75	<div></div>
The panel facilitator was helpful during the workshop	4.75	<div></div>
Everyone had equal opportunity to contribute ideas and opinions	4.50	<div></div>
My ideas and opinions were listened to and respected by the group	4.50	<div></div>
My panel facilitator clearly and promptly addressed my questions	4.50	<div></div>
The other materials shared by my facilitator were useful	4.50	<div></div>
The whole-group training facilitator was helpful during the workshop	4.50	<div></div>
It was easy to access the evaluation and demographics forms	4.25	<div></div>
It was easy to access the items on the Content Review Tool	4.25	<div></div>
Other support staff were helpful during the workshop	4.25	<div></div>
Practicing making ratings as a group in my assigned panel helped me better understand the alignment activities	4.25	<div></div>
The Content Review Tool allowed me to effectively accomplish my tasks during the workshop	4.25	<div></div>
The Google Rating Sheet was useful for recording alignment ratings	4.25	<div></div>
The hands-on training in my assigned panel helped me better understand the alignment activities	4.25	<div></div>
The hands-on training in my assigned panel was an effective use of time	4.25	<div></div>
The materials hosted on Google Drive were useful (e.g., standards)	4.25	<div></div>
The panel-specific hands-on training was well organized	4.25	<div></div>
The Google Rating Sheet provided a comprehensive platform for capturing alignment of standards	4.00	<div></div>
The group-wide training session provided a useful overview of the alignment activities for the week	4.00	<div></div>
The group-wide training session was well organized	4.00	<div></div>
The group-wide training was an effective use of time	4.00	<div></div>
The group-wide training session clearly described my role as a panelist	3.75	<div></div>
The group-wide training session effectively outlined the purpose of the alignment workshop	3.75	<div></div>

Note: Strongly Disagree = 1 to Strongly Agree = 5



**Table L2. Panelist Evaluation Survey Results – Grade 4**

Question	Average	
Practicing making ratings as a group in my assigned panel helped me better understand the alignment activities	4.67	<div></div>
The hands-on training in my assigned panel was an effective use of time	4.67	<div></div>
My panel facilitator clearly and promptly addressed my questions	4.67	<div></div>
Everyone had equal opportunity to contribute ideas and opinions	4.67	<div></div>
My ideas and opinions were listened to and respected by the group	4.67	<div></div>
The hands-on training in my assigned panel helped me better understand the alignment activities	4.33	<div></div>
The panel-specific hands-on training was well organized	4.33	<div></div>
My panel facilitator did an effective job of facilitating discussion and ensuring that all panelists' perspectives were heard	4.33	<div></div>
The panel facilitator was helpful during the workshop	4.33	<div></div>
The group-wide training session effectively outlined the purpose of the alignment workshop	4.00	<div></div>
The group-wide training session provided a useful overview of the alignment activities for the week	4.00	<div></div>
The group-wide training session clearly described my role as a panelist	4.00	<div></div>
The group-wide training session was well organized	4.00	<div></div>
The group-wide training was an effective use of time	4.00	<div></div>
The materials hosted on Google Drive were useful (e.g., standards)	4.00	<div></div>
The Google Rating Sheet provided a comprehensive platform for capturing alignment of standards	4.00	<div></div>
It was easy to access the items on the Content Review Tool	4.00	<div></div>
The Content Review Tool allowed me to effectively accomplish my tasks during the workshop	4.00	<div></div>
It was easy to access the evaluation and demographics forms	4.00	<div></div>
The whole-group training facilitator was helpful during the workshop	4.00	<div></div>
The other materials shared by my facilitator were useful	3.67	<div></div>
Other support staff were helpful during the workshop	3.67	<div></div>
The Google Rating Sheet was useful for recording alignment ratings	3.33	<div></div>

Note: Strongly Disagree = 1 to Strongly Agree = 5

**Table L3. Panelist Evaluation Survey Results – Grade 5**

Question	Average	
My panel facilitator clearly and promptly addressed my questions	5.00	
My panel facilitator did an effective job of facilitating discussion and ensuring that all panelists' perspectives were heard	5.00	
The panel facilitator was helpful during the workshop	5.00	
The materials hosted on Google Drive were useful (e.g., standards)	4.80	
It was easy to access the evaluation and demographics forms	4.80	
Practicing making ratings as a group in my assigned panel helped me better understand the alignment activities	4.60	
The Google Rating Sheet was useful for recording alignment ratings	4.60	
The Google Rating Sheet provided a comprehensive platform for capturing alignment of standards	4.60	
The other materials shared by my facilitator were useful	4.60	
The group-wide training session was well organized	4.50	
The group-wide training session effectively outlined the purpose of the alignment workshop	4.40	
The group-wide training was an effective use of time	4.40	
The panel-specific hands-on training was well organized	4.40	
The hands-on training in my assigned panel was an effective use of time	4.40	
Everyone had equal opportunity to contribute ideas and opinions	4.40	
My ideas and opinions were listened to and respected by the group	4.40	
The whole-group training facilitator was helpful during the workshop	4.40	
Other support staff were helpful during the workshop	4.40	
The group-wide training session provided a useful overview of the alignment activities for the week	4.20	
The group-wide training session clearly described my role as a panelist	4.20	
The hands-on training in my assigned panel helped me better understand the alignment activities	4.20	
It was easy to access the items on the Content Review Tool	4.00	
The Content Review Tool allowed me to effectively accomplish my tasks during the workshop	4.00	

Note: Strongly Disagree = 1 to Strongly Agree = 5

**Table L4. Panelist Evaluation Survey Results – Grade 6**

Question	Average
My panel facilitator clearly and promptly addressed my questions	5.00
My panel facilitator did an effective job of facilitating discussion and ensuring that all panelists' perspectives were heard	5.00
The materials hosted on Google Drive were useful (e.g., standards)	5.00
The Google Rating Sheet was useful for recording alignment ratings	5.00
The hands-on training in my assigned panel helped me better understand the alignment activities	4.83
Practicing making ratings as a group in my assigned panel helped me better understand the alignment activities	4.83
The hands-on training in my assigned panel was an effective use of time	4.83
Everyone had equal opportunity to contribute ideas and opinions	4.83
My ideas and opinions were listened to and respected by the group	4.83
The Google Rating Sheet provided a comprehensive platform for capturing alignment of standards	4.83
The other materials shared by my facilitator were useful	4.83
The whole-group training facilitator was helpful during the workshop	4.83
The panel facilitator was helpful during the workshop	4.83
Other support staff were helpful during the workshop	4.83
The panel-specific hands-on training was well organized	4.67
The Content Review Tool allowed me to effectively accomplish my tasks during the workshop	4.50
It was easy to access the evaluation and demographics forms	4.50
The group-wide training session effectively outlined the purpose of the alignment workshop	4.17
The group-wide training session provided a useful overview of the alignment activities for the week	4.17
The group-wide training session clearly described my role as a panelist	4.17
The group-wide training session was well organized	4.17
It was easy to access the items on the Content Review Tool	4.17
The group-wide training was an effective use of time	4.00

Note: Strongly Disagree = 1 to Strongly Agree = 5



**Table L5. Panelist Evaluation Survey Results – Grade 7**

Question	Average
My panel facilitator clearly and promptly addressed my questions	5.00
My panel facilitator did an effective job of facilitating discussion and ensuring that all panelists' perspectives were heard	5.00
Everyone had equal opportunity to contribute ideas and opinions	5.00
My ideas and opinions were listened to and respected by the group	5.00
The group-wide training session effectively outlined the purpose of the alignment workshop	4.80
The group-wide training session provided a useful overview of the alignment activities for the week	4.80
The group-wide training session clearly described my role as a panelist	4.80
The group-wide training session was well organized	4.80
The group-wide training was an effective use of time	4.80
Practicing making ratings as a group in my assigned panel helped me better understand the alignment activities	4.80
The panel-specific hands-on training was well organized	4.80
The hands-on training in my assigned panel was an effective use of time	4.80
The materials hosted on Google Drive were useful (e.g., standards)	4.80
The Google Rating Sheet was useful for recording alignment ratings	4.80
The Google Rating Sheet provided a comprehensive platform for capturing alignment of standards	4.80
The other materials shared by my facilitator were useful	4.80
It was easy to access the evaluation and demographics forms	4.80
The whole-group training facilitator was helpful during the workshop	4.80
The panel facilitator was helpful during the workshop	4.80
Other support staff were helpful during the workshop	4.80
The hands-on training in my assigned panel helped me better understand the alignment activities	4.60
The Content Review Tool allowed me to effectively accomplish my tasks during the workshop	4.60
It was easy to access the items on the Content Review Tool	4.20

Note: Strongly Disagree = 1 to Strongly Agree = 5

**Table L6. Panelist Evaluation Survey Results – Grade 8**

Question	Average	
Practicing making ratings as a group in my assigned panel helped me better understand the alignment activities	5.00	
The panel-specific hands-on training was well organized	5.00	
The hands-on training in my assigned panel was an effective use of time	5.00	
My panel facilitator clearly and promptly addressed my questions	5.00	
My panel facilitator did an effective job of facilitating discussion and ensuring that all panelists' perspectives were heard	5.00	
Everyone had equal opportunity to contribute ideas and opinions	5.00	
My ideas and opinions were listened to and respected by the group	5.00	
The materials hosted on Google Drive were useful (e.g., standards)	5.00	
The panel facilitator was helpful during the workshop	5.00	
The group-wide training session clearly described my role as a panelist	4.86	
The Google Rating Sheet was useful for recording alignment ratings	4.86	
The Google Rating Sheet provided a comprehensive platform for capturing alignment of standards	4.86	
The other materials shared by my facilitator were useful	4.86	
It was easy to access the evaluation and demographics forms	4.86	
The group-wide training session effectively outlined the purpose of the alignment workshop	4.71	
The group-wide training session provided a useful overview of the alignment activities for the week	4.71	
The group-wide training session was well organized	4.71	
The group-wide training was an effective use of time	4.57	
The hands-on training in my assigned panel helped me better understand the alignment activities	4.57	
The Content Review Tool allowed me to effectively accomplish my tasks during the workshop	4.57	
The whole-group training facilitator was helpful during the workshop	4.57	
Other support staff were helpful during the workshop	4.43	
It was easy to access the items on the Content Review Tool	4.14	

Note: Strongly Disagree = 1 to Strongly Agree = 5

**Table L7. Overall Alignment – All grades**

Answers	Count	Percentage
Strongly aligned	15	50%
Partially aligned	15	50%
Not at all aligned	0	0%

**Table L8. Overall Alignment – Grade 3**

Answers	Count	Percentage
Strongly aligned	2	50%
Partially aligned	2	50%
Not at all aligned	0	0%

**Table L9. Overall Alignment – Grade 4**

Answers	Count	Percentage
Strongly aligned	3	100%
Partially aligned	0	0%
Not at all aligned	0	0%

**Table L10. Overall Alignment – Grade 5**

Answers	Count	Percentage
Strongly aligned	4	80%
Partially aligned	1	20%
Not at all aligned	0	0%

**Table L11. Overall Alignment – Grade 6**

Answers	Count	Percentage
Strongly aligned	1	16.7%
Partially aligned	5	83.3%
Not at all aligned	0	0%

**Table L12. Overall Alignment – Grade 7**

Answers	Count	Percentage
Strongly aligned	1	20%
Partially aligned	4	80%
Not at all aligned	0	0%

**Table L13. Overall Alignment – Grade 8**

Answers	Count	Percentage
Strongly aligned	4	57.1%
Partially aligned	3	42.9%
Not at all aligned	0	0%