2006-2007 Alaska Alternate Assessment

Technical Documentation

June 25, 2007

Dillard Research Associates
jerry@dillardresearchassociates.com

TECHNICAL MANUAL TABLE OF CONTENTS

The Alaska Alternate Assessment provides parents, teachers, and administrators grade level proficiency level information for students with significant disabilities. This document supplements two previously published reports on the test development (content-related evidence (with appendices) and Training (with appendices) Following is a table of contents used to organize the technical documentation supporting this assessment.

DOMAIN-SAMPLING PLANS EXCERPT FROM TECHNICAL MANUAL)

READING

The Alternate Reading Assessment measures a progression of discrete and dynamic behaviors associated with reading acquisition. These discrete skills include: early symbol associations and names, sound-symbol associations, word, sentence, and passage reading, and comprehension. The Alternate Reading Assessment begins with the early skills of image representations of signs and symbols and then moves to the alphabetic principles of letter name, sounds, blends, and decoding. Lastly, the ability to read connected text is measured by sentence and passage reading as well as by comprehension (both listening and reading).

The Alternate Reading Assessment is comprised of tasks designed to measure essentialized reading skills that are pre-requisite to grade level content standards. Specifically, each of the Alternate Reading Tasks is designed to measure the degree to which students with significant disabilities are able to read at the symbol, word, and text levels as well as make interpretations from text. Each task within the Alternate Reading Assessment is anchored to the critical attributes associated with acquiring basic reading as defined by the extended grade level expectations:

Extended Standard Attributes/Strands

- Word Identification Skills: This attribute is conceptualized as the knowledge that print/symbols is spoken words in written form with specific meaning; the recognition that sentences in print are comprised of separate words; and the recognition and naming of all upper and lower case letters This knowledge is demonstrated by applying the above structural analysis skills to read words. Word Identification Skills are operationally defined by the following behaviors:
  - The Identification of pictures;
  - The identification of symbols;
  - The identification of letter names;
  - Orally segment words into single syllable spoken phonemes;
  - The identification of the sound of each letter in the alphabet;
  - Orally blend single syllable spoken phonemes into recognizable words;
  - The identification or reading words of increasing complexity (e.g. ≥5 letters or ≥2 syllables);
  - The reading of simple sentences of 2 to 5 words.

- Forming a General Understanding: This attribute is best conceptualized as comprehension—both listening and reading--and the awareness that comprehension can be facilitated by predicting, questioning, and linking information to personal experiences. This attribute is operationally defined by the following behaviors:
  - The Identification of details from a story read aloud;

- Analysis of General Content and Structure: This attribute centers on reading of connected text and answering questions related to literal comprehension of story grammar elements (character, setting, problem), evaluative comprehension (solution and emotion), and inferential comprehension (prediction). This attribute is operationally defined by the following behaviors:
  - Identification of the main character;
  - Identification of the setting;
  - Provide a description of the conflict/problem;
  - Provide a description of the resolution;
  - Provide a description of character traits: feelings, motivation, physical traits.

- Fluency: Fluency is conceptualized as the automatic application of decoding skills to facilitate the reading of familiar and unfamiliar words automatically and effortlessly. Fluency is operationally defined by the following behavior:
  - Read a passage aloud with rhythm, flow, and expression.

This attribute framework was used to guide the development of the specific test specifications used to build the bank of tasks found in the Alternate Reading Assessments. These tasks utilize a developmental scale of proficiency. The attributes framework helped to operationally define the dynamic skills associated with mastery of basic skills and as such the tasks are a representative sample of the most basic skills in the areas of reading, with each successive task increasing in cognitive demand and level of skill proficiency. For example, letter identification precedes word reading in isolation, then word reading in text (which also varies by readability), and finally reading text and answering questions.

Test and Item Specifications

The Alternate Reading Assessment is comprised of reading tasks that provide sensitive administration and scoring systems to extend below the current statewide tests that rely on multiple-choice or performance task formats. The framework of attributes allowed the development of test specifications with a variety of tasks that eventually were placed into a single construct using a developmental scale of difficulty. The item specifications for the Alternate Reading Assessment are used to develop annual secure versions.

Task 1: Identify Pictures: This is a picture naming exercise. Responses are scored as incorrect, partially correct, or fully correct. It is comprised of eight items in which students are asked to identify common objects. Common objects include those objects most children encounter every day, and are not limited to "functional" sings. Objects are presented in black and white, high contrast resolution.

Task 2: *Identify Symbols*: Common symbols used in signage in the community.

Task 3: *Identify Letter Names*: Responses are scored as incorrect, partially correct, or fully correct. It is comprised of the following items:
4 items: Lower case letter. Font = any "Sans" (without serif).
4 items: Upper case letters. Font = Times is O.K.

Task 4: Comprehend Oral Text: This is a listening comprehension task. Students are read a short story. Students are then asked six specific, guided comprehension questions. Each question taps a specific element of story structure: character, setting, problem, solution, emotion, and prediction. Answers are scored on a three point scale (0, 1, 2), earning points based on the completeness of the answer.

Task 5: Segment Phonemes: This is a phoneme segmentation exercise. Students are asked to identify the sounds in a given word. Responses are scored for number of correct phonemes identified. This task is comprised of eight items:
2 items: one-syllable words comprised of two phonemes.
6 items: one-syllable words comprised of three phonemes.

Task 6: Identify Letter Sounds: All letters and consonant digraphs (n=42) are sampled for the primary sounds represented.

Task 7: Blend Sounds: This is a word reading exercise. Students are asked to identify all of the sounds in a word presented to them on a flashcard. Responses are scored for number of correct sounds produced. This task is comprised of eight items:
1 item: one-syllable phonemically regular word comprised of three sounds.
7 items: one-syllable phonemically regular words comprised of four sounds.

Task 8: Identify Beginning Reading Words: This is a word reading exercise. Students are asked to read eight different words, each individually presented on a flashcard. Responses are scored as incorrect, partially correct, or fully correct. This task is comprised of the following eight items:
1 item: sight word comprised of two letters.
2 items: sight words comprised of three letters.
5 items: sight words comprised of four items. The sight words are selected from the list of 400 Common Sight Words (Direct Instruction Reading [1991] Carnine, Silver & Kame'enui).

Task 9: Identify Advanced Reading Words: A sample of words with various blends, consonant digraphs, vowel digraphs, prefixes and suffixes, and representing various root words and word families are sampled from the Frye Book of Lists.

Task 10: Read Sentences: This is a sentence reading exercise. Students are asked to read sentences individually presented on flashcards. Responses are scored for number of words read correctly. This task is comprised of five items:
1 item: two-word sentences comprised of phonetically regular words, each word not more than four letters long.

1 item: three-word sentences comprised of phonetically regular words and sight words.
1 item: three-word sentences comprised of phonetically regular words and sight words, each word comprised of not more than six letters.
1 item: four-word sentences comprised of phonetically regular words and sight words, each comprised of words not more than six letters.
1 item: five-word sentence comprised of phonetically regular words and sight words, each comprised of not more than six letters.

Task 11: Read Passages: This is a connected test reading exercise. Responses are scored for words read correct per minute. There are three levels of text, each constructed according to the following guidelines. However, each student is initially administered the first passage. Decision rules and student performance dictate which, if any, successive passage(s) is administered. Refer to Appendix F for the Read Passages Decision Guide. Below are the passage construction guidelines:
Story: 150 – 200 word multi-paragraph, Flesh-Kincaid Reading Level < grade 1.2.
Story: 200 – 250 words. Multi-paragraph, Flesch-Kincaid Reading Level of 1.2 to 2.5.
Story: 225 – 300 words. Multi-paragraph, Flesch-Kincaid Reading Level of 2.6 to 3.0.

Flesch-Kincaid is a formula to determine difficulty of reading level that takes into consideration the length of words, # of words per sentence, average length of word, number of syllables and the frequency and complexity of sentences.

Task 12: Comprehend Printed Text: This is a reading comprehension task. Students are asked to read two different passages (Level 1 and Level 2) Students restate the main ideas of a story that they have read. The retell is scored for number of important items mentioned. There are three levels of retell tasks.

WRITING

The Alternate Writing Assessment samples a progression of emerging writing skills. Mastery of these initial skills leads to proficiency in writing. Specifically, each task within the Alternate Writing Assessment is designed to document student skill acquisition in composing written units of increased complexity (letters, words, sentences, paragraphs) to communicate meaning. These skills are the foundational skills necessary to meet Extended Grade Level Expectations (EGLEs). Initial tasks use simple symbols of letters and structured tasks of copying. More complex tasks include the generation of words, sentences, and stories. These performance tasks are scored to reflect varying degrees of complexity and skill development. Each task within the Alternate Writing Assessment is anchored to the critical attributes associated with written language. Three attributes were used to define the Writing domain and identify specific, discrete skills that reflect mastery of each attribute:

Extended Standard Attributes/Strands

- Writes using a variety of forms: This attribute is conceptualized as the behaviors needed to facilitate the ability to communicate ideas through the written language. The following discrete behaviors operationally define this behavior:
  - Identify Letter;
  - Copy Letters;
  - Identify Words;
  - Copy Words;
  - Identify Sentences;
  - Copy Sentences;
  - Write Own name;
  - Write words from dictation;
  - Write sentences from dictation;
  - Write a sentence in response to a given prompt;
  - Write a story about a series of pictorial prompts;
  - Write a story in response to a provided prompt.

- Structures and conventions of writing: This attribute is conceptualized as the ability to apply the basic rules of punctuation and capitalization, grammar and the use a variety of sentence structures when communicating ideas. The following behaviors operationally define this attribute:
  - Identify errors in punctuation;
  - Identify correct end mark punctuation;
  - Identify correct word tense;
  - Identify correct word use within a provided sentence.

- Revisions
  - Write a story with pictures
  - Write a story without pictures

This attribute framework was used to guide the development of the specific test specifications used to build the bank of tasks found in the Alternate Writing Assessments. These tasks reflect a representative sample of the most basic skills in the areas of written language, with successive tasks increasing in cognitive demand and level of skill to achieve proficiency.

Test and Item Specifications

The Alternate Writing Assessment is comprised of 10 writing tasks with sensitive administration and scoring systems that extend below the current statewide tests that rely on multiple-choice or performance task formats. The item specifications for the Alternate Writing Assessment are presented below and are used to develop secure versions of the Alternate Writing Assessment. A total of 10 tasks are used in the assessment, each successive task is a more difficult, complex skill underlying written language development. Each task is presented within a standardized format.

Task 1: Copy Letters: Responses are scored as letters formed fully, partially, or not at all correct. It is comprised of ten items:

5 items: Copying of individual lower case letters.

5 items: Copying of individual uppercase letters.

General representation of letter groupings:

Big Bellies- a, c, d, o, g, q (start with circle shape)

Tall Guys- b, f, l, h, k, t (start at top line)

Sinkers- j, p, y, (g, q) (go below the line)

Short Sticks- n, m, i, r, u, v, w (start at mid line)

Zigs and Zags- e, s, z, x (change in direction, diagonal lines)

Task 2: Copy Words: Responses are scored as letters formed and spelling of words as fully, partially, or not at all correct. A general representation of letters in the alphabet is used. Letters are spread out and which letters are used is noted. It is comprised of eight items:

8 items: Copying of words comprised of three graphemes (visual representation of sound) presented in lower case lettering.

Task 3: Copy Sentences: Responses are scored for number of correct letter sequences, and on a scale of 0-2 for letter placement, letter form, and spacing. It is comprised of the following exercises:

1 item: Copying a three-word sentence comprised of common words, and is a mix between phonemically regular words and sight words each comprised of not more than 6 letters.

2 items: Copying a four-word sentence comprised of common words, and is a mix between phonemically regular words and sight words, each comprised of not more than 6 letters. Punctuation should include one possessive, and one exclamation.

1 item: Copying a five-word sentence comprised of common words, and is a mix between phonemically regular words and sight words, each comprised of not more than 6 letters. Punctuation should include a question mark.

Task 4: *Write Your Own Names:* Students write their first and last names. Responses are scored for number of correct letter sequences written. It is comprised of the following exercise:

1 item: Student writes first name.

1 item: Student writes last name.

Task 5: *Write Words from Dictation:* Responses are scored for fully, partially, or not at all correct. It is comprised of the following exercises:

3 items: One-syllable, phonemically regular words, each comprised of two letters.

3 items: One-syllable, phonemically regular words, each comprised of three letters.

2 items: One-syllable, phonemically irregular words, each comprised of three letters.

2 items: One-syllable, phonetically irregular words, each comprised of four letters.

1 item: One-syllable, consonant-vowel-consonant-silent e combination.

Task 6: Writing Sentences from Dictation: Responses are scored for number of correct letter sequences written. It is comprised of the following exercises:

1 item: Sentence totaling 14-17 correct letter sequences, with punctuation only at the end of the sentence, comprised of words no longer than 6 letters long.

1 item: Sentence totaling 24-27 correct letter sequences, with one instance of within-sentence punctuation, comprised of words no longer than 6 letters long.

Task 7: Sentence Mechanics: A series of sentences are sampled that represent various grammatical, syntactical, and semantic constructions in grade level materials.

Task 8: Write a Sentence: This task requires students to produce a sentence in response to a verbal prompt. Responses are scored for number of Correct Letter Sequences written. It is comprised of one item scored on a continuous scale.

Students write a sentence in response to a prompt. The student is given two choices to write about.

Task 9: Write a Story with Pictures: Three pictures are presented with a setting scene in the first picture, an action scene in the second picture, and an outcome scene in the third picture. Above the pictures are words that depict all critical objects displayed in the scenes.

Task 10: Write a Story without Pictures: The student is the primary source for this task with a general topical prompt used to frame the content (for example, write a story about school…).

MATHEMATICS

The Alternate Mathematics Assessment samples a progression of emerging mathematics concepts and skills necessary for successful completion of computation and application problem solving with numbers and algorithms that govern their use. Concepts measured include: differences, numerical value and order, money, time, size, and various mathematical algorithms. Skills are expressed through solution of problems requiring expression of numbers through dictation and copying, determining quantities and units, identifying dimensions (objects, values, units), and solving problems with addition, subtraction, and multiplication.

The Alternate Mathematics Assessment is comprised of 22 tasks designed to measure emerging student performance preparatory to meeting the Alternate grade level expectations (ExGLEs). Each of the Alternate Mathematics Tasks measures the degree to which students with significant disabilities have developed numerical understanding. Each task within the Alternate Mathematics Assessment is anchored to critical attributes associated with the following Extended Grade Level expectations (ExGLEs): Numeration, Measurement, Estimation and Computation, Functions and Relations, Geometry, Statistics and Probability, and Communication.

Alternate Standard Strands/Attributes

- Numeration –This domain is conceptualized as the understanding that numbers are unique symbols, the different ways to represent numbers, and the relationship among numbers and number systems as defined by the following behaviors:
  - Identifying numerals;
  - Copying numerals;
  - Writing numerals;
  - Identifying numbers on a number line;
  - Ordering numbers;
  - Counting on dictation;
  - Identifying the place value of a number within a larger number;
  - Identifying fractions.

- Measurement: This domain focuses on identifying and using units and metrics for scaling and making comparisons and is defined by the following behaviors:
  - Identifying items by their size;
  - Identifying the time of day;
  - Telling time;
  - Counting money.

- Estimation and Computation: Estimation and Computation is conceptualized as the utilization of the processes, conceptual understanding and skills needed to perform basic arithmetic functions, make reasonable estimates, and select and use appropriate methods/tools to perform these functions. This attribute is operationally defined by the following behaviors:
  - Counting items;
  - Adding and subtracting numbers under timed and un-timed conditions.
  - Identifying fractions

- Functions and Relationships: This domain addresses mathematical patters and symbols and is expressed in the following behaviors:
  - Discriminating Differences between Shapes and Numbers;
  - Using mathematical symbols
  - Identifying the largest group;
  - Identifying the smallest group;
  - Identifying groups based on the words: less than, more than, equal to.

- Geometry: Identifying basic shapes and with three and four sides and includes the following tasks:
  - Identifying shapes

- Statistics and Probability: This domain addresses fractions and probabilities, including display and interpretation of (likely) outcomes.

- Communication: This domain address the use of mathematical skills and concepts in everyday situations and employs the following types of behaviors:
  - Identifying specific dates and days in a month;
  - Identifying coins and bills.

This framework was used to guide the development of the specific test specifications used to build the bank of tasks found in the Alternate Mathematics Assessments. These tasks utilize a developmental scale of proficiency. The attributes framework helped to operationally define the skills with each successive task generally increasing in cognitive demand and level of skill proficiency.

Test and Item Specifications

The Alternate Mathematics Assessment is comprised of Mathematics tasks with sensitive administration and scoring systems that extend below the current statewide tests that rely on multiple-choice or performance task formats. The Alternate Mathematics Assessment is designed to evaluate whether or not students with significant disabilities are developing numerical understanding at the most basic levels. The attributes allowed the development of test specifications with a variety of tasks and items that eventually were placed into a single construct using a developmental scale of proficiency. The item specifications for the Alternate Mathematics Assessment are presented below.

Task 1: Identify Numbers: This task is a numeral naming exercise. Responses are scored as correct or incorrect. It is comprised of the following exercises:
8 Items: Naming one-digit (0-9) numerals that are presented randomly on cards.

Task 2: *Copy Numbers:* This is task is a copying exercise. Responses are scored as digits fully, digits partially or digits not at all correct. All numerals are represented and attempts have been made to distribute them evenly. This task is comprised of the following items:
5 items: Copying single digit numbers.
2 items: Copying two-digit numbers.
1 item: Copying three-digit numbers.

Task 3: *Identify Shapes*: This is a shape naming exercise. Responses are scored as correct or incorrect. Items are chosen from the following two-dimensional shapes: square, triangle, circle, oval, rectangle, trapezoid, pentagon, hexagon, or octagon. Diamond or rhombus can be used but are allowed to have interchangeable responses. This task is comprised of eight items:
5 items from k-1 Foundations: Naming two-dimensional shapes, circle, triangle, rectangle, square, and oval.
3 items from grade 2 Foundations: Naming two-dimensional shapes rhombus, trapezoid, diamond, parallelogram, hexagon, pentagon, or octagon.

Task 4:Write Numerals: This task is a dictation exercise. All ten digits (0-9) must be included at least once. Responses are scored for fully correct, partially correct, or not at all correct. It is comprised of the following exercises:
4 items: One-digit numbers.
6 items: Two-digit numbers.

Task 5: Manipulate Mathematics Concepts-Discriminate Differences: This task requires students to discriminate differences between size, shape or design. Responses are scored as correct or incorrect. This task is comprised of four items:
2 items: Student identifies one object, out of a set of no more than four, which is different.
2 items: Student identifies two objects, out of a set of no more than four, which are the same.

Task 6: Measurement-Size: This task requires students to be aware of size differences and comparisons such as larger than, smaller than, longer than, shorter than, equal to, and equivalent sizes between objects. Responses are either correct or incorrect. This task is comprised of four items:
3 items: Student identifies one picture out of a set of no more than four where two are the same size, which is larger, smaller, longer, shorter, or equal.
1 item: Student estimates equivalent quantities of one object to measure the height of another object.

Task 7: Identify Money: Using real coins and bills is preferred. Note: none of the eligible content in the standards addresses $20 bills. Bills in general are not addressed until 3rd grade instruction. However, bills are important for students following the Life Skills curriculum. The option is given as a choice in the last item. This task has two parts. The first is a money identifying exercise. The second is a money value task. In both, responses are scored as correct or incorrect.

Task 8: Measurement-Time of Day: This task requires students to understand the concepts such as morning, afternoon, evening, day, and night. Responses are scored for each correct identification of the flashcard that represents the teacher's verbal prompt. This task consists of three items: student identifies the appropriate picture out of the five possible scenarios that describe the time of day.

Task 9: Number Line: This task relates to the student's understanding of a number line. Responses are scored either correct or not. It is comprised of the following exercises:

a. 1 item: Saying or pointing to the missing number on a number line with 11 out of 12 numbers indicated, ranging between 0 and 19.
b. 1 item: Saying or pointing to the number that comes BEFORE (NOT the missing number).
c. 1 item: Saying or pointing to the number that comes AFTER (NOT the missing number).
d. 1 item: Saying or pointing to the number that comes FIRST.
e. 1 item: Saying or pointing to the number that comes LAST.

Task 10: This task requires students to identify the time presented on a clock face to the nearest hour, half-hour or minute. Responses are scored for each correct identification of hour and minutes. This task is comprised of four items:

2 items: Time presented on analog clock. Any hour with at least one item on the hour or half-hour, with minutes presented in no less than 5-minute increments.

2 Items: Time presented on digital clock. Any hour with at least one item on the hour or half-hour, with minutes presented in no less than in 1-minute increments.

Task 11 Order Numbers: This task is an ordering exercise. Responses are scored as correct or incorrect. Patterns are avoided in the foil. It is okay to skip numbers and no more than two numbers are to be sequential. This task is comprised of one item:

1 item: Four single digit numbers are presented in random order. Students reconfigure the numbers in ascending order.

Task 12 Manipulate Mathematics Concepts – Ordinal Numbers: This task requires students to exhibit knowledge of vocabulary regarding ordinal position. Responses are scored as correct or incorrect. This task is comprised of four items:

2 items: Student responds to question regarding ordinal position by identifying object in appropriate place (up to tenth place) OR

2 items: Student responds to question regarding "before" and "after" positions.

Task 13 Calendar: This task requires students to be able to read and interpret a calendar. Responses are scored as either correct or not. The calendar has a traditional "Sunday first" format. It is comprised of five items:

1 item: Saying the missing date on the calendar.
1 item: Pointing to the day BEFORE a certain date. The certain date must not be a Sunday.
1 item: Pointing to the day AFTER a certain date. The certain date must not be a Saturday.
1 item: Saying the name of the day of the week when indicating the date.
1 item: Counting the number of days in the week in a certain month (e.g., How many Tuesdays are there in this month?).

Task 14 Manipulate Mathematics Concepts – Counting/Take Away: This is a counting task that requires students to count and provide a number that represents a group of objects. Responses are scored as correct or incorrect. The task is comprised of seven items:

4 items: Student counts objects and provide the correct numeral for the summative amount. All summative amounts are less than 10.
2 items: Student compiles a set of objects to represent a number presented orally.
1 item: Student identifies the missing subtrahend.

Task 15 Manipulate Mathematics Concepts – Quantity: This task requires students to exhibit knowledge of vocabulary related to quantities (most, same, less than, more than). Responses are scored as correct or incorrect. This task is comprised of four items: Student identifies group of objects in response to a question requiring determination of quantities.

Task 16: Manipulate Mathematics Concepts – Fractions: This task requires students to identify fractional parts of wholes and sets. Responses are scored as correct or incorrect. This task is comprised of five items:

2 items: Student identifies the number of objects that represent stated fraction (using wholes, halves, and fourths).
1 item: Student identifies the number of objects that represent a stated fraction (using thirds, eighths, or tenths only).
2 items: student states the fraction of objects represented (using halves, fourths, and thirds only).

Task 17 Count Money: This task requires calculation of multiple coin values. Responses are scored for partially, fully, or not at all correct. It is comprised of five items:
1 item: Adding up to five coins of not more than two values, sum is less than twenty cents.
3 items: Adding up to six coins of not more than three values, sum is less than one dollar.
1 item: Identifying the correct unit of cents within any calculation of coin value tasks.

Task 18: Manipulate Mathematics Concepts – Place Value: This task requires students to exhibit knowledge of place value. Responses are scored as correct or incorrect. This task is comprised of four items:
1 item: Student identifies the digit that occupies a specific place value in a two-digit number.
2 items: Student tells the number of 1's 10's or 100's represented in a 3-digit number.
1 item: Student tells the number represented by a picture of a mix of Base-10 blocks ("Cuisenaire" type) with 1's, 10's and 100's.

Task 19: Count on Dictation: This task is an oral counting task. Responses are scored for correct number sequences produced over the total correct number sequences possible. It is comprised of the following items:
1 item: Student recites the string of consecutive numbers when provided with a prompt that requires them to continue a number sequence, with increments of 1, including numbers in the next place value.
1 item: Student recites the string of consecutive numbers when provided with a prompt that requires them to continue a number sequence, with increments of 2 or 5.
1 item: Student recites the string of consecutive numbers when prompted with a prompt that requires them to continue a number sequence, with increments of 10 or 100.

Task 20: Timed Computation – Addition Facts: This task requires calculation. It also tests how many calculations can be made within one minute (fluency sums). The responses are scored for digits correct. All problems are written vertically. They are comprised of the following items:
8 items: Addition of two one-digit numbers, sum is less than or equal to 10.
7 items: Addition of two one-digit numbers, sum is greater than or equal to 10.
5 items: Addition of a one digit number with a two digit number that is less than twenty, sum less than 20.

Task 21: Timed Computation – Subtraction Facts: This task requires calculation. It also tests how many calculations can be made within one minute (fluency sums). The responses are scored for digits correct. All problems are written vertically. They are comprised of the following items:
15 items: One-digit number subtracted from one-digit number, difference is less than 10.
5 items: One-digit number subtracted from two-digit number, difference is less than 10.

Task 22: Mixed Computation – Addition and Subtraction: This task requires calculation. The responses are scored for digits correct. Numerical problems are vertically formed. It is comprised of the following items:
3 items: Addition of two one-digit numerals, sum less than 10.
2 items: Addition of two one-digit numerals, sum is greater than 10.
1 item: Addition of a one-digit numeral and a two-digit numeral, carrying from 1's to 10's.
3 items: Subtraction of two one-digit numerals, difference is less than 10.
2 items: Subtraction of a one-digit numeral from a two-digit numeral, no borrowing required.
1 item: Subtraction of a two-digit numeral from a two-digit numeral, no borrowing required.

ADMINISTRATION TYPE SUMMARY: NT, NA-I, ELOS, AND STD

Although the decision to administer the assessment is global at the overall academic skill level, teachers still need to make decisions about which individual tasks to administer. The decision to give the Alternate Writing Assessment does not imply that all tasks are given. Rather, some may be given and others not given based on two important principles. It is important to actively make this decision rather than leave the task blank as the data analysis routine encodes a "0" when either of these two options IS NOT invoked.

Not Administered-Inappropriate: Teachers may decide that the task is inappropriate for a variety of reasons: Physical or sensory impairments may preclude administration of a task; interfering behaviors keep the student from effectively participating, or the behavioral repertoire of the student is extremely restricted for the tasks being considered. In general, not giving a task because it is inappropriate leads to a blank score for those tasks and the student is not scored. Later tasks are scored as taken and the students overall rating is adjusted accordingly.

How Should Students Take the Assessment?

Teachers may decide to have students take the test (or individual subtasks) with changes that help the student perform and the score to be more meaningful. The following two types of changes are very different from each other in that the first one **(accommodations)** results in students' scores being aggregated and the overall rating made while the second change **(expanded levels of support)** results in the changes being made with a "0" being scored for the task. The distinction between accommodated and expanded levels of support is a function of three aspects of the administration. First, the domain sampled within any single task needs to reflect random selection. This allows others to make inferences about a student's skill in general. Second, the task needs to be neutral to the topography of the behavior, or the form in which students interact with the specific tasks. Third, the administration and scoring systems need to be relatively standardized so that small changes are ignored if they do not threaten the meaning of what is being measured. As a consequence, students can participate in the statewide testing program in two ways: (a) accommodated, or (b) expanded levels of support.

Expanded Levels of Support (ELOS): This type of change generally refers to content and often is done to allow the teacher to capture at least some behavior in the domain of the task but not to make serious inferences about general performance in that domain. The most common modifications are part of physical assistance or enhancing correctly guessing.

> *Facilitating correct responses* – When cards are used for a selection response (pointing or nodding), the full set of cards needs to be deployed as any subset would increase a correct response by guessing alone.
> *Physical guidance* – When students are given any kind of physical guidance (e.g., moving the hand in writing a number or using a tracing line), this change would be considered a modification.

Accommodated: When changes are made in the manner for giving or taking the assessment AND these changes are in format only, they are considered accommodations that do not change the construct being measured. Many changes are possible and only a few of the more general types are listed below.

> *Flashcards* – All tasks that have flashcards can be changed so that the student does not have to make a constructed response (vocalize or print). For example the student can point to (nod an affirmative, vocalize agreement, etc.) when the correctly dictated letter or number is made while the administrator is saying it.
> *Computerized response* – All tasks that require a written constructed behavior (usually requiring a pencil) can be exhibited with a computer or any other kind of production assist.

Following are the results from all types of administration: (a) NT, Not tested, (b) NA-I, Not – Inappropriate, (b) ELOS, Extended Level of Support, and (d) STD, Standard with or without accommodations. It is important to note that the high level of NT is that tasks were aligned with grade level expectations and Adequate Yearly Progress was established only with grade-aligned tasks.

Table 1. Reading Administration and Test Participation Frequencies

Reading Tasks	**NT**	**NA-I**	**ELOS**	**STD**		**NT**	**NA-I**	**ELOS**	**STD**
1. Identify Pictures/Representations of Objects	5	0	64	387		1%	14%	0%	85%
2. Identify Signs and Symbols	4	0	69	383		1%	15%	0%	84%
3. Identify Letter Names	8	0	64	381		2%	14%	0%	84%
4. Comprehend Oral Text (Listening)	49	0	35	352		11%	8%	0%	81%
5. Segment Phonemes	55	0	19	332		14%	5%	0%	82%
6. Identify Letter Sounds	60	0	27	341		14%	6%	0%	80%
7. Blend Sounds	113	0	12	302		26%	3%	0%	71%
8. Identify Beginning Reading Words	121	0	19	310		27%	4%	0%	69%
9. Identify Advanced Reading Words	125	0	17	304		28%	4%	0%	68%
10. Read Sentences	125	0	12	307		28%	3%	0%	69%
11. Read Passages – SP	190	0	1	249		43%	0%	0%	57%
11. Read Passages – L1	229	0	6	204		52%	1%	0%	46%
11. Read Passages – L2	296	0	2	135		68%	0%	0%	31%
12. Comprehend Printed Text	216	0	8	213		49%	2%	0%	49%

Table 2. Writing Administration and Test Participation Frequencies

Writing Tasks	**NT**	**NA-I**	**ELOS**	**STD**		**NT**	**NA-I**	**ELOS**	**STD**
1. Identify/Copy Letters	5	10	76	349		1%	17%	2%	79%
2. Identify/Copy Words	8	10	78	345		2%	18%	2%	78%
3. Identify/Copy Sentences	10	19	69	343		2%	16%	4%	78%
4. Write Your Own Name	28	9	62	343		6%	14%	2%	78%
5. Write Words from Dictation	82	26	27	308		19%	6%	6%	70%
6. Write Sentences from Dictation	92	31	21	299		21%	5%	7%	67%
7. Sentence Mechanics	97	27	11	307		22%	2%	6%	69%
8. Write a Sentence	119	30	16	277		27%	4%	7%	63%
9. Write a Story with Pictures	195	22	6	219		44%	1%	5%	50%
10. Write a Story Without Pictures	214	23	5	201		48%	1%	5%	45%

Table 3. Mathematics Administration and Test Participation Frequencies

Mathematics Tasks	**NT**	**NA-I**	**ELOS**	**STD**		**NT**	**NA-I**	**ELOS**	**STD**
1. Identify Numbers	3	3	67	384		1%	1%	15%	84%
2. Identify/Copy Numerals	16	11	72	358		4%	2%	16%	78%
3. Identify Shapes	9	3	65	382		2%	1%	14%	83%
4. Write Numerals	54	21	39	344		12%	5%	9%	75%
5. Discriminate Differences	55	11	23	370		12%	2%	5%	81%
6. Measurement (Size)	62	13	22	362		14%	3%	5%	79%
7. Identify Money	71	11	21	355		16%	2%	5%	78%
8. Time of Day	81	8	22	348		18%	2%	5%	76%
9. Number Line	85	16	19	338		19%	3%	4%	74%
10. Tell Time	142	11	11	295		31%	2%	2%	64%
11. Order Numbers	145	11	8	295		32%	2%	2%	64%
12. Ordinal Numbers	147	10	8	294		32%	2%	2%	64%
13. Calendar	149	13	6	291		32%	3%	1%	63%
14. Take Away	155	9	4	290		34%	2%	1%	63%
15. Quantity	156	11	4	287		34%	2%	1%	63%
16. Fractions	186	13	5	255		41%	3%	1%	56%
17. Count Money	188	14	5	252		41%	3%	1%	55%
18. Place Value	192	14	4	249		42%	3%	1%	54%
19. Count on Dictation	197	15	4	243		43%	3%	1%	53%
20. Addition Facts	201	14	7	237		44%	3%	2%	52%
21. Subtraction Facts	200	15	8	236		44%	3%	2%	51%
22. Addition and Subtraction	207	15	9	228		45%	3%	2%	50%

For students who took the test using ELOS on any task, teachers also completed a survey on the student's receptive and expressive communication skills. Following are the results of this survey. Some highlights of this survey indicate that, of the approximately 170 students for whom teachers completed the survey, a large percentage…

- using *receptive communication*, DO NOT read written words (56%) or understand manual signs (50%), complex gestures (70%), (43%) or objects as representations (44%).

- using *expressive communication*, DO NOT communicate in sentences (62%), phrases-word combinations (51%), or single words (39%); DO NOT use simple gestures (33%) or complex gestures (77%) or use non-speech symbols (59%); DO NOT point to pictures or drawings (33%), symbols-textures, shapes (58%) or use communication boards (57%). Only 1 in 3 can hold an object.

Table 4. Frequency Data on Receptive Skills for Students Participating in Extended Levels of Support (ELOS)

Receptive Communication	Doesn't Do	Does Somewhat	Does Routinely	Total
Hear spoken words/language	9	41	119	169
Read written words	97	49	27	173
Read Braille or read lips	173	0	0	173
Understand manual signs	86	75	11	172
Understand complex gesture	120	38	14	172
Understand simple gestures	22	90	60	172
Understand pictures used to represent objects, events	30	68	75	173
Understand line drawings used to represent objects, events	64	63	45	172
Understand symbols (textures, shapes) to represent objects, events	74	68	29	171
Understand miniature objects used to represent objects, events	76	61	35	172
Other Receptive Mode of communication used by student	58	24	21	103

Table 5. Percentage Data on Receptive Skills for Students Participating in Extended Levels of Support (ELOS)

Receptive Communication	Doesn't Do	Does Somewhat	Does Routinely
Hear spoken words/language	5%	24%	70%
Read written words	56%	28%	16%
Read Braille or read lips	100%	0%	0%
Understand manual signs	50%	44%	6%
Understand complex gesture	70%	22%	8%
Understand simple gestures	13%	52%	35%
Understand pictures used to represent objects, events	17%	39%	43%
Understand line drawings used to represent objects, events	37%	37%	26%
Understand symbols (textures, shapes) to represent objects, events	43%	40%	17%
Understand miniature objects used to represent objects, events	44%	35%	20%
Other Receptive Mode of communication used by student	56%	23%	20%

Table 6. Frequency Data on Expressive Skills for Students Participating in Extended Levels of Support (ELOS)

Expressive Communication	Doesn't Do	Does Somewhat	Does Routinely	Total
Communicates in sentences	108	40	25	173
Communicates in phrases or word combinations	88	28	57	173
Communicates using single words	67	48	58	173
Verbalizes/Vocalizes	16	66	90	172
Signs (sign language)	117	49	6	172
Gestures (simple)	57	83	32	172
Gestures (complex)	133	30	9	172
Operates electronic system	108	47	18	173
Uses non-speech symbols	101	55	16	172
Indicates (points to) objects	40	71	63	174
Indicates (points to) pictures/drawings	56	62	54	172
Uses communication board	98	54	21	173
Communicates using miniature objects	144	25	5	174
Indicates (points to) symbols, textures, shapes	101	52	22	175
Touches person	31	75	69	175
Touches object	36	72	62	170
Manipulates person (physical)	68	71	31	170
Extends (holds up) object	56	76	38	170
Other expressive mode of communication	24	23	29	76

Table 7. Frequency Data on Expressive Skills for Students Participating in Extended Levels of Support (ELOS)

Expressive Communication	Doesn't Do	Does Somewhat	Does Routinely
Communicates in sentences	62%	23%	14%
Communicates in phrases or word combinations	51%	16%	33%
Communicates using single words	39%	28%	34%
Verbalizes/Vocalizes	9%	38%	52%
Signs (sign language)	68%	28%	3%
Gestures (simple)	33%	48%	19%
Gestures (complex)	77%	17%	5%
Operates electronic system	62%	27%	10%
Uses non-speech symbols	59%	32%	9%
Indicates (points to) objects	23%	41%	36%
Indicates (points to) pictures/drawings	33%	36%	31%

Expressive Communication – Continued	Doesn't Do	Does Somewhat	Does Routinely
Uses communication board	57%	31%	12%
Communicates using miniature objects	83%	14%	3%
Indicates (points to) symbols, textures, shapes	58%	30%	13%
Touches person	18%	43%	39%
Touches object	21%	42%	36%
Manipulates person (physical)	40%	42%	18%
Extends (holds up) object	33%	45%	22%
Other expressive mode of communication	32%	30%	38%

Currently, an analysis is being completed on the performance of this population (those for whom ELOS was used).

ITEM LEVEL FUNCTIONING (INTERNAL CONSISTENCY RELIABILITY)

The first analysis of the Alternate assessment focuses on the item level functioning within each of the tasks (providing internal consistency reliability). The second analysis focuses on the task functioning (providing descriptive statistics). In both analyses, data are reported by grade bands. Furthermore, the entire set of tasks is analyzed, even though not all tasks are used in the grade band for establishing AYP.

READING

Table 8. Reading Grade Band 3/4

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 - Identify Pictures/Representations of Objects	.788	.789	8
2 – Identify Signs and Symbols	.711	.712	8
3 – Identify Letter Names	.885	.883	8
4 - Comprehend Oral Text (Listening)	.865	.867	6
5 – Segment Phonemes	.956	.958	8
6 – Identify Letter Sounds	.898	.901	8
7 – Blend Sounds	.951	.955	8
8 – Identify Beginning Reading Words	.879	.878	8
9 – Identify Advanced Reading Words	.918	.919	8
10 – Read Sentences	.898	.916	5
12E – Comprehend Printed Text Level 1	.844	.847	6
12D – Comprehend Printed Text Level 2	.902	.902	6

Table 9. Reading Grade Band 5/6

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 - Identify Pictures/Representations of Objects	.672	.779	8
2 – Identify Signs and Symbols	.712	.717	8
3 – Identify Letter Names	.908	.910	8
4 - Comprehend Oral Text (Listening)	.852	.850	6
5 – Segment Phonemes	.943	.945	8
6 – Identify Letter Sounds	.838	.846	8
7 – Blend Sounds	.938	.941	8
8 – Identify Beginning Reading Words	.810	.802	8
9 – Identify Advanced Reading Words	.892	.892	8
10 – Read Sentences	.903	.918	5
12E – Comprehend Printed Text Level 1	.806	.807	6
12D – Comprehend Printed Text Level 2	.884	.884	6

Table 10. Reading Grade Band 7/8

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 - Identify Pictures/Representations of Objects	.895	.904	8
2 – Identify Signs and Symbols	.812	.817	8
3 – Identify Letter Names	.899	.901	8
4 - Comprehend Oral Text (Listening)	.918	.919	6
5 – Segment Phonemes	.968	.970	8
6 – Identify Letter Sounds	.899	.906	8
7 – Blend Sounds	.965	.967	8
8 – Identify Beginning Reading Words	.923	.923	8
9 – Identify Advanced Reading Words	.875	.876	8
10 – Read Sentences	.895	.916	5
12E – Comprehend Printed Text Level 1	.869	.870	6
12D – Comprehend Printed Text Level 2	.914	.915	6

Table 11. Reading Grade Band 9/10

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 - Identify Pictures/Representations of Objects	.799	.863	8
2 – Identify Signs and Symbols	.696	.684	8
3 – Identify Letter Names	.874	.873	8
4 - Comprehend Oral Text (Listening)	.862	.862	6
5 – Segment Phonemes	.945	.947	8
6 – Identify Letter Sounds	.861	.868	8
7 – Blend Sounds	.922	.924	8
8 – Identify Beginning Reading Words	.931	.933	8
9 – Identify Advanced Reading Words	.917	.917	8
10 – Read Sentences	.893	.910	5
12E – Comprehend Printed Text Level 1	.806	.815	6
12D – Comprehend Printed Text Level 2	.877	.884	6

Summary

All internal consistency reliability coefficients are high to very high, with only two tasks reflecting high-moderate levels (Identify Pictures/Representations of Objects and Identify Signs and Symbols).

WRITING

Table 12. Writing Grade Band 3/4

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify/Copy Letters	.944	.946	10
2 – Identify/Copy Words	.949	.949	8
3 – Identify/Copy Sentences	.935	.975	3
5 – Write Words from Dictation	.946	.948	10
6 – Write Sentences from Dictation	.916	.959	3
7 – Sentence Mechanics	.621	.627	8

Table 13. Writing Grade Band 5/6

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify/Copy Letters	.929	.930	10
2 – Identify/Copy Words	.945	.945	8
3 – Identify/Copy Sentences	.908	.963	3
5 – Write Words from Dictation	.934	.938	10
6 – Write Sentences from Dictation	.924	.968	3
7 – Sentence Mechanics	.601	.610	8

Table 14. Writing Grade Band 7/8

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify/Copy Letters	.939	.941	10
2 – Identify/Copy Words	.975	.975	8
3 – Identify/Copy Sentences	.907	.959	3
5 – Write Words from Dictation	.935	.942	10
6 – Write Sentences from Dictation	.909	.957	3
7 – Sentence Mechanics	.689	.696	8

Table 15. Writing Grade Band 9/10

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify/Copy Letters	.932	.935	10
2 – Identify/Copy Words	.941	.944	8
3 – Identify/Copy Sentences	.907	.963	3
5 – Write Words from Dictation	.944	.949	10
6 – Write Sentences from Dictation	.905	.954	3
7 – Sentence Mechanics	.661	.678	8

Summary

All asks reflect very high internal consistency coefficients with only Sentence mechanics reflecting moderate levels.

MATHEMATICS

Although 22 tasks appear in the Alternate mathematics assessment, the last three do not have item information, thereby precluding any analysis at this level.

Table 16. Mathematics Grade Band 3/4

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify Numbers	.899	.901	8
2 – Identify/Copy Numerals	.887	.930	8
3 – Identify Shapes	.780	.777	8
4 – Write Numerals	.908	.942	10
5 – Manipulate Mathematical Concepts- Discriminate Differences	.815	.815	4
6 – Manipulate Mathematical Concepts- Measurement (size)	.813	.812	4
7 – Identify Money	.853	.873	9
8 – Measurement (Time of Day)	.729	.729	3
9 – Number Line	.819	.819	5
10 – Tell Time	.784	.783	4
12 – Manipulate Mathematical Concepts- Ordinal Numbers	.689	.689	2
13 – Calendar	.776	.774	5
14 – Manipulate Mathematical Concepts – Take Away	.804	.817	7
15 – Manipulate Mathematical Concepts – Quantity	.733	.724	4
16 – Manipulate Mathematical Concepts – Fractions	.545	.573	5
17 – Count Money	.831	.829	5
18 – Manipulate Mathematical Concepts- Place Value	.627	.628	4
19 – Count on Dictation	.709	.792	3

Table 17. Mathematics Grade Band 5/6

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify Numbers	.901	.908	8
2 – Identify/Copy Numerals	.906	.942	8
3 – Identify Shapes	.697	.690	8
4 – Write Numerals	.906	.942	10
5 – Manipulate Mathematical Concepts- Discriminate Differences	.691	.687	4
6 – Manipulate Mathematical Concepts- Measurement (size)	.739	.740	4
7 – Identify Money	.748	.803	9
8 – Measurement (Time of Day)	.746	.745	3
9 – Number Line	.692	.693	5
10 – Tell Time	.780	.781	4
12 – Manipulate Mathematical Concepts- Ordinal Numbers	.629	.629	2
13 – Calendar	.770	.764	5
14 – Manipulate Mathematical Concepts – Take Away	.736	.780	7
15 – Manipulate Mathematical Concepts – Quantity	.767	.763	4
16 – Manipulate Mathematical Concepts – Fractions	.534	.542	5
17 – Count Money	.903	.902	5
18 – Manipulate Mathematical Concepts- Place Value	.572	.574	4
19 – Count on Dictation	.679	.750	3

Table 18. Mathematics Grade Band 7/8

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify Numbers	.440	.870	8
2 – Identify/Copy Numerals	.913	.961	8
3 – Identify Shapes	.768	.787	8
4 – Write Numerals	.904	.936	10
5 – Manipulate Mathematical Concepts- Discriminate Differences	.768	.781	4
6 – Manipulate Mathematical Concepts- Measurement (size)	.780	.783	4
7 – Identify Money	.787	.839	9
8 – Measurement (Time of Day)	.794	.793	3
9 – Number Line	.827	.828	5
10 – Tell Time	.815	.814	4
12 – Manipulate Mathematical Concepts- Ordinal Numbers	.757	.757	2
13 – Calendar	.809	.810	5
14 – Manipulate Mathematical Concepts – Take Away	.583	.590	7
15 – Manipulate Mathematical Concepts – Quantity	.768	.764	4
16 – Manipulate Mathematical Concepts – Fractions	.724	.727	5
17 – Count Money	.876	.877	5
18 – Manipulate Mathematical Concepts- Place Value	.692	.693	4
19 – Count on Dictation	.703	.761	3

Table 19. Mathematics Grade Band 9/10

Task	Cronbach's Alpha	Cronbach's Alpha w Standardized Items	N of Items
1 – Identify Numbers	.298	.822	8
2 – Identify/Copy Numerals	.853	.913	8
3 – Identify Shapes	.687	.670	8
4 – Write Numerals	.898	.929	10
5 – Manipulate Mathematical Concepts- Discriminate Differences	.812	.817	4
6 – Manipulate Mathematical Concepts- Measurement (size)	.821	.826	4
7 – Identify Money	.820	.870	9
8 – Measurement (Time of Day)	.823	.823	3
9 – Number Line	.820	.824	5
10 – Tell Time	.824	.837	4
12 – Manipulate Mathematical Concepts- Ordinal Numbers	.807	.807	2
13 – Calendar	.668	.683	5
14 – Manipulate Mathematical Concepts – Take Away	.642	.634	7
15 – Manipulate Mathematical Concepts – Quantity	.712	.713	4
16 – Manipulate Mathematical Concepts – Fractions	.555	.553	5
17 – Count Money	.876	.877	5
18 – Manipulate Mathematical Concepts- Place Value	.658	.655	4
19 – Count on Dictation	.596	.661	3

Summary

All tasks reflect moderate to moderate-high internal consistency coefficients. In general, the coefficients are lower for those tasks that have fewer items or that are taken by students who are not appropriately learning these skills (e.g. identifying numbers for grade 9-10 students or computing answers to fractions problems for students in the 3-4 grade band).

TASK LEVEL FUNCTIONING (DESCRIPTIVE STATISTICS BY SUBJECT AREA)

In Appendices 1, 2, and 3, tables of descriptive statistics provide information on the functioning of all tasks in reading, writing, and mathematics, respectively. For all grades and then for each grade band, the following information is presented:

Case Processing Summary –N and % of cases that are valid, excluded, and total
Descriptive Statistics –N (count), minimum, maximum, mean, and standard deviation
Reliability Statistics – Cronbach's Alpha, Cronbach's Alpha based on standardized items, and N of items
Item Statistics – mean, std. deviation, and N (of cases)
Inter-Item Correlation Matrix – all items with all items
Summary Item Statistics – mean, minimum, maximum, range, max/min, variance, and N of items
Item-Total Statistics – scale mean if item is deleted, scale variance if item is deleted, corrected item-total correlation, squared multiple correlation, and Cronbach's Alpha if item is deleted
Scale Statistics – mean, variance, std. deviation, and N of items

SUBJECT AREA GRADE BAND STRAND ANALYSIS:

TASKS, POINTS, WEIGHTS, AND PERFORMANCE

Tables 20a-20c. Reading: Grade 3-4

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Word Identification	1-Pictures	16		
	2-Signs	16		
	6-Sounds	8		
	7-Blends	25		
	Total	65	0.8	50
Form Gen Understand	4-Comp	12		
	Total	12	4.2	50
Total		77		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Word Identification	128	.00	65.00	29.8594	20.45611
Valid N (listwise)	128				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Form Gen Understand	128	.00	12.00	2.7500	3.80220
Valid N (listwise)	128				

Tables 21A-21C. Reading: Grade 5-6

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Word Identification	8-Bgn words	16		
	9-Adv Words	16		
	10-Sentences	17		
	Total	49	0.7	33
Form Gen Understand	4-Comp	12		
	Total	12	2.8	33
Analysis of Cnt/Struc	4-Listen Comp (Item 1 used in SS)			
	12-Comp - 2 Passages (E & D) (1 item each used in SS)	4		
	Total	4	8.3	33
Total		65		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Word Identification	123	.00	49.00	24.1138	19.32602
Valid N (listwise)	123				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Form Gen Understand	123	.00	12.00	4.4390	4.22938
Valid N (listwise)	123				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Analysis of Cnt/Struc	48	.00	24.00	13.9583	7.71833
Valid N (listwise)	48				

Tables 22A-22B. Reading: Grade 7-8

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Word Identification	9-Adv Words	16		
	10-Sentences	17		
	11-Rdg Passage (SP)	10		
	Total	43	1.2	50
Form Gen Understand	No ExGLEs			
Analysis of Cnt/Struc	4-Listen Comp	12		
	12-Comp - 2 Passages (E & D)	24		
	Total	36	1.4	50
Total		79		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Word Identification	69	1.00	44.00	28.4638	9.14842
Analysis of Cnt/Struc	55	.00	36.00	23.0182	11.10971
Valid N (listwise)	55				

Tables 23A-23B. Reading: Grade 9-10

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Word Identification				
	9-Adv Words	16		
	10-Sentences	17		
	11-Rdg Passage (SP, L1, L2)	30		
	Total	63	0.8	50
Form Gen Understand	No ExGLEs			
Analysis of Cnt/Struc	4-Listen Comp-2 pass	12		
	12-Comp - 2 Passages (E & D)	24		
	Total	36	1.4	50
Total		99		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Word Identification	69	4.00	66.00	53.1014	13.74253
Analysis of Cnt/Struc	87	.00	36.00	26.6552	8.35523
Valid N (listwise)	65				

Tables 24A-24B. Writing: Grade 3-4 Reading

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Write Using Forms	1-ID Letters	20		
	2-ID Words	32		
	3-ID Sentence	50		
	4-Write Name	10		
	5-Words Dict	39		
	6-Sent. Dict	57		
	8-Write Sent.	10		
	9-Story w Pic.	10		
	10-Story w/o Pic.	10		
	Total	238	0.4	100
Total				100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Write Using Forms	26	149.00	235.00	199.8462	23.38494
Valid N (listwise)	26				

Table 25A-25B. Writing: Grade 5-6

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Write Using Forms	4-Write Own Name	10		
	5-Words Dict	39		
	8-Write Sent.	10		
	9-Story w Pic.	10		
	10-Story w/o Pic.	10		
	Total	79	0.6	50
Structures/Conventions	4-Write Own Name*			
	6-Write Sentences Fr Dic	57		
	7-Sentence Mech	8		
	8-Write Sent.*			
	9-Story w Pic.*			
	10-Story w/o Pic.*			
	Total	65	0.8	50
Total		144		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Write Using Forms	37	31.00	79.00	60.8108	12.49186
Structures/Conventions	84	.00	58.00	27.4286	20.64519
Valid N (listwise)	37				

Tables 26A-26B. Writing: Grade 7-8

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Write Using Forms	6-Write Sent. Fr Dictate	57		
	8-Write Sent.	10		
	9-Story w Pic.	10		
	9-Ideas	5		
	9-Org	5		
	10-Story w/o Pic.	10		
	10-Ideas	5		
	10-Org	5		
	Total	107	0.5	50
Structures/Conventions				
	6-Write Sentences Fr Dic*			
	7-Sentence Mech	8		
	8-Write Sent.*			
	9-Story w Pic.*			
	10-Story w/o Pic.*			
	Total	8	6.25	50
Revises	8-Write Sent.*			
	9-Story w Pic.*			
	10-Story w/o Pic.*			
	Total			
Total		115		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Write Using Forms	49	15.00	100.00	74.1429	18.74944
Structures/Conventions	88	.00	3.00	2.0909	.96652
Valid N (listwise)	49				

Tables 27A-27B. Writing: Grade 9-10

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Write Using Forms	8-Write Sent.	10		
	9-Story w Pic.	10		
	10-Story w/o Pic.	10		
	Total	30	0.6	50
Structures/Conventions	7-Sentence Mech	8		
	Total	8	2.50	50
Revisions	8-Write Sent.*			
	9-Story w Pic.*			
	10-Story w/o Pic.*			
	Total			
Total		99		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Write Using Forms	81	9.00	102.00	72.8272	21.70875
Structures/Conventions	126	.00	3.00	1.9841	.91200
Valid N (listwise)	81				

Tables 28A-28B. Mathematics: Grade 3-4

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Numeration	2-Copy No.	24		0
	9-No. Line	2		0
	Total	26	1.0	25
Measurement	No ExGLE			
Est./Comp.	No ExGLE			
Functions-Rltions	5-Diffs	4		
	6-Msmt	4		
	Total	8	3.1	25
Geometry	3-Shapes	8		0
	5-Diffs*			
	Total	8	3.1	25
Stats/Prob	No ExGLE			
Communication	7-Money	2		
	8-Time/Day	3		
	Total	5	5.0	25
Total		47		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
N34	127	.00	29.00	15.5039	11.84682
Functions-Rltions	127	.00	8.00	3.1890	3.11864
Geometry	127	.00	8.00	3.2283	2.69388
Communication	127	.00	13.00	4.9449	4.93177
Valid N (listwise)	127				

Tables 29A-29B. Mathematics: Grade 5-6

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Numeration	1-ID No.	8		
	2-Copy No.	24		
	4-Wr Num.	44		
	9-No. Line	2		
	12-Ordinal No.	1		
	14-Take Away	1		
	15-Quant	4		
	Total	84	0.2	20
Measurement	6-Msmt	4		
	7-Money	4		
	15-Quant*			
	Total	8	2.5	20
Est/Comp.	20-Add	6		
	22-Mixed	2		
	Total	8	2.5	20
Funcs/Rltns	No ExGLE			
Geometry	3-Shapes	8		
	5-Diffs	4		
	12-Ord*			
	Total	12	1.7	20
Stats/Prob	No ExGLE			
Communication	7-Money	2		
	8-Time/Day	3		
	Total	5	4.0	20
Total		117		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Numeration	122	.00	94.00	55.2049	35.23333
Measurement	122	.00	14.00	7.2049	5.10702
Est/Comp.	46	.00	49.00	27.3261	14.31558
Geometry	122	.00	12.00	5.7459	3.57115
Communication	122	.00	13.00	6.6639	4.75101
Valid N (listwise)	46				

Tables 30A-30B. Mathematics: Grade 7-8

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Numeration	No ExGLEs			
Msmt	7-Money	8		
	17-Count Money	9		
	Total	17	1.2	20
Est/Comp.	20-Addition	16		
	21-Subtraction	5		
	22-Mixed	6		
	Total	27	0.7	20
Funcs/Rltns	20-Addition*	32		
	21-Subtraction*	20		
	22-Mixed*	17		
	Total* (-Est/CompTot)	42	0.5	20
Geometry	3-Shapes	8		
	5-Diffs	4		
	12-Ord No.	2		
	Total	14	1.4	20
Stats/Prob	No Ex GLE			
Communication	7-Money*			
	8-Time/Day	3		
	Total	3	6.7	20
Total		103		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Measurement	128	.00	19.00	8.6797	7.07545
Est/Comp.	62	.00	98.00	50.7903	19.86358
Geometry	128	.00	14.00	6.8281	4.66366
Communication	128	.00	3.00	1.5859	1.36623
Valid N (listwise)	62				

Tables 31A-31B. Mathematics: Grade 9-10

Strand-Attribute	Task Name	Points	Strand Equal	Strand Tot
Numeration	No ExGLEs			
Msmt	7-ID Money	10		
	8-Time/Day	3		
	10-Time	12		
	13-Calendar	5		
	17-Count Money	9		
	Total	39	0.6	25
Est/Comp.	20-Add	16		
	21-Sub	5		
	22-Mix	6		
	Total	27	0.9	25
Function/Relations	20-Add*	16		
	21-Sub*	15		
	22-Mix*	11		
	Total	42	0.6	25
Geometry	No Ex GLEs			
Communication	7-Money	2		
	8-Time/Day*			
	Total	2	12.5	25
Total		110		100

**The value is already counted*

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Measurement	169	.00	39.00	20.0355	14.51145
Est/Comp.	100	.00	69.00	46.7000	19.85885
Communication	169	.00	10.00	6.4320	3.98917
Valid N (listwise)	100				

STANDARD SETTING – SUMMARY OF PROCEDURES

This summarizes the 10 steps used in the process for setting standards in the Alaska Alternate Assessment. The following sections are included: (I) executive summary, (II) description of the steps, (III) extended notes from the final meeting (articulation and extended levels of support), (IV) evaluation from the participants, and (V) appendices of the forms.

Overview

An integrated judgment standard setting process was implemented with the Alaska Alternate Assessment in reading, writing, and mathematics over the course of four days. The evening before the first day, 3 table leaders were oriented in facilitation to ensure the process was completed with integrity. We used 3 tables – one for each subject area – with 8 participants in each table.

On the first day, participants were oriented to the assessment system, how to administer and score the tasks; they also were trained on a 10-step process and completed a first trial judgment in the first grade band (3-4) of their subject area. Participants made cut score judgments to distinguish levels of performance: (a) advanced, (b) proficient, (c) below, and (d) far below, including both individual judgments and consensus on a common cut score for that grade band.

On the second day, each table proceeded in a similar fashion, making individual judgments at each grade band and then reaching consensus as a group. The process of reaching consensus ostensibly involved a rather detailed discussion of the differences in individual judgments and consideration of the population and the specific skills being tested as well as opportunity to learn. At the end of the second day, participants were provided an overview of basic statistical summaries and how to interpret frequency distributions.

On the third day, each table was provided the outcomes of students' performance for each grade band, including how students performed as a group (reflecting grade band average, standard deviation, minimum, maximum, and range) as well as a frequency distribution for each score value and task. They were directed to use this information to adjust their initial judgments.

For the last half of day three, participants were divided into two groups: (a) an impact analysis group that viewed the number and percentage of students reaching each level of performance (advanced, proficient, below, and far below) and (b) a review of administration using extended levels of support for students who took the test under 'changed' conditions. In both groups, the focus was on eliciting feedback relevant to a final recommendation for adequate yearly progress reporting at the systems level.

An evaluation of the standard setting process was conducted to ascertain participant's perception of both the process and the certainty of their judgments. This section summarizes the entire process and provides all forms used throughout the three days.

Description of the Steps

The training began with an overview of all 10 steps that were to be used in the standard setting process. Participants were provided a Powerpoint® presentation with handouts. After the orientation, they were provided test booklets for their subject area and standard setting booklets for each subject area and grade band. See Appendix B.

Step 1. Train on Test Administration, attending to both scoring protocols and student materials. On the first morning of standard setting, all participants were trained on the test administration and scoring. A power point slide show was presented with examples from reading, writing, and mathematics to illustrate some of the nuances that were specific to each subject area as well as in common across subject areas. All participants were then directed to work with partners at their subject area table and practice administration and scoring.

Step 2. Review Extended GLEs. In this step, participants examined all of the Extended Grade Level Expectations (ExGLEs) for each grade band (3-4, 5-6, 7-8, 9-10). These ExGLEs included those aligned to tasks as well as those tested locally or not tested at all.

Step 3. Review cross walk document and relevant tasks; consider other tasks to grade level band. Of the entire group of ExGLE's, only those that were to be tested were cross walked to the tasks in the alternate assessment. In this step, a matrix was presented with rows and columns. The rows reflected the ExGLE's organized into strands (attributes). Each column represented the task number, the number of tasks, the number of items, and the number of points (from each task) for each relevant ExGLE. Participants also were provided each task and all items within to use in reviewing the cross walk. They were permitted to add other tasks if they thought it necessary.

Step 4: Review draft proficiency level descriptors (PLDs). Draft proficiency level descriptors for each grade band and subject area were provided in the booklet. These descriptors included language that depicted four levels of proficiency: (a) advanced, (b) proficient, (c) below, and (d) far below.

Step 5. Establish proficiency levels for <u>each task</u> and for each level. Participants were directed to award the number of points for EACH relevant task at that grade band) reflecting an appropriate cut score for defining the bottom boundary of (a) advanced (b) proficient, (c) below, and (d) far below. They were trained to begin with proficient and then move up and down the 'scale'.

Step 6. Confer with partner and articulate your rationale. Consider the issues in item 1 above. Reach consensus. These judgments of minimum cut scores at each proficiency level were made independently within each table and then consensus was reached in which the table compared their individual cut scores and discussed whether it was too high or too low (relative to others at the table). Often the discussion focused on the specific skills being tested in the task, the population being addressed (students with significant disabilities) relative to the general education students, and their opportunity to learn.

Step 7. Review grade band data set to confirm your reasoning of <u>each task</u>. For every cross-walked task at every grade band within each subject area, relevant student performance data were presented: (a) group statistics that reflected the average and normal variance around it (e.g. standard deviation) as well as the count, the minimum and maximum scores, and the range; (b) frequency distributions for every score reflecting the number of students. See Appendix C.

Step 8. As a group, devise a combined (across task) judgment for each proficiency level. Use a weighting system to confer a value for individual tasks that allows them to be combined.

This step allowed participants to change the weighting for their grade band so that the equal weights of each strand could be adjusted to reflect a different value. No group to chose to do so.

Step 9. Complete the Standard Setting Form for your grade level. The standard setting form was completed, transposing the values from individuals to a common form for each table. These data were then entered into statistical software for each proposed cut scores to determine frequency counts of students at each proficiency level within grade bands.

Step 10. Review Impact Data. On the last day, half of the participants were presented impact data reflecting the number and percent of students achieving each of the proficiency levels for each grade band and subject area. Differences were presented between the use of raw scores and weighted scores.

Extended Notes from the Final Meeting

5/4/07 – 12:00 – Regroup and summarize what the groups discussed and decided on:

ELOS Group:
Simply put, ELOS was used and points were accrued, the student was far below. No one disagreed with this judgment. The key piece for next year is to feather in some scoring options, integrated scoring with use of pictures in reading and writing. This will give more structure to the whole set.

The AT survey was very informing. The low population did not have a lot of communication. DRA received AT surveys for 175 students. The pattern showed that ELOS was used in the early tasks, and then the administration shifted to NT in the later, more difficult tasks. Typically in the early tasks about 14-16% were ELOS, 2-3% were NT. These numbers flipped by the later tasks and bumped up to about 30% NT.

The Oregon scaffolded test was reviewed and introduced as a possibility for scaffolding in the 2008 Alaska test.

Articulation Group:
Reading – was interesting to look at the numbers and then see them on a graph and how they shook out. More of a bold curve, a huge number of students had zeros, and some had all questions right. Pattern was, either they have it or they don't have it. Slight discrepancy in 5-6th gr. Quite a few proficient and advanced, making recommendations for someone to look at the trends over the grade levels to smooth out the transition better. Particularly for reading and writing that DRA could smooth out the transition. Interesting how the graph changed when the weighted pieces were factored in.

Weighting vs. not weighting – writing didn't change much because there were only 2 strands. Reading was interesting. Paul doesn't think weighting makes that much difference, other than reading 5-6, which he will double check. But everywhere else, it didn't make a huge change. Paul believes it's more accurate, but also raises some other issues.

As states report out, districts report out, and AYP gets distributed, the reports are not scarily different from the portfolio. Only about 10-15% difference in percentage points, which is pretty reasonable. As long as the difference is in that ball-park, it seems fairly practical. Classifications were such in grade bands that a small majority are in the just below proficient than in the far below. More who are below than proficient and advanced, but it's fairly evenly distributed between the two major sides of the spectrum.

A question arose that some students were possibly wrongly put in the 1% that need to take the alternate assessment.

2007/2008 testing window proposed training and test changes:

- Modifying some tasks, adding pictures
- Having some definitions/guidelines on how to administer and score ELOS. Make this ELOS report more useful for the teachers. One suggestion to train teachers to put in zeros for all ELOS tasks, so when they printed out the red line report, a score will come up. A good example was the write your own name, instead of ELOS for the whole task, put in standard so the kid gets points for the first name, and zeros are entered for the last name.
- The whole process will become better with time as the teachers become more comfortable with the test.
- Math in particular – felt there needed to be a larger number of tasks/items. Kids were scored on a really small task samples. The math group felt there needed to be an opportunity to get more points.
- Guidelines on training/refreshing recommendations, should they do certain items in writing, 40%, 80%? What are the requirements so people will actually follow through?
- The ELOS samples on the state website or on the testing site would be wonderful, if there was a link to examples of accommodations and ELOS approaches.
- **Tasks for refreshing** – scoring on correct letter and correct word sequences. Items with partial scoring. The early tasks don't need refreshers. The harder tasks, the tasks that were harder for people to get through, maxed out trials, should do at least 3 of each subject area.
- No administration scoring where you watch the video.

- Was helpful for teachers to print out what was wrong and go over correcting the answers with them. The administration samples actually showed better examples than the initial training.

SEAS
-Really helps to process ahead of time and get some structure on how to administer the test.
-Understanding Accom/ELOS up front will help the teachers administer the test. Those who administered under ELOS were trying to design on the fly, and this would help them know ahead of time.

Scoring Subtleties and Protocol changes
-ELOS, how far you keep administering if you're giving a full score.
-Clarification on what to point for, when is an ELOS administered, and what kind.
-Clean up the scoring sheets, timed vs. un-timed
-Get rid of the NA-I scoring instructions, confusing, and not correct.
-Comment section for NA-I administration
-Specific place on protocol for accommodation that is being used, this document is to help teachers, so if the student does not have the same teacher the next year, they can reference the scoring protocol for accommodations.
-Writing – write your name could be earlier in the test, not task 4.
-Teacher name and qualified assessor name on front of packet.
-Cover sheet that should be filled out for each student who should be taking this test even if they are not being tested (extended absence, etc.). Need to document why a student included in the AYP is not being tested.
-Reading, since we're not going to do timed, make the passages all 100 words to make the math easy, and work on more high interest stuff for the older kids. They will get careless if they're reading 'baby stories.'
-More age appropriate pictures.
-Directions – sometimes it was hard to find what you were supposed to say to the student. GenEd has it more formatted. Particularly Task 7, writing, people might administer in different ways. More clarification and standardize the formatting for the instructions.
-Concern about the standardization of the instructions – that rephrasing the question is an acceptable accommodation, but needs to be noted.
-Writing – where student had to write own story, then having student read story back to administrator was really difficult. Sometimes the student would read something different than they had written. Confusing for teacher to score for CLS.
-Vote not to have that "This task is going to be used for this grade" keep the growth model.

Quantitative Evaluation from Participants

Table 32. Evaluation Summary [a]

Evaluation Topic	Strongly Disagree	Disagree	Agree	Strongly Agree
1. Training Mat.	0	2	9	12
2. 10 step process	0	0	13	11
3. admin/score test	0	1	6	17
4. ExGLEs	0	0	8	16
5. PLDs	0	0	5	19
6. cut scores accurate	0	0	8	16
7. Consensus reached	0	0	3	21
8. Cut scores are defensible	0	0	9	14
9. AA realistic/ accountable	0	2	18	3

[a] *See Appendix 4D for the Evaluation Form*

Table 33. Comments from Participants

Improvement	Positive Comm.
Great organization, no ideas for improvement; , Some time line give-up	There was an excellent diversity of individuals. The statistics were explained very well by Paul.
Fine tune PLDs	Enjoyed your instruction and directions - I actually enjoyed statistics today.
As with the GLE's, the PLD's are not all inclusive, there is still a lot of room for improvement	Very thorough, very knowledgeable and proficient DRA staff, helpful handouts. Good step by step walk through of needed steps and process
If time would have been beneficial to incorporate our cut scores into the data to see how it would look - Do this beside the other outcome data.	Very important to have a reading specialist in group. Conversations were much appreciated to come to a consensus
The materials were not user friendly. Binders to allow for movement of papers would be helpful. I knew very little about this assessment and didn't feel that I needed to know. Now I feel that this is critical information for a regular ed teacher.	Paul was extremely helpful with infor! Easy to understand. Jerry is great with explanations! Sevrina with her support.
Going through the statistic packet for one task and what the figures meant as a group would have been helpful.	Paul is like the Mr. Rogers of statistics! Good process. Great group.
Further discussion and decisions related to ELOS. Next time generate summary of scores proficient plus above and other stakeholders, principals? Parents? BP plus FBP	Interesting to see the process. Everything was very organized
Fix the small inconsistencies/flows on actual test that were noted	This process was very good to facilitate a group product with all input. Having general ed people involved helped give insights of expectations.
I thought it was well organized, no time to divert the process for the end result	Good session, clear process of individual work and group consensus. Good group of people with different experiences at the table.
Practice the process before doing the process	Great team work made by our members! Paul and Jerry were great! Thanks!
	Jerry and staff were very helpful and patient through a difficult process
On time coffee	I thought the process was well thought out. Our team seemed to work extremely well together doing things individually first and then coming to a consensus was a smooth transition for our team. Running the #'s against the data served to confirm our original thoughts in setting the cut scores.
	I really enjoyed interacting with other colleagues from across the state. I thought Jerry, Paul, and their assistant helped to make this a great experience.
	The group process was very effective
	Lots of chocolate! :) All Dillard Assoc. were very knowledgeable and able to respond to our team's concerns. Despite high expectations, I felt at ease with no stress. Thanks!

CLASSIFICATION RESULTS WITH IMPACT DATA AND JUSTIFICATION

Weighted Content Area Scoring

All grade level expectations were organized into strands or attributes (both terms are used in EED documents). These grade level expectations were then cross-walked to relevant grade band tasks (3-4, 5-6, 7-8, and 9-10). For each content area (Math, Reading, Writing), Alternate assessment raw scores were then computed using only the tasks that had been cross-walked to the grade-level expectations. However, the number of relevant tasks varied considerably by strand (within and across all grade bands), making it possible for a student to have one strand out value another strand if raw scores were used to set standards.

Therefore, all strands were weighted to make them equal within each of the three subject areas and grade bands. To do this, the total test was set to 100 points and then divided by the number of strands. If a subject area had two strands (e.g. grade 3-4 had *Word Identification* and *Forming a General Understanding*), then each strand was weighted at 50 points. If the subject area had three strands, then each would be worth 33.3 points. Then the points from all tasks that had been cross-walked to grade level expectations within that strand were multiplied by a weight to make the total equal the requisite points needed. For example, in grade 3-4 reading, the *Word Identification* strand had 56 points from relevant tasks and *Forming a General Understanding* had 12 points. To make them equal in weight, the former strand total points was multiplied by .9 (to equal 50 points) and the latter strand total points was multiplied by 4.2 (to equal 50 points). A default assumption was made that the content strands per grade band would be valued as equally important. The standard setting panels, however, were free to change those strand value allocations. No changes to the 'default' values were made.

Procedure for Weighting

Following are the procedures for computing (a) content area strand weights, (b) weighted scores within grade band, (c) strand weighted performance standards, (d) content area weighted performance standards.

1. Compute the raw scores per examinee
2. Compute the total possible points per strand
3. Compute weights per point within strand: strand weights=(100/#strands)/#points per strand
4. Compute weighted strand scores: weighted strand score=strand raw score **X** strand weight
5. Compute the content area weighted total score: weighted content total score=sum(strand weighted scores)
6. Compute strand weighted performance standards (pstandard): weighted strand pstandard=strand raw pstandard **X** strand weight
7. Compute strand raw p-standard strand weight
8. Compute content area weighted performance standards: content area weighted performance standard=sum(strand pstandards)

In addition to weighting, a standard error of measurement (SEM) was applied to adjust the cut scores. The SEM can be used to account for normal measurement error and used to either control for false positives (students who are deemed as proficient but are not) or false negatives (students are deemed as not proficient but really are). After the articulation meeting, it was decided that an adjustment would be to the weighted score with 1 SEM; in effect the cut score would be lowered at each category value (far below-below, below-proficient, and proficient-advanced) to reduce the number of false negatives. Following are the cut scores for raw and weighted values, each adjusted by 1 SEM.

Results

Table 34. Reading – Cut Scores for Weighted Scores (Scaled to 100) with 1 SEM

	Advanced	Proficient	Below	Far Below
3-4	63	32	8	<8
5-6	77	46	11	<11
7-8	52	33	12	<12
9-10	57	43	22	<22

Table 35. Writing – Cut Scores for Weighted Scores (Scaled to 100) with 1 SEM

	Advanced	Proficient	Below	Far Below
3-4	76	38	7	<7
5-6	67	33	10	<10
7-8	76	41	16	<16
9-10	82	47	24	<24

Table 36. Reading/Writing – Cut Scores for Weighted Scores (Scaled to 100) with 1 SEM
\

	Advanced	Proficient	Below	Far Below
3-4	139	70	15	<15
5-6	144	79	21	<21
7-8	128	74	28	<28
9-10	139	90	46	<46

Table 37. Math – Cut Scores for Weighted Scores (Scaled to 100) with 1 SEM

	Advanced	Proficient	Below	Far Below
3-4	62	33	6	<6
5-6	61	25	8	<8
7-8	74	52	22	<22
9-10	81	63	24	<24

Table 38. Reading/Writing Counts and Percentages of Examinees in Proficiency Categories by Grade Band (Weighted Scoring and Weighted with 1 SEM)

		Grade Band							
		34		56		78		910	
		N	%	N	%	N	%	N	%
Weighted Score Proficiency Level (-1sem)	far below	31	25.0%	41	33.3%	53	42.7%	67	38.7%
	below	20	16.1%	33	26.8%	21	16.9%	28	16.2%
	proficient	44	35.5%	27	22.0%	25	20.2%	52	30.1%
	advanced	29	23.4%	22	17.9%	25	20.2%	26	15.0%
	Total	124	100.0%	123	100.0%	124	100.0%	173	100.0%

Table 39. Math Counts and Percentages of Examinees in Proficiency Categories by Grade Band (Weighted Scoring and Weighted with 1 SEM)

		Grade Band							
		34		56		78		910	
		N	%	N	%	N	%	N	%
Weighted Score Proficiency Level (-1sem)	far below	28	22.6%	26	21.1%	49	39.5%	66	38.2%
	below	18	14.5%	10	8.1%	25	20.2%	25	14.5%
	proficient	31	25.0%	47	38.2%	12	9.7%	30	17.3%
	advanced	47	37.9%	40	32.5%	38	30.6%	52	30.1%
	Total	124	100.0%	123	100.0%	124	100.0%	173	100.0%

Justification for Use of Weighted Scores with 1 SEM

The following supports the use of weighted scores and presents cut scores and percentages in AYP categories from raw scores as well as weighted scores without 1 SEM adjustment.

1. The overall percentages of passing are actually very reasonable and not at all surprising. The standard setting group endorsed them with a strong degree of confidence..
2. A close inspection of the data reveals only 1 anomaly comparing the weighted with the raw score: In grade 5-6, there is a considerable decrease in the percentage passing when the weighted scores are used. In most other grades, the weights are neutral or help (by adjusting in areas where more items are needed). This outcome is likely to be muted when the average of reading and writing is used for reporting AYP.
3. The outcomes from using the weighted scores are very close to those of the raw scores in most grade bands and subject areas. There are very few outcomes that are not explainable. For example, the shift downward in passing rates for high school is very understandable as this group of students has the shortest history of learning in an academic classroom and is likely on a diet of preparation for post secondary environments that is less likely to include a strict emphasis on reading, writing, and mathematics.
4. The use of the weighted scores allows a better transition next year as the field will be trained in interpretations of passing relative to a 100-point scale. In fact, we will use the summer to better capture items within tasks so the weighting for each strand should be negligible (but is still important to consider as the strands are never the same across all the grade bands). Nevertheless, the nature of this test is both growth-oriented (teachers can view change over time) AND standards oriented. To do this, requires a complex interleaving of tasks and items that are best assembled with a very simple and straightforward weighting process.
5. The tasks had very high reliability coefficients with very high internal consistency. When the data were analyzed for 'item deletion' (in other words, what would be the effect if an item within a task was removed), very few differences were found. This result is very important because it means that the task is a

simple linear combination of items (and the total can result from a simple addition of items; an extrapolation of this finding is that the total can be based on any combination of items – e.g. multiplied to increase or decrease the amount).

6. The findings from an analysis of adjusting the cut scores to a lower bound using the standard error of measurement (SEM) results in more acceptable outcomes politically: The results obviously increase the percentages passing.

Table 40. Reading – Cut Scores for Unweighted Raw Scores

	Advanced	Proficient	Below	Far Below
3-4	61	37	19	<19
5-6	57	35	15	<15
7-8	110	79	42	<42
9-10	115	90	53	<53

Table 41. Reading – Cut Scores for Unweighted Raw Scores with 1 SEM

	Advanced	Proficient	Below	Far Below
3-4	51	27	9	<9
5-6	52	30	10	<10
7-8	99	28	31	<31
9-10	107	82	45	<45

Table 42. Reading – Cut Scores for Weighted Scores (Scaled to 100)

	Advanced	Proficient	Below	Far Below
3-4	78	45	21	<21
5-6	90	59	24	<24
7-8	61	42	21	<21
9-10	62	48	27	<27

Table 43. Writing – Cut Scores for Unweighted Raw Scores

	Advanced	Proficient	Below	Far Below
3-4	140	82	36	<36
5-6	119	71	37	<37
7-8	85	55	30	<30
9-10	52	33	21	<21

Table 44. Writing – Cut Scores for Unweighted Raw Scores with 1 SEM

	Advanced	Proficient	Below	Far Below
3-4	115	57	11	<11
5-6	96	48	14	<14
7-8	69	39	14	<14
9-10	44	25	13	<13

Table 45. Writing – Cut Scores Weighted Scores (Scaled to 100)

	Advanced	Proficient	Below	Far Below
3-4	93	54	24	<24
5-6	82	49	26	<26
7-8	91	57	32	<32
9-10	94	59	37	<37

Table 46. Math – Cut Scores for Unweighted Raw Scores

	Advanced	Proficient	Below	Far Below
3-4	41	25	11	<11
5-6	125	54	27	<27
7-8	101	72	35	<35
9-10	98	77	35	<35

Table 47. Math – Cut Scores for Unweighted Raw Scores with 1 SEM

	Advanced	Proficient	Below	Far Below
3-4	31	15	1	<1
5-6	107	36	9	<9
7-8	91	62	25	<25
9-10	85	64	22	<22

Table 48. Math – Cut Scores for Weighted Scores (Scaled to 100)

	Advanced	Proficient	Below	Far Below
3-4	76	47	19	<19
5-6	69	32	17	<17
7-8	80	58	28	<28
9-10	91	74	35	<35

Table 49. Original Reading Counts and Percentages of Examinees in Proficiency Categories by Grade Band (Weighted Scoring and Weighted with 1 SEM)

		Grade Band								Total	
		34		56		78		910			
		N	%	N	%	N	%	N	%	N	%
Weighted Score Proficiency Level	far below	55	43.0%	51	41.5%	72	56.7%	81	47.9%	259	47.3%
	below	30	23.4%	42	34.1%	13	10.2%	26	15.4%	111	20.3%
	proficient	32	25.0%	25	20.3%	35	27.6%	48	28.4%	140	25.6%
	advanced	11	8.6%	5	4.1%	7	5.5%	14	8.3%	37	6.8%
	Total	128	100.0%	123	100.0%	127	100.0%	169	100.0%	547	100.0%
Weighted Score Proficiency Level (-1sem)	far below	26	20.3%	37	30.1%	60	47.2%	78	46.2%	201	36.7%
	below	40	31.3%	39	31.7%	20	15.7%	21	12.4%	120	21.9%
	proficient	38	29.7%	31	25.2%	23	18.1%	39	23.1%	131	23.9%
	advanced	24	18.8%	16	13.0%	24	18.9%	31	18.3%	95	17.4%
	Total	128	100.0%	123	100.0%	127	100.0%	169	100.0%	547	100.0%

Table 50. Original Writing Counts and Percentages of Examinees in Proficiency Categories by Grade Band (Weighted Scoring and Weighted with 1 SEM)

		Grade Bands									
		34		56		78		910		Total	
		N	%	N	%	N	%	N	%	N	%
Weighted Score Proficiency Level	far below	18	18.0%	43	42.6%	31	34.4%	34	26.8%	126	30.1%
	below	16	16.0%	15	14.9%	17	18.9%	31	24.4%	79	18.9%
	proficient	50	50.0%	28	27.7%	37	41.1%	62	48.8%	177	42.3%
	advanced	16	16.0%	15	14.9%	5	5.6%	0	.0%	36	8.6%
	Total	100	100.0%	101	100.0%	90	100.0%	127	100.0%	418	100.0%
Weighted Score Proficiency Level (-1sem)	far below	16	16.0%	28	27.7%	17	18.9%	14	11.0%	75	17.9%
	below	9	9.0%	19	18.8%	19	21.1%	37	29.1%	84	20.1%
	proficient	28	28.0%	24	23.8%	26	28.9%	60	47.2%	138	33.0%
	advanced	47	47.0%	30	29.7%	28	31.1%	16	12.6%	121	28.9%
	Total	100	100.0%	101	100.0%	90	100.0%	127	100.0%	418	100.0%

Table 51. Original Math Counts and Percentages of Examinees in Proficiency Categories by Grade Band (Weighted Scoring and Weighted with 1 SEM)

		Grade Band								Total	
		34		56		78		910			
		N	%	N	%	N	%	N	%	N	%
Weighted Score Proficiency Level	far below	37	29.1%	32	26.2%	58	45.3%	77	45.6%	204	37.4%
	below	28	22.0%	10	8.2%	21	16.4%	31	18.3%	90	16.5%
	proficient	30	23.6%	46	37.7%	19	14.8%	33	19.5%	128	23.4%
	advanced	32	25.2%	34	27.9%	30	23.4%	28	16.6%	124	22.7%
	Total	127	100.0%	122	100.0%	128	100.0%	169	100.0%	546	100.0%
Weighted Score Proficiency Level (-1sem)	far below	31	24.4%	25	20.5%	50	39.1%	66	39.1%	172	31.5%
	below	19	15.0%	11	9.0%	25	19.5%	25	14.8%	80	14.7%
	proficient	30	23.6%	46	37.7%	13	10.2%	27	16.0%	116	21.2%
	advanced	47	37.0%	40	32.8%	40	31.3%	51	30.2%	178	32.6%
	Total	127	100.0%	122	100.0%	128	100.0%	169	100.0%	546	100.0%

In the figures below, a comparison can be made between the weighted scores and the use of 1 SEM with the weighted scores to lower it. The upper graph of each subject area shows the percentage of students at each performance category when only the weighted score is used; the lower graph of each subject area shows the same percentages for each category. It is important to note the -1 SEM causes the percentages of students deemed 'far below' and 'below' to generally decrease; at the same time, the percentage of students deemed 'proficient' and 'advanced' generally increases when -1 SEM is used.

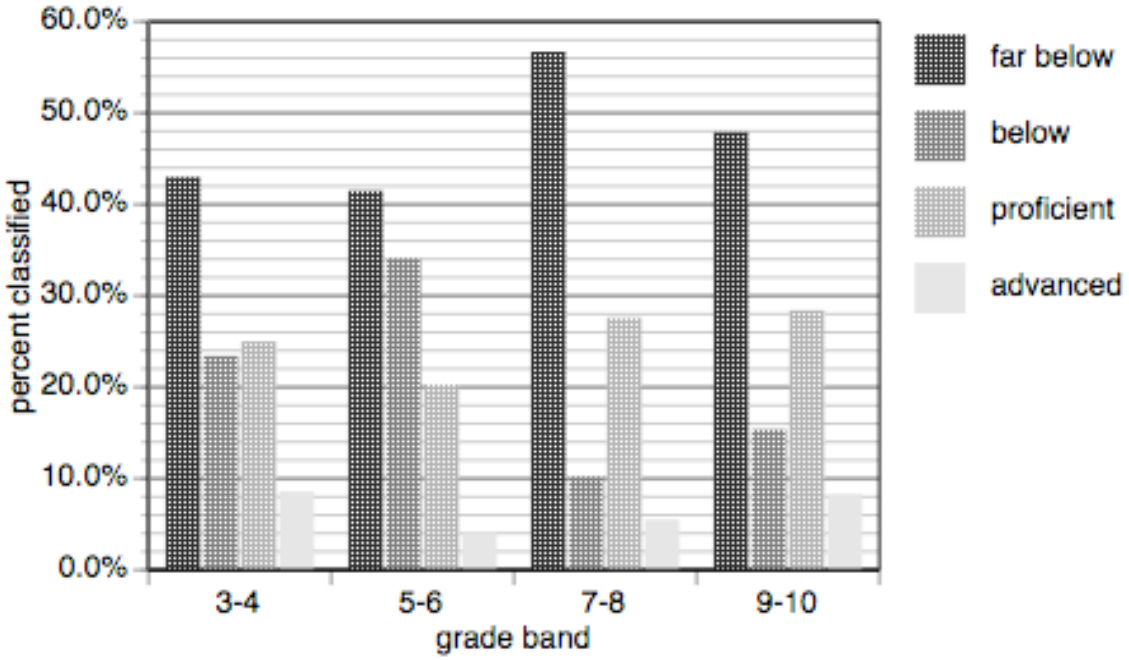


Figure 1. Reading Weighted Scoring and Weighted Cut Score

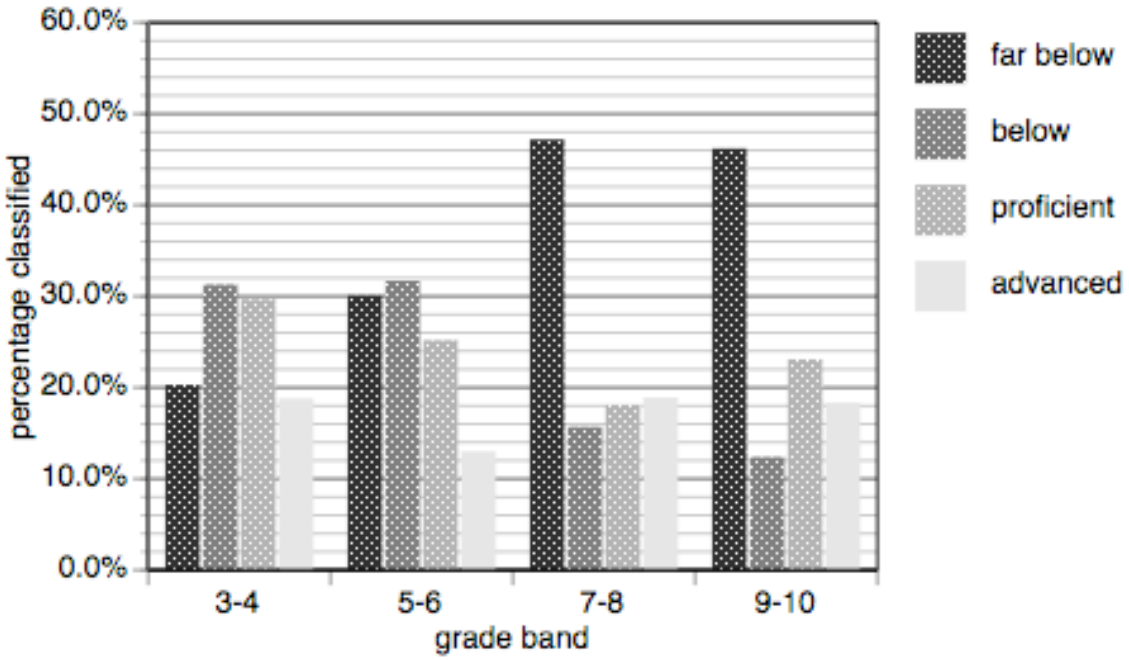


Figure 2. Reading Weighted Scoring and Weighted Cut Score (-1 SEM)

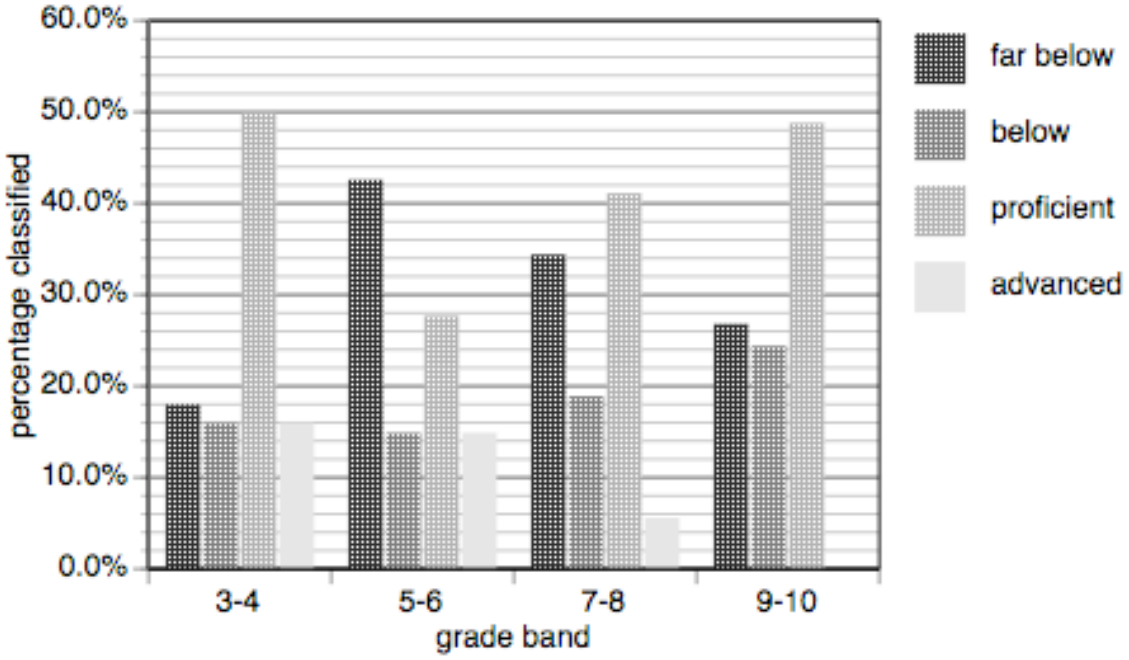


Figure 3. Writing Weighted Scoring and Weighted Cut Score

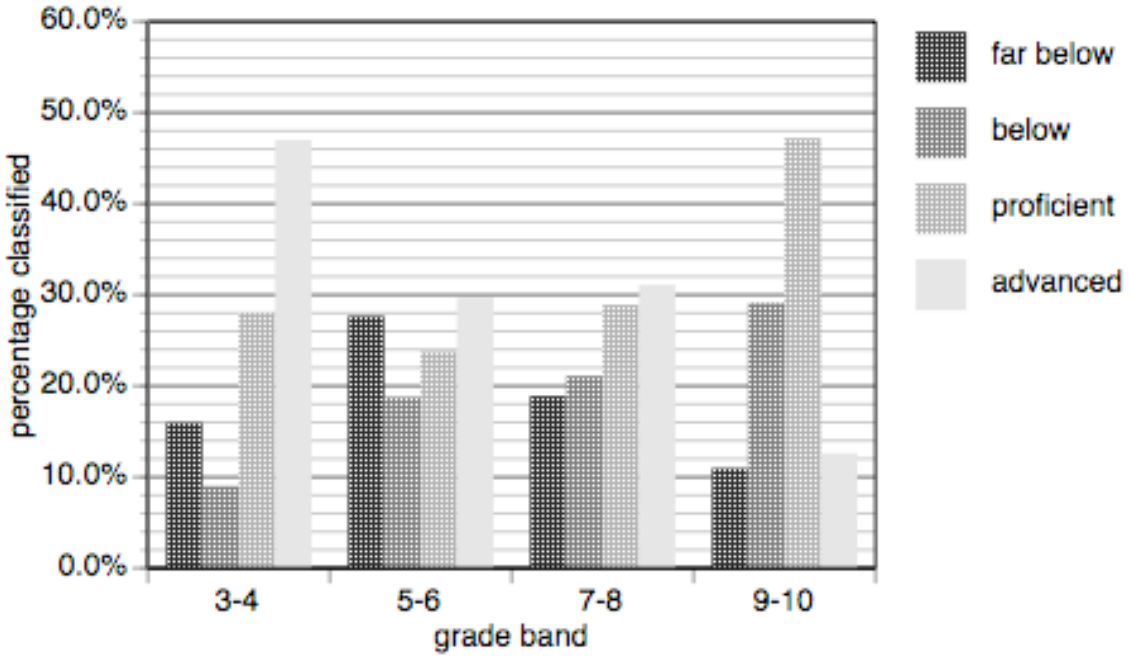


Figure 4. Writing Weighted Scoring and Weighted Cut Score (-1 SEM)

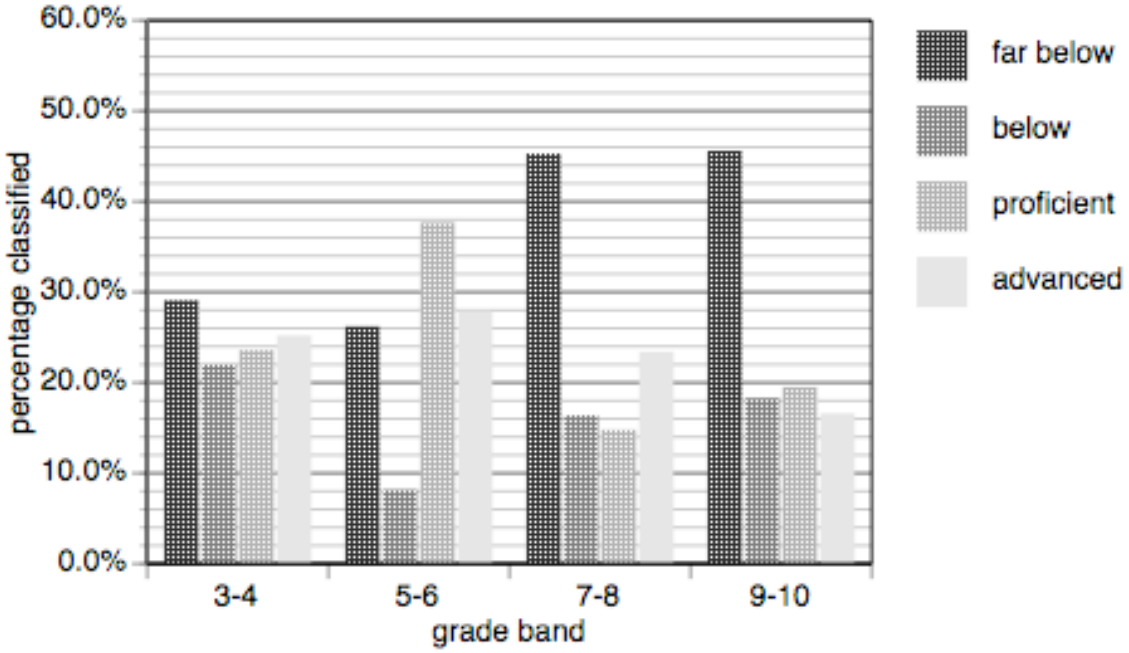


Figure 5. Math Weighted Scoring and Weighted Cut Score

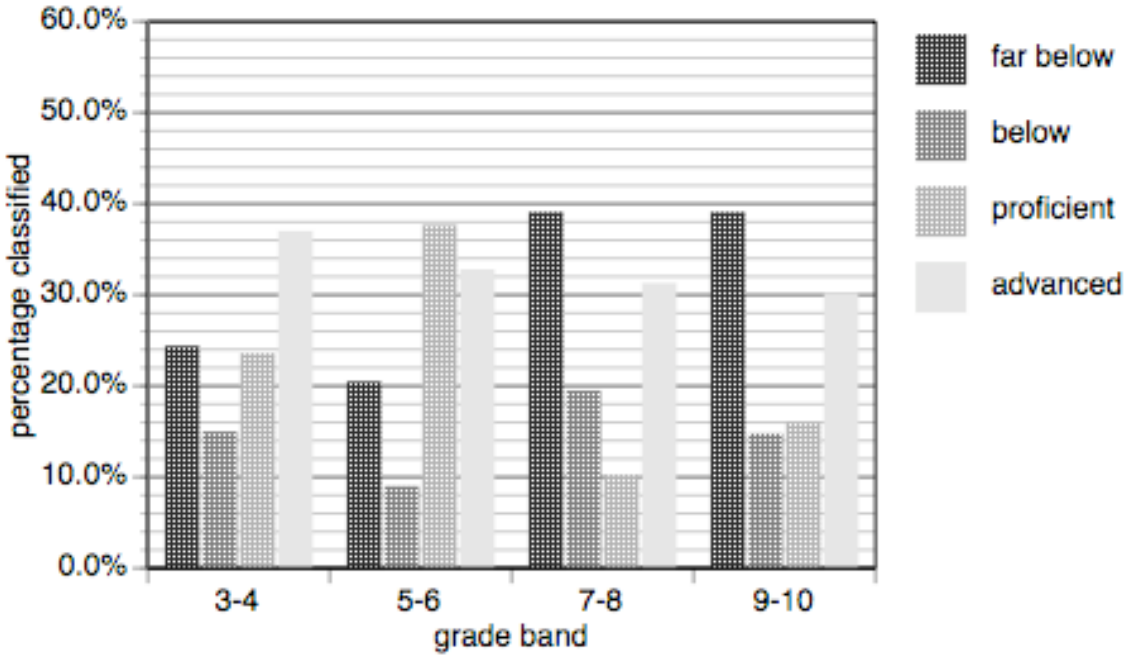


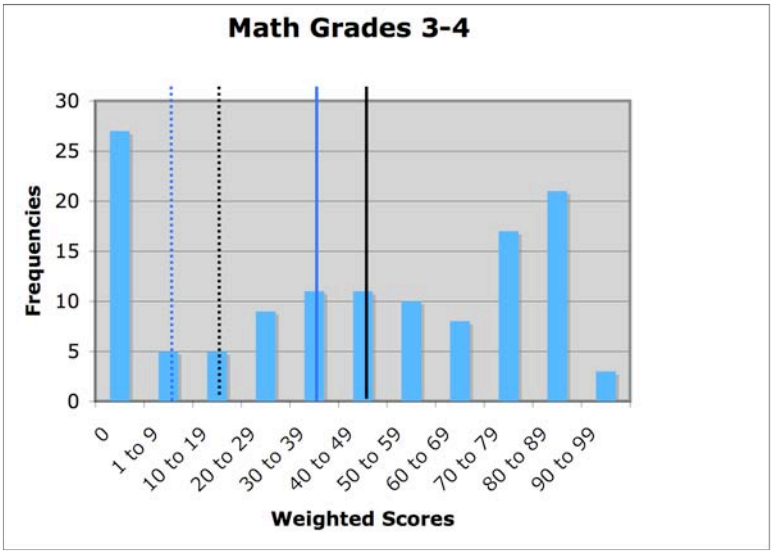
Figure 6. Math Weighted Scoring and Weighted Cut Score (-1 SEM)

In the remainder of this report, each subject area is graphed at each grade band with a distribution of students. Two graphs are presented on each page and for each grade band. The first is a bar chart of score bands with vertical lines and the second is a pie chart with three values of changes. Following is an explanation of these graphs.

Bar graph. All weighted scores within a subject area were organized into a distribution so that a bar chart would show the lowest interval (0) to the highest interval (90-99). Eleven intervals were selected (usually between 10 and 20 intervals are suggested). For example, in grade 3-4 math, the following number of students scored in each of these 11 intervals.

Score interval	No. students
0	27
1 to 9	5
10 to 19	5
20 to 29	9
30 to 39	11
40 to 49	11
50 to 59	10
60 to 69	8
70 to 79	17
80 to 89	21
90 to 99	3

The graph of this distribution is displayed on the right.

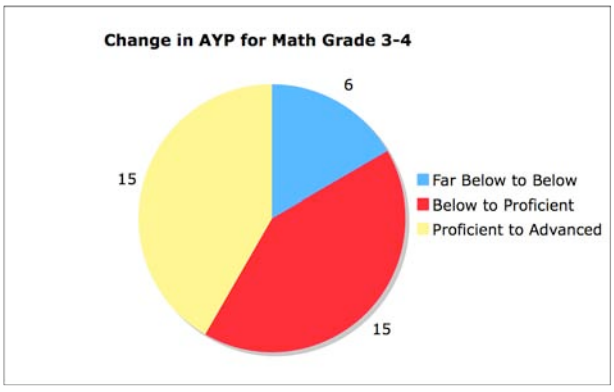


The vertical lines reflect four values from left to right:
• Left-most dotted blue-gray line is -1 SEM for the weighted cut score depicting *Below Proficient* (to the left is *Far Below*)
• Dotted black line is the weighted cut score for below *proficient* (to the left is *Far Below*)
• Gray-blue solid line to the left is -1 SEM for the weighted cut score for *Proficient* (to the left is *Below Proficient*)
• Right-most solid black line is the weighted cut score for *Proficient* (to the left is *Below Proficient*)

In effect the use of the cut score lowers the value needed for any given proficiency category.

Pie chart. In addition, the number of students who move into adjacent performance categories is depicted in a pie chart **when the -1 SEM is added to the weighted cut score**.

In the example to the right, 6 students move from *Far Below* to *Below*; 15 students move from *Below* to *Proficient*, and an additional 15 students move from *Proficient* to *Advanced*.



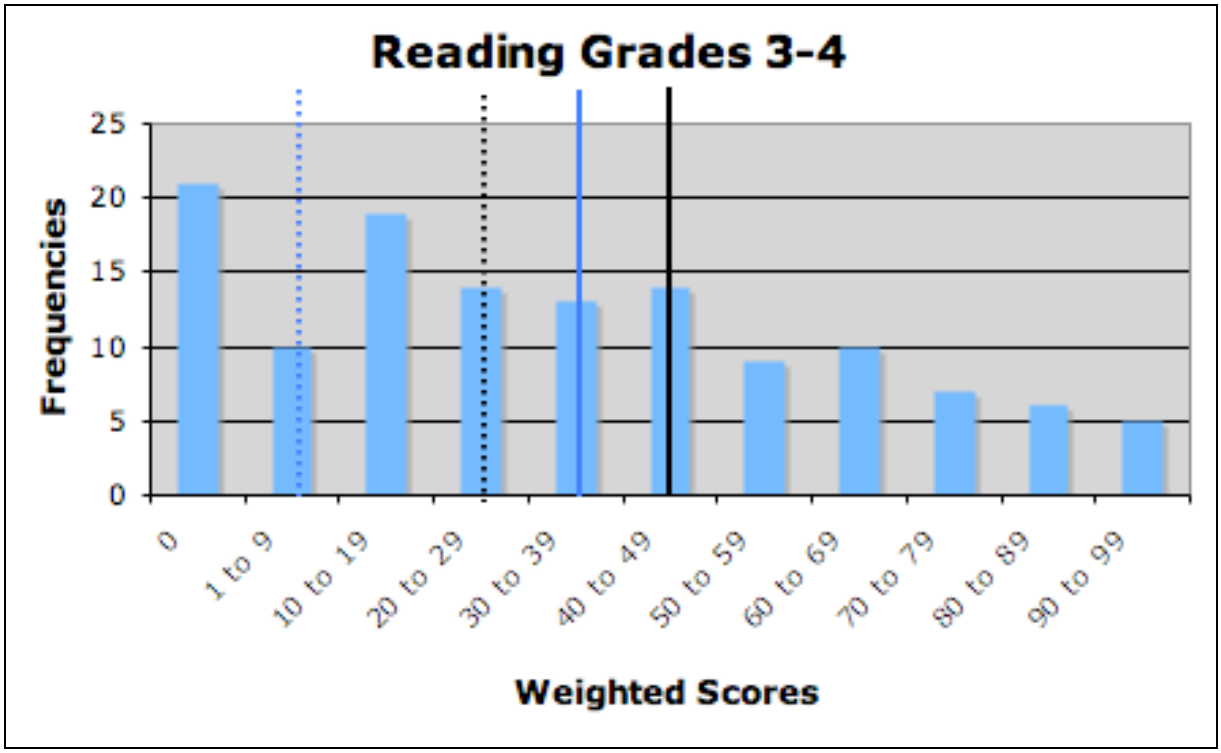


Figure 7. Reading Grade 3-4: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

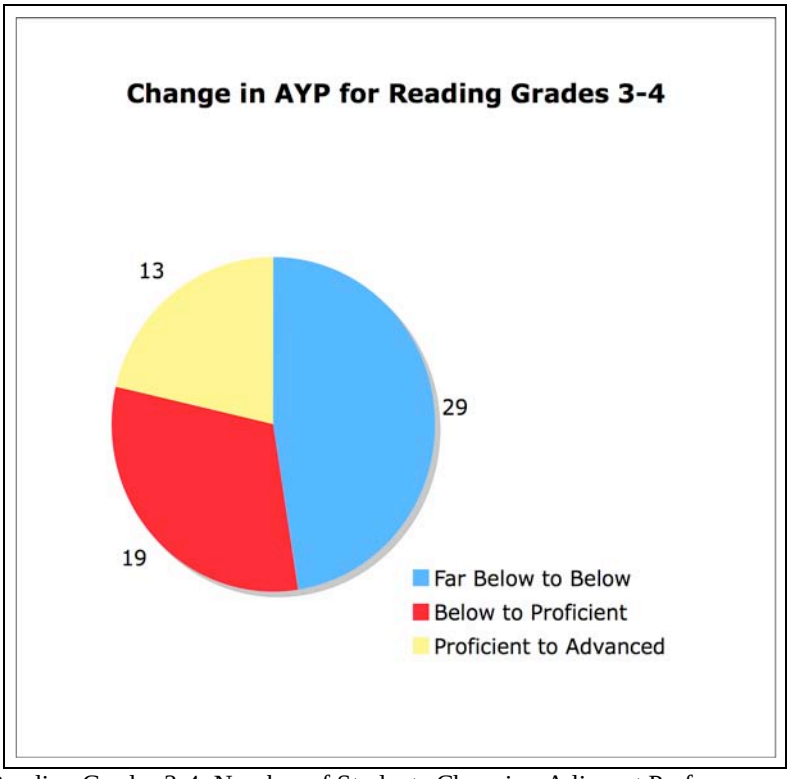


Figure 8. Reading Grades 3-4: Number of Students Changing Adjacent Performance Categories

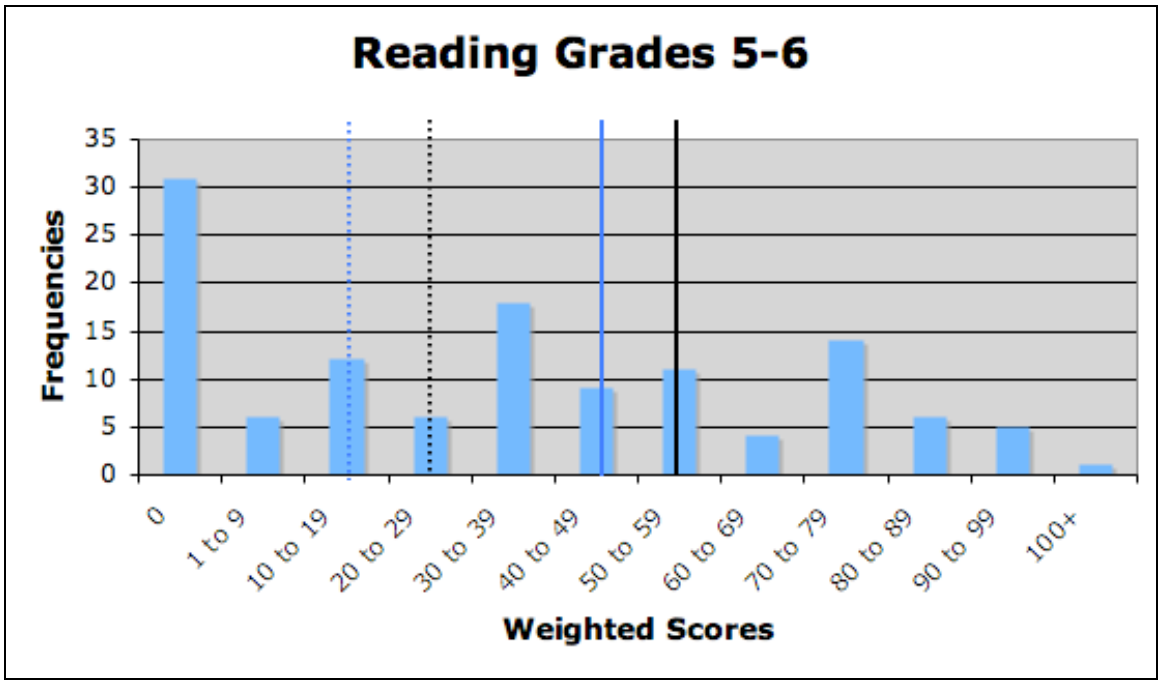


Figure 9. Reading Grades 5-6: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

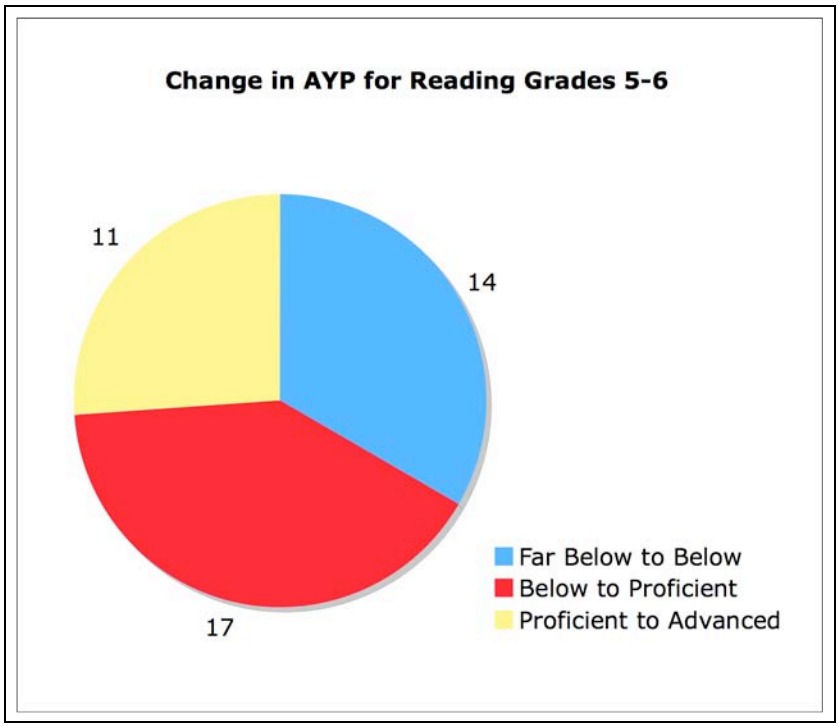


Figure 10. Reading Grades 5-6: Number of Students Changing Adjacent Performance Categories

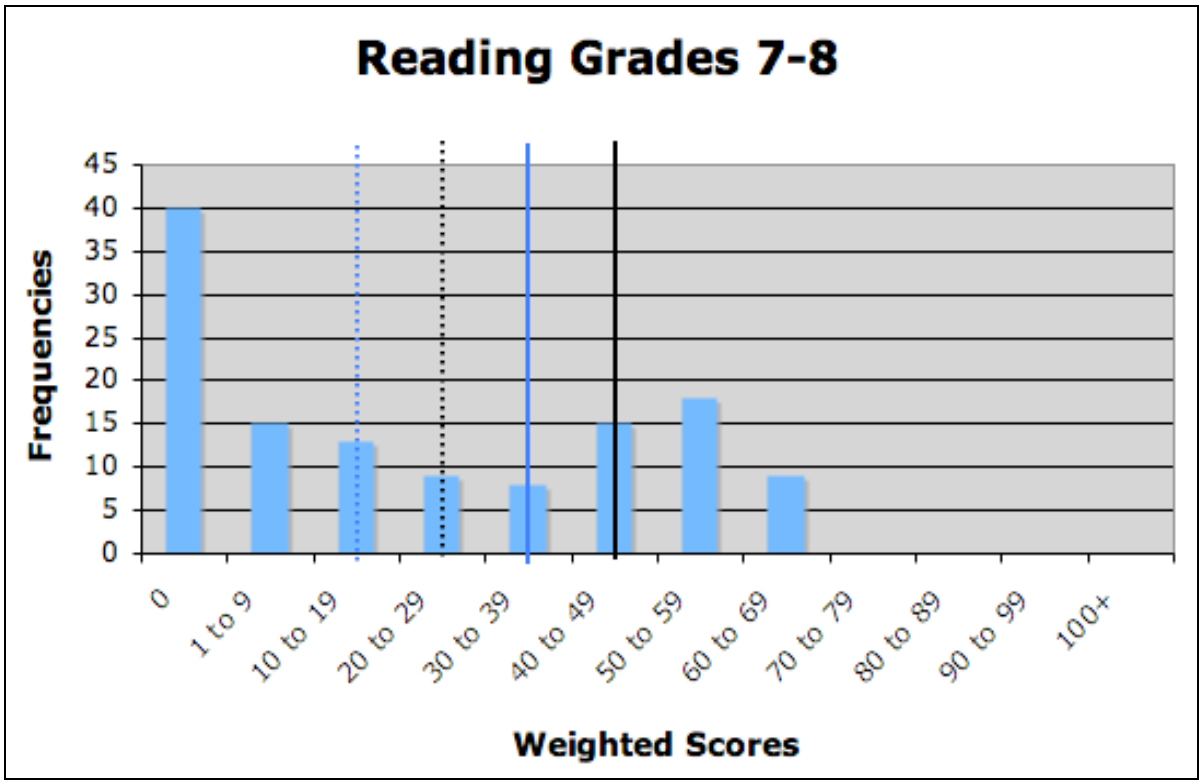


Figure 11. Reading Grades 7-8: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

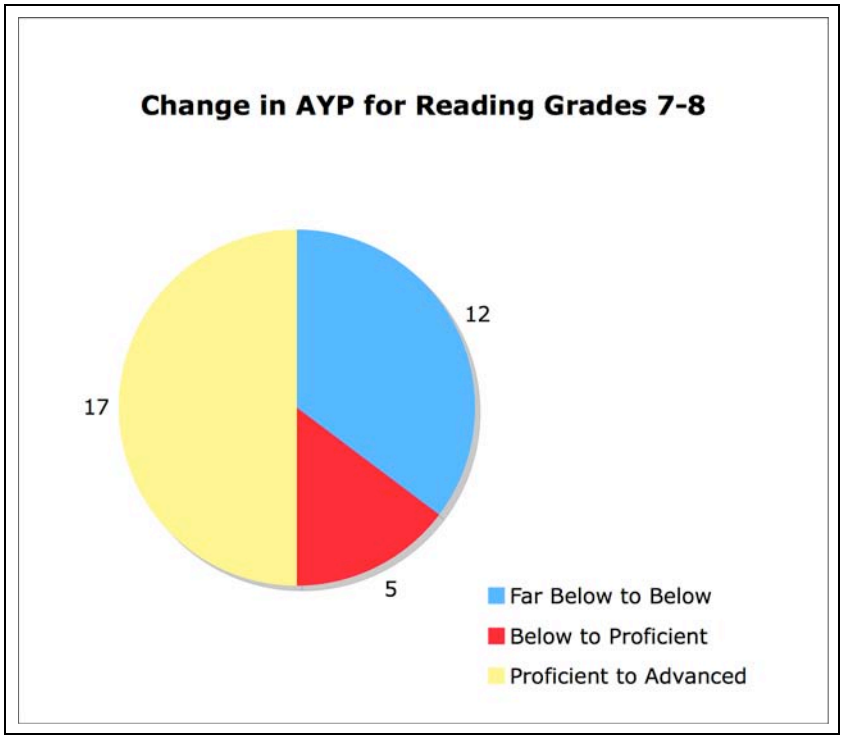


Figure 12. Reading Grades 7-8: Number of Students Changing Adjacent Performance Categories

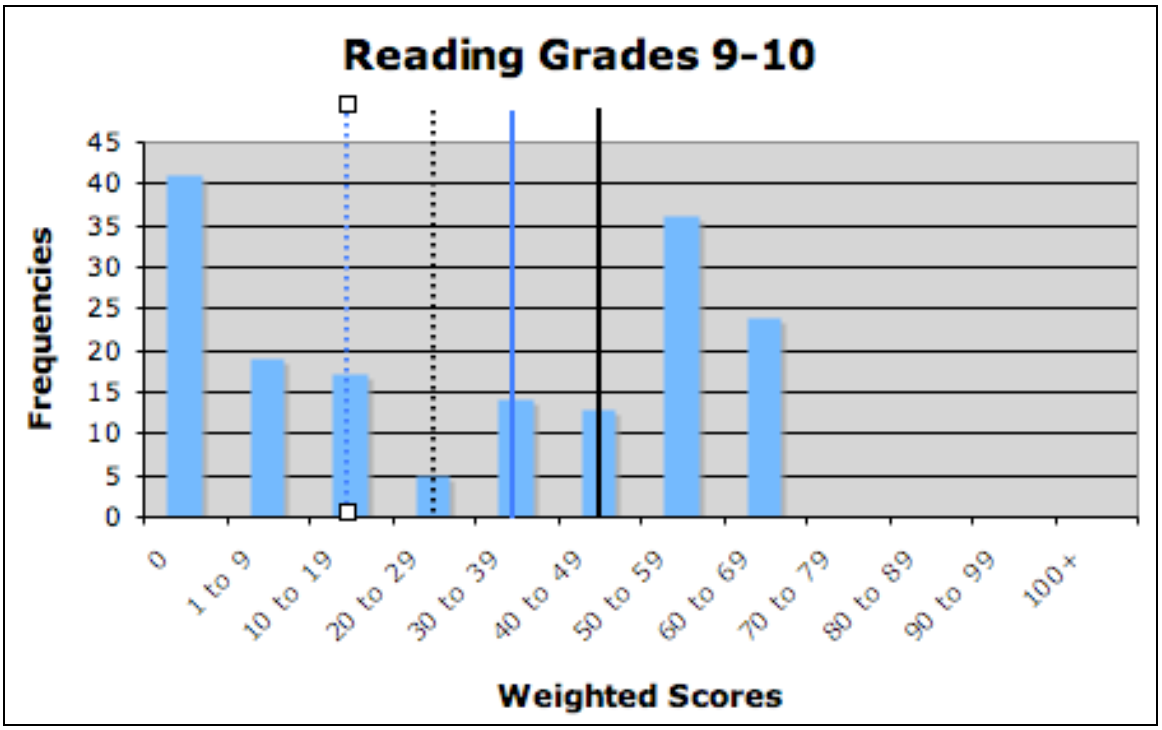


Figure 13. Reading Graded 9-10: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

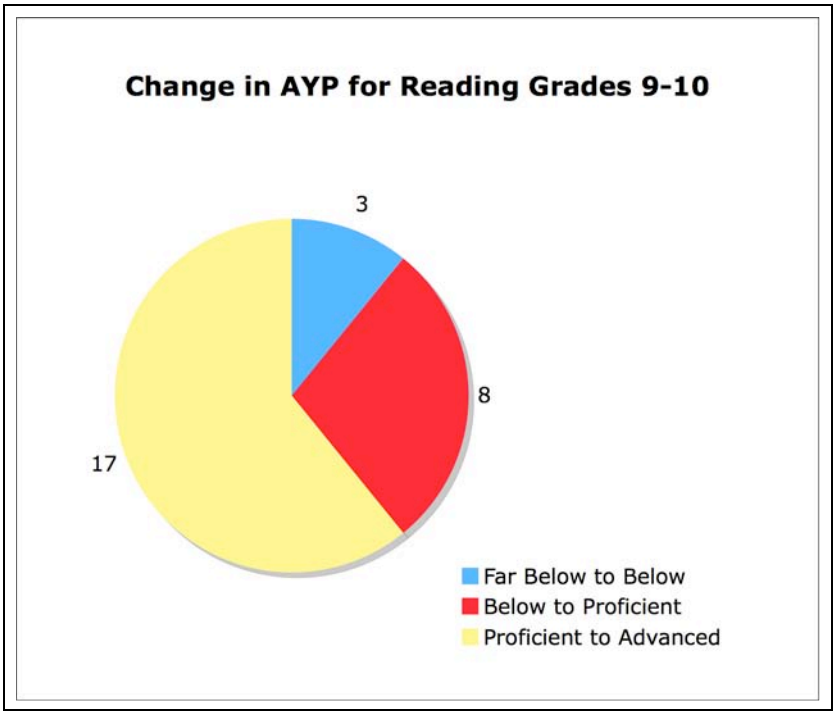


Figure 14. Reading Grades 9-10: Number of Students Changing Adjacent Performance Categories

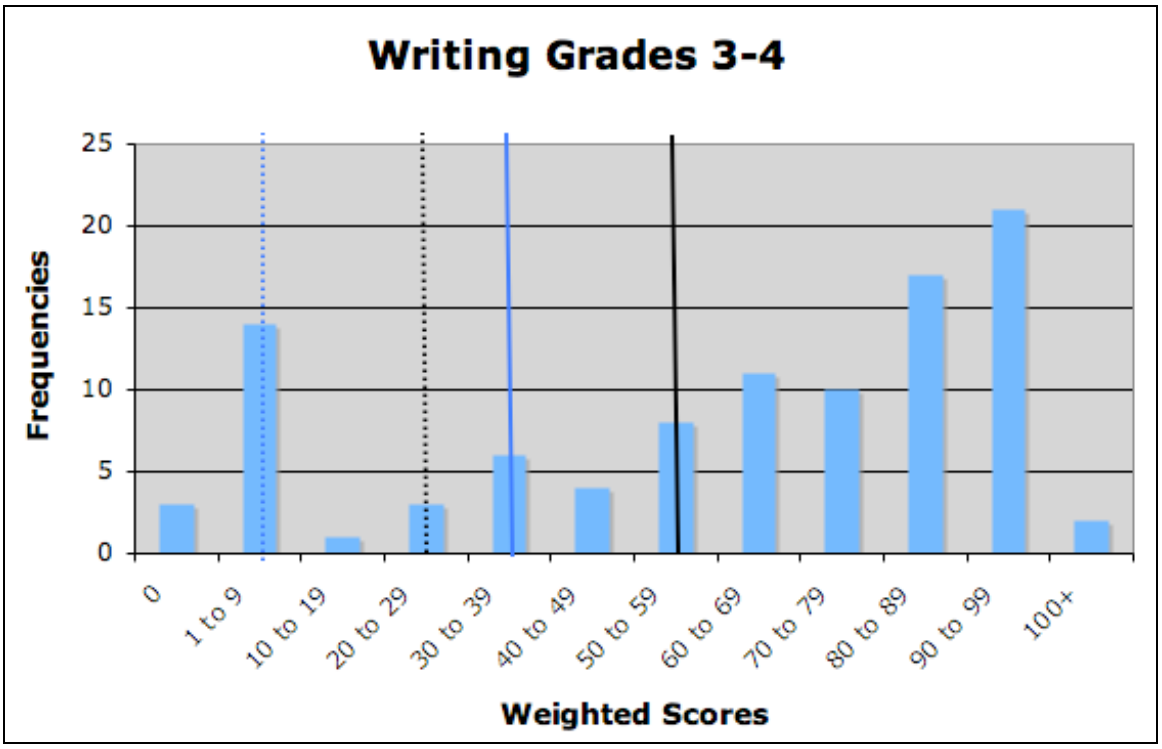


Figure 15. Writing Grades 3-4: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

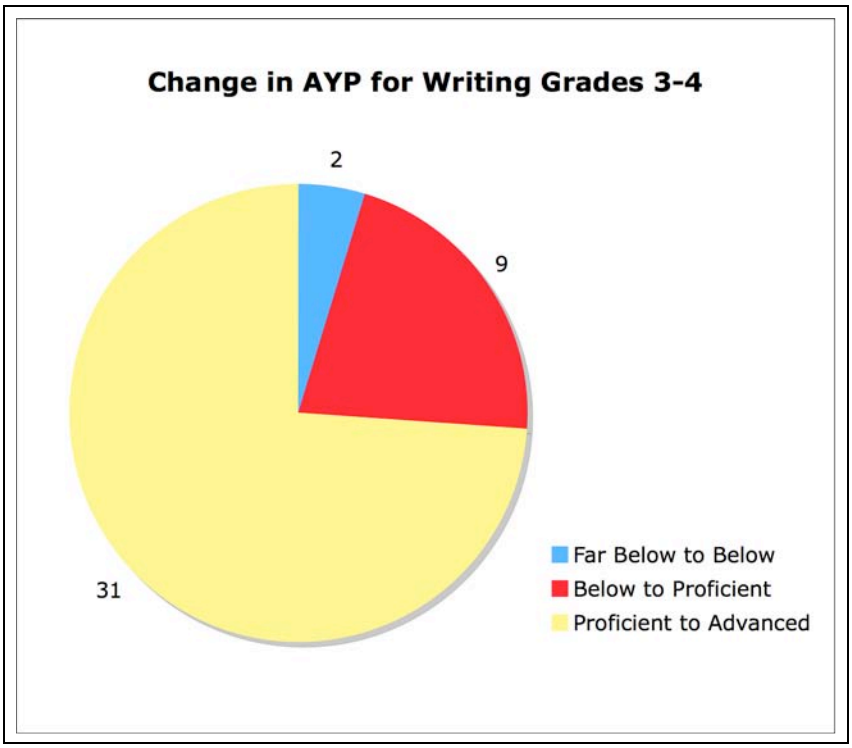


Figure 16. Writing Grades 3-4: Number of Students Changing Adjacent Performance Categories

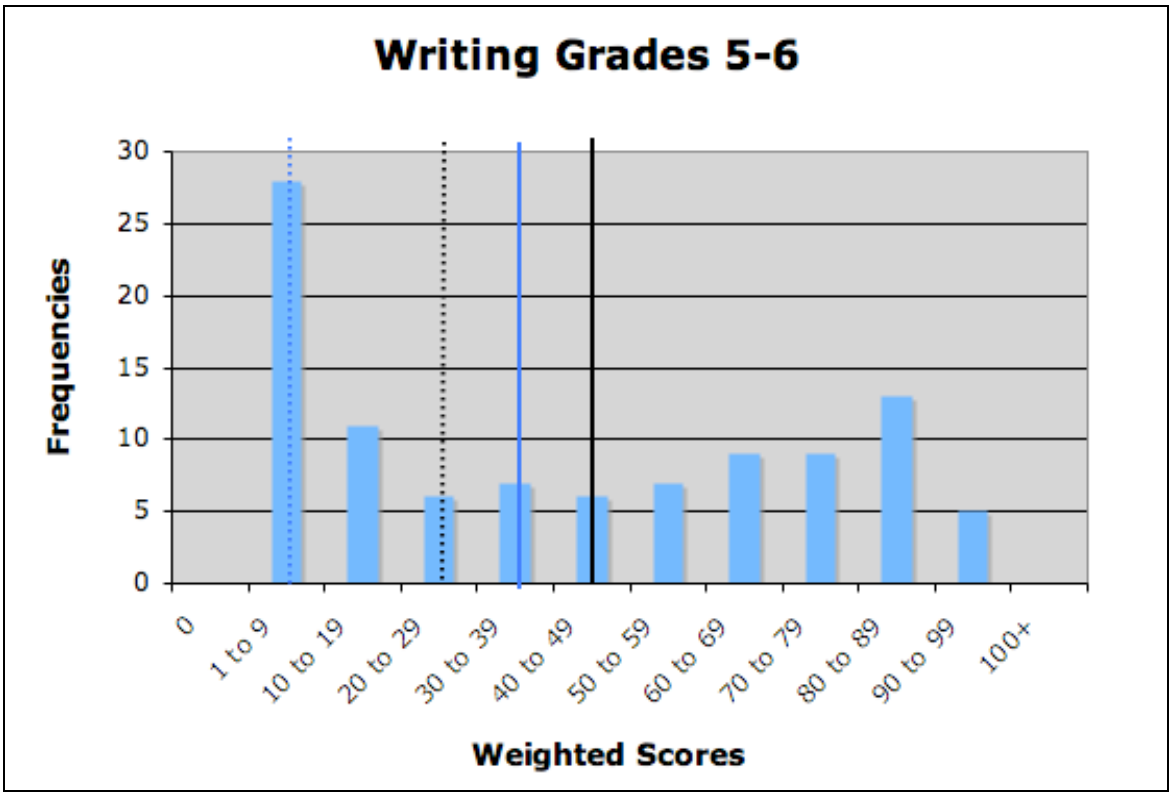


Figure 17. Writing Grades 5-6: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

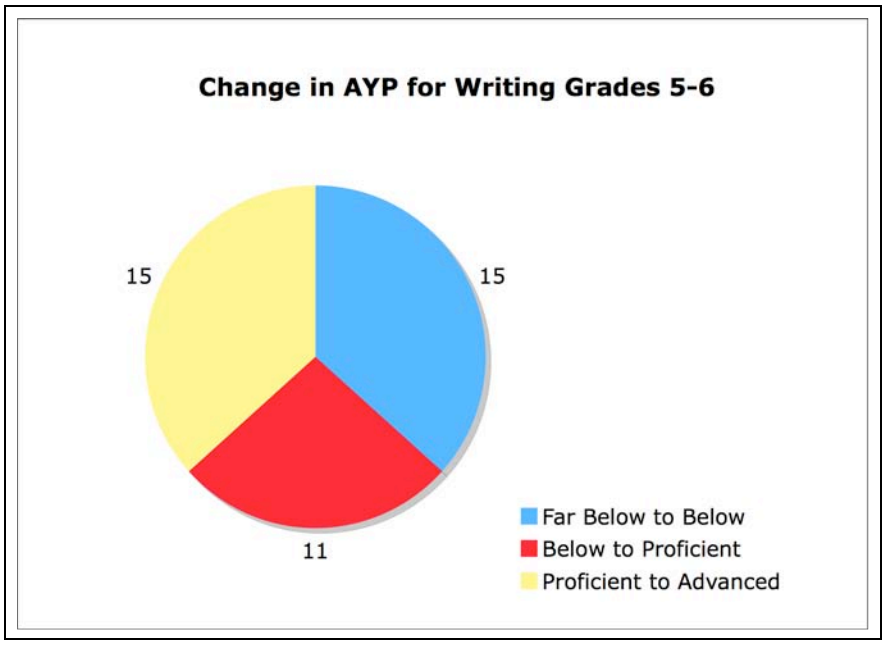


Figure 18. Writing Grades 5-6: Number of Students Changing Adjacent Performance Categories

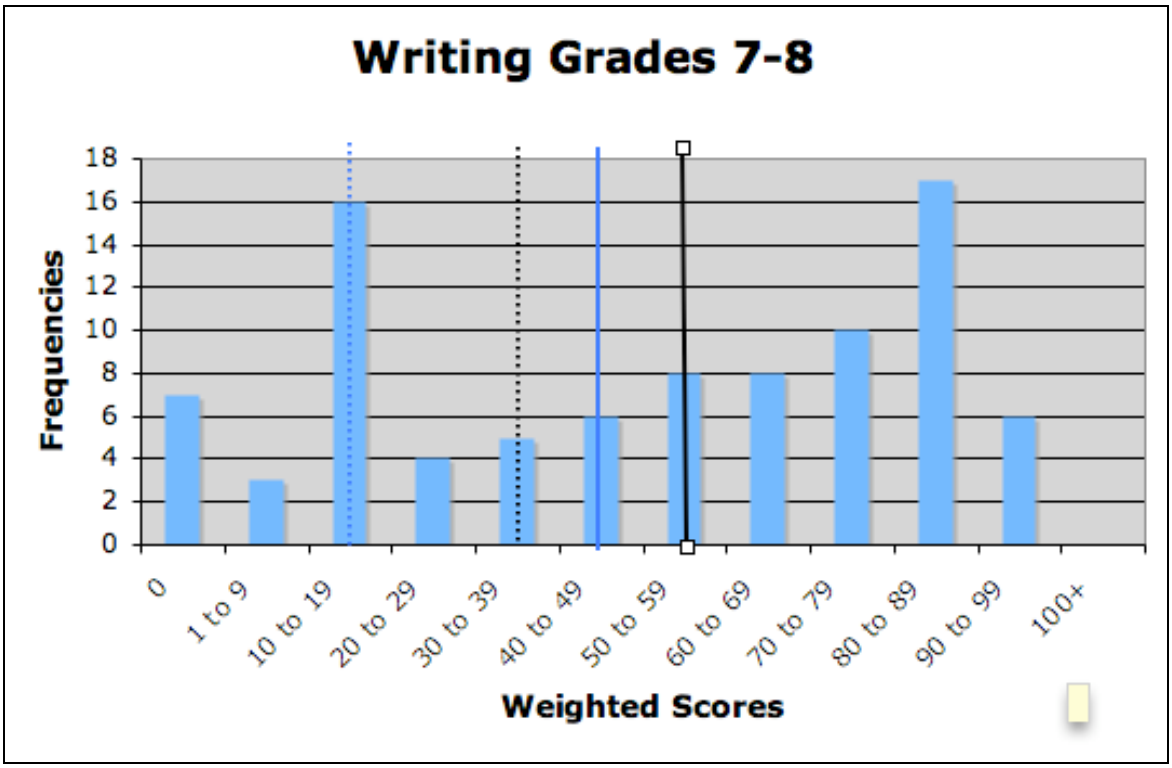


Figure 19. Writing Grades 7-8: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

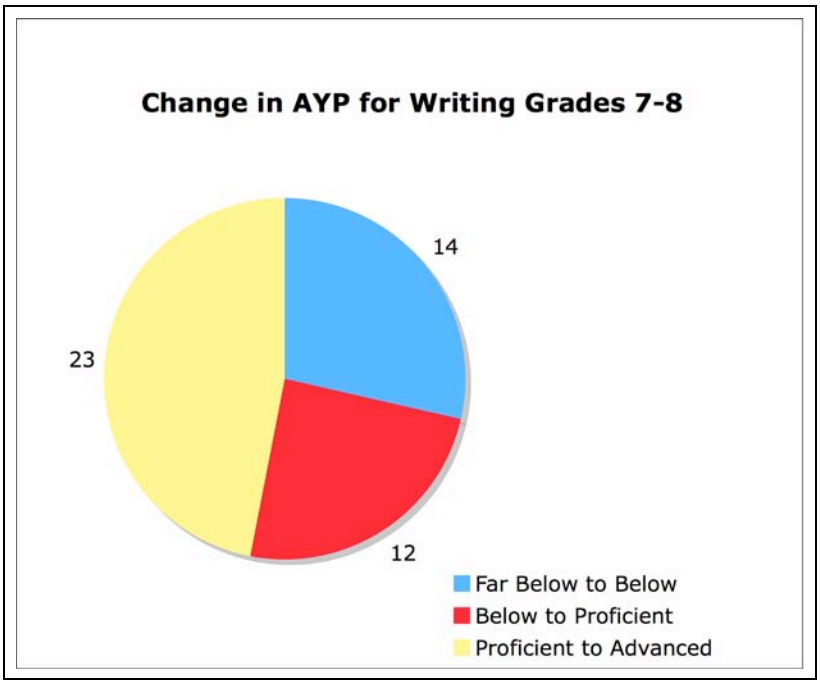


Figure 20. Writing Grades 7-8: Number of Students Changing Adjacent Performance Categories

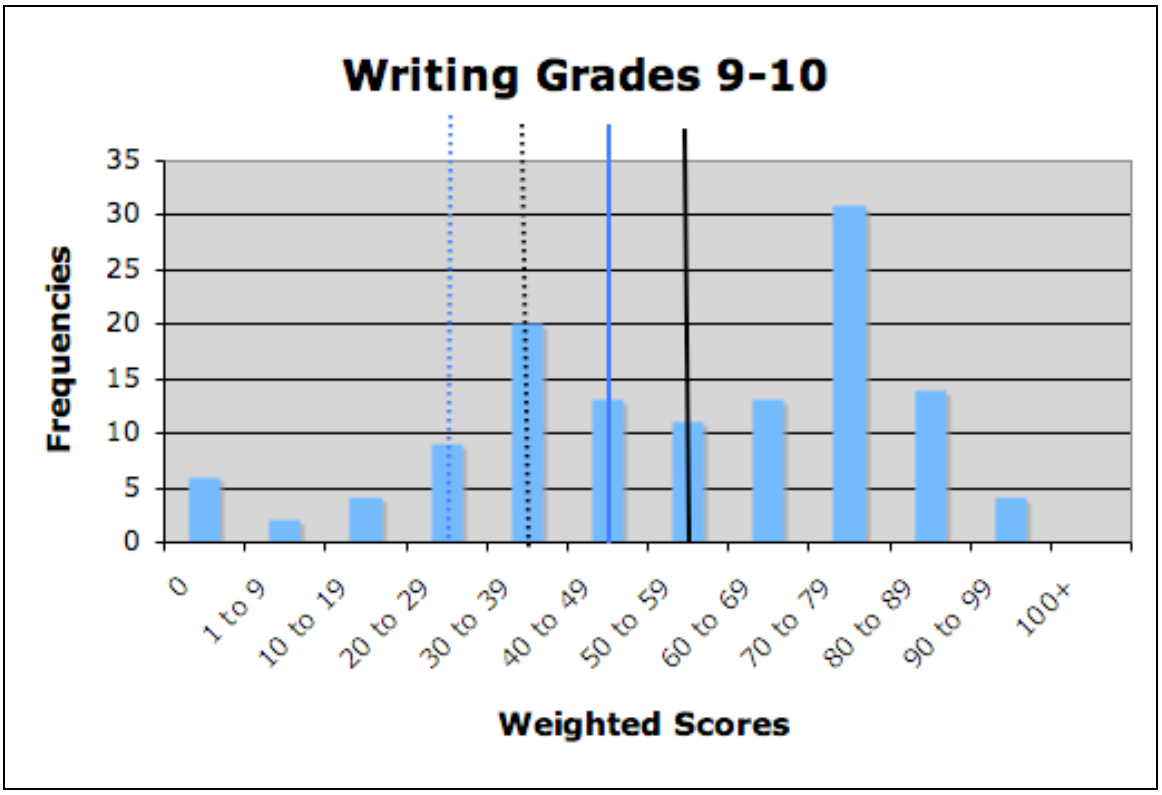


Figure 21. Writing Grades 9-10: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

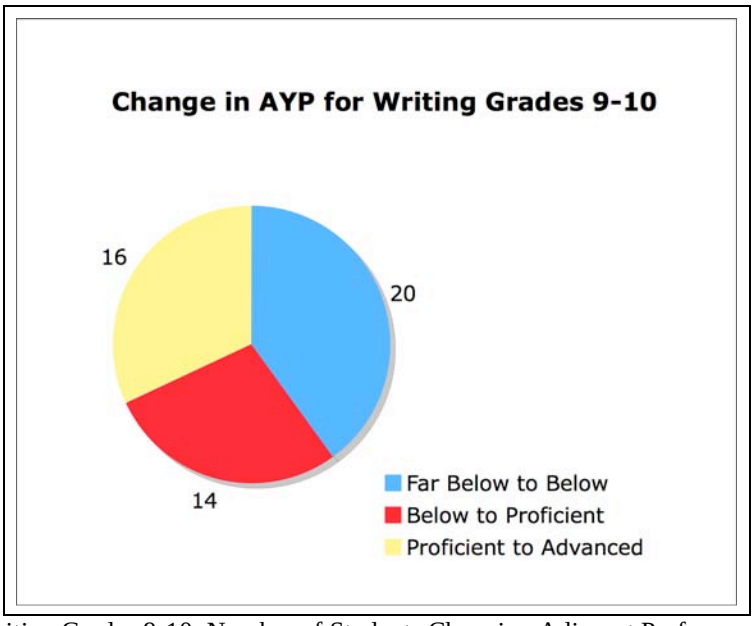


Figure 22. Writing Grades 9-10: Number of Students Changing Adjacent Performance Categories

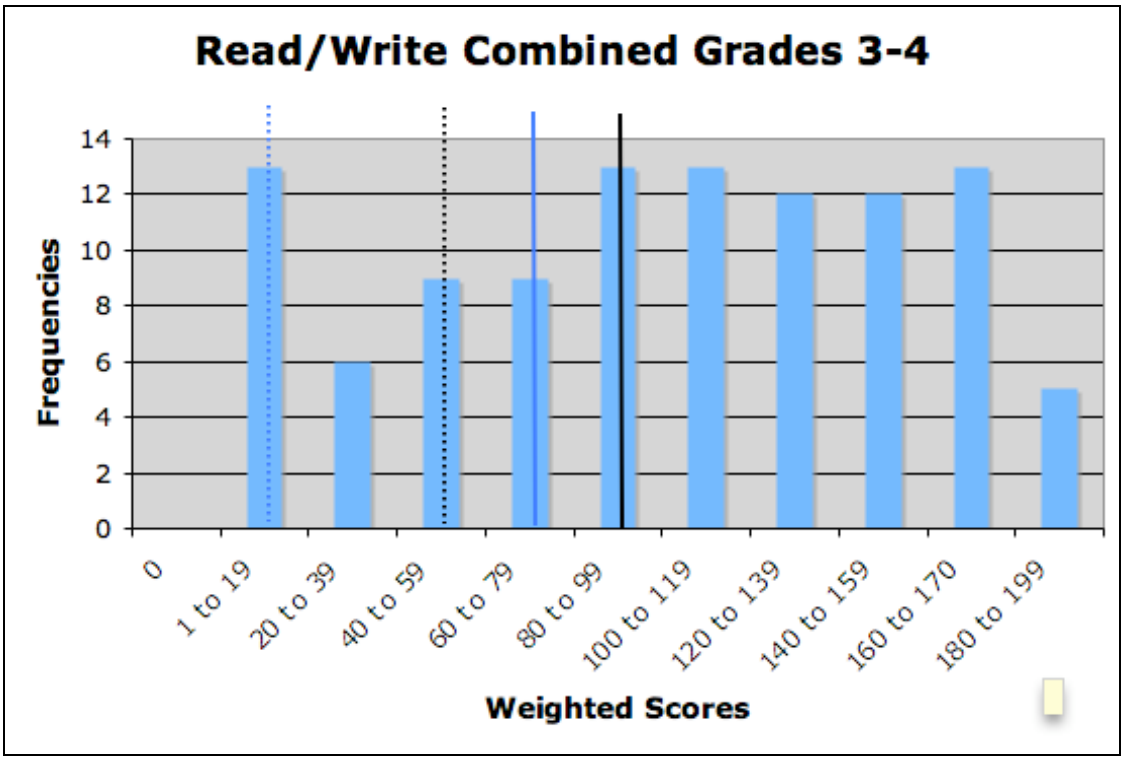


Figure 23. Reading/Writing Grades 3-4: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

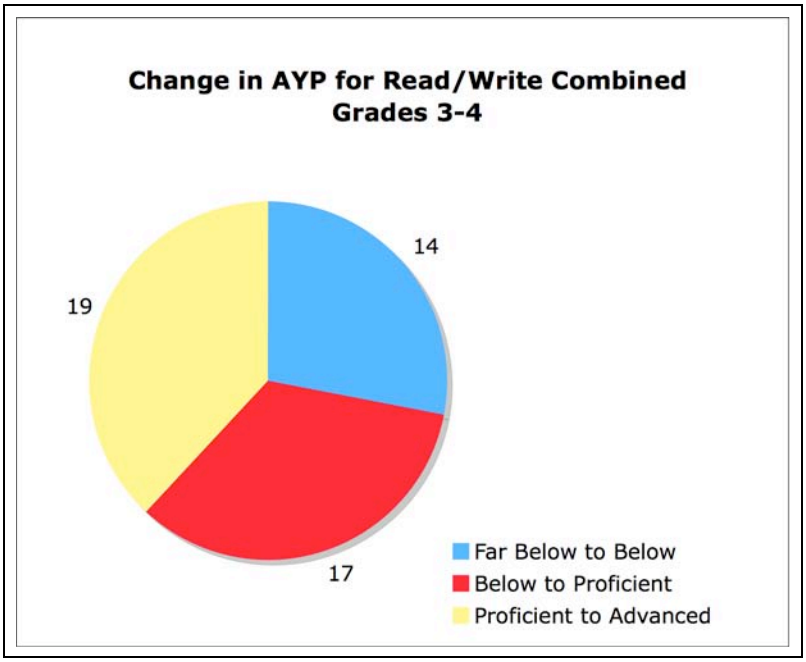


Figure 24. Reading/Writing Grades 3-4: Number of Students Changing Adjacent Performance Categories

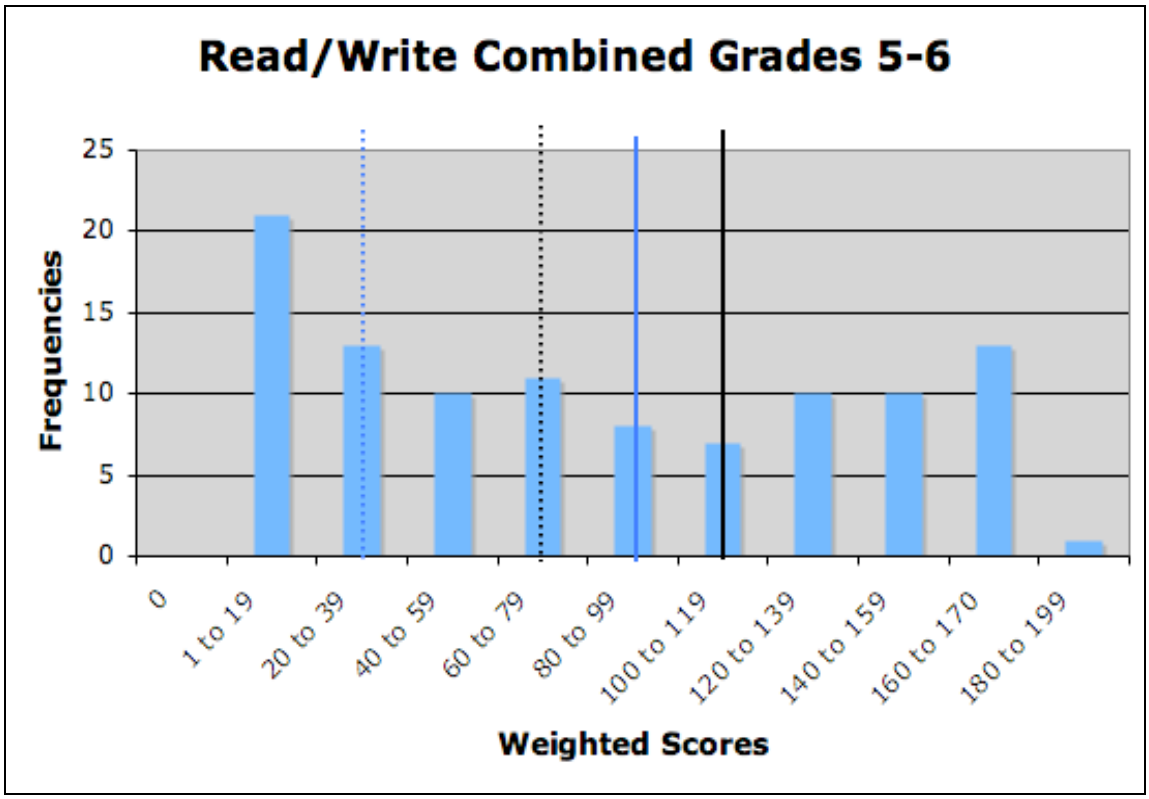


Figure 25. Reading/Writing Grades 5-6: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

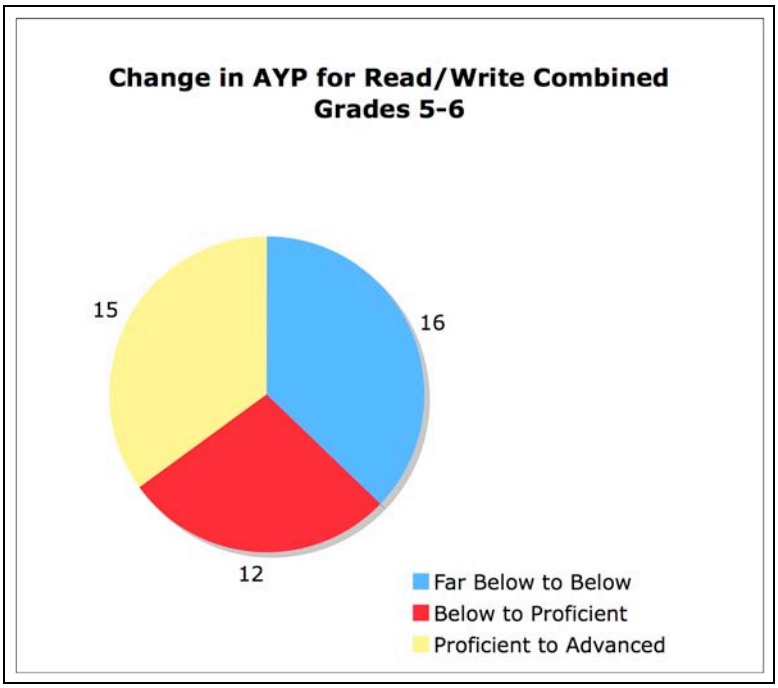


Figure 26. Reading/Writing Grades 5-6: Number of Students Changing Adjacent Performance Categories

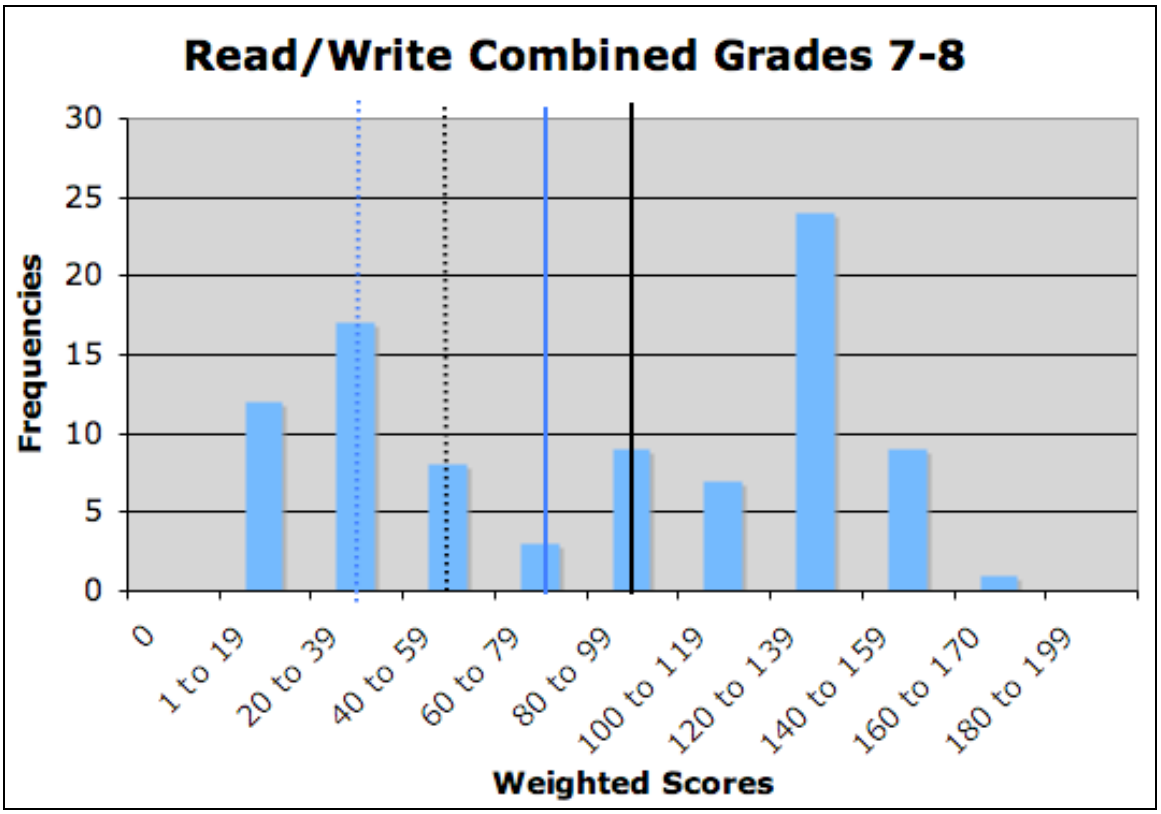


Figure 27. Reading/Writing Grades 7-8: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

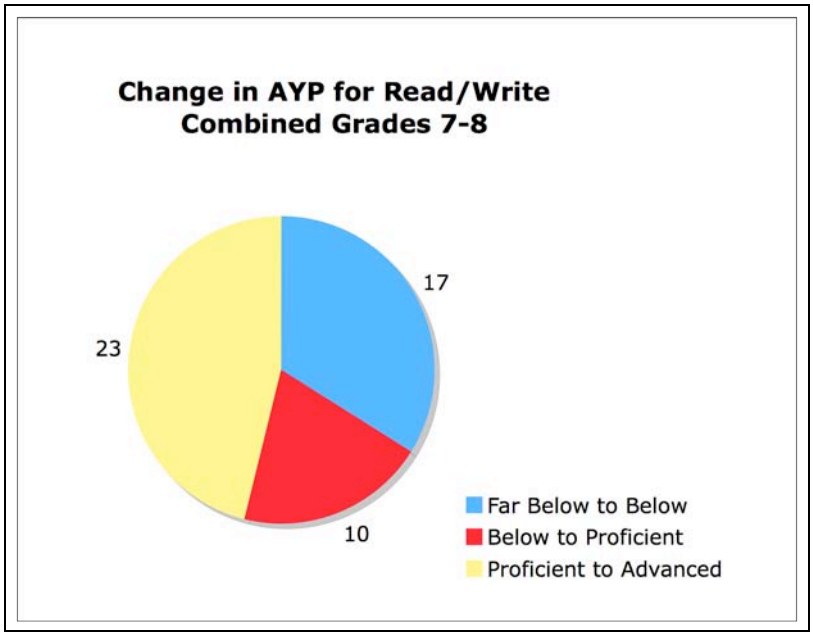


Figure 28. Reading/Writing Grades 7-8: Number of Students Changing Adjacent Performance Categories

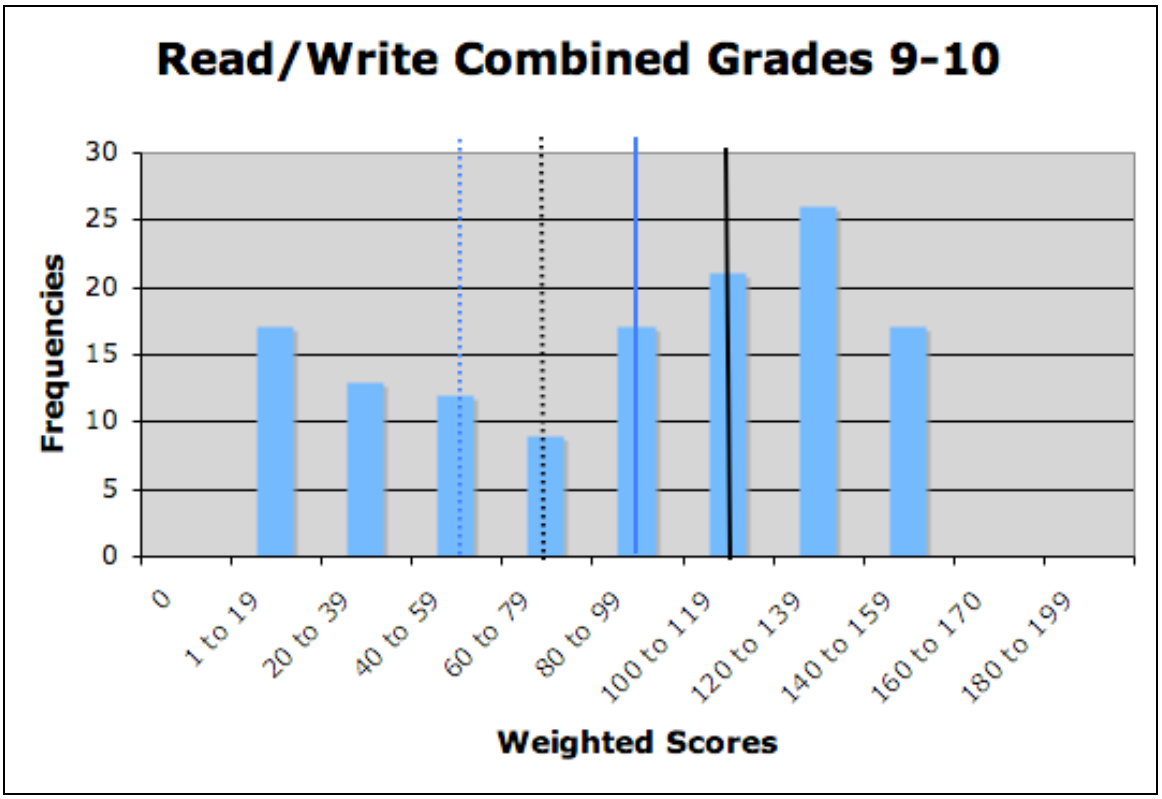


Figure 29. Reading/Writing Grades 9-10: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

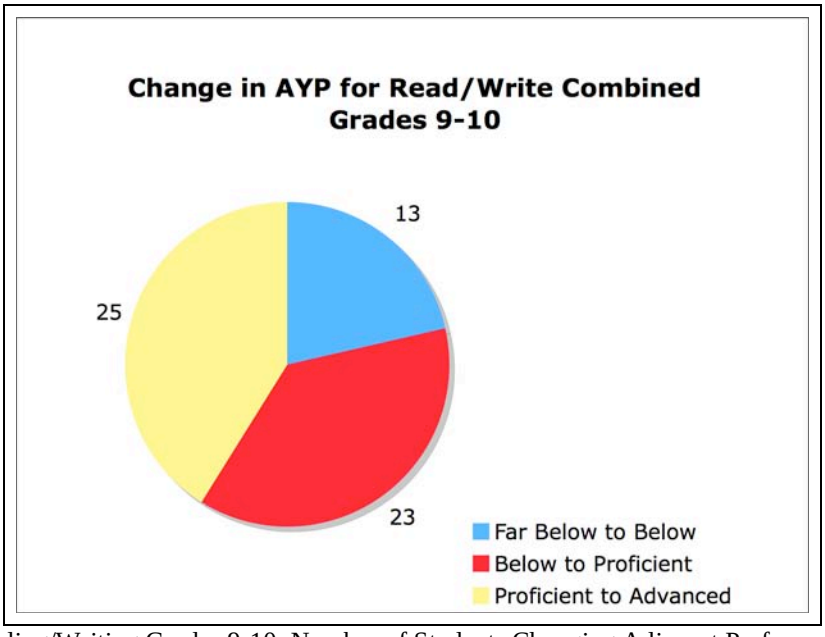


Figure 30. Reading/Writing Grades 9-10: Number of Students Changing Adjacent Performance Categories

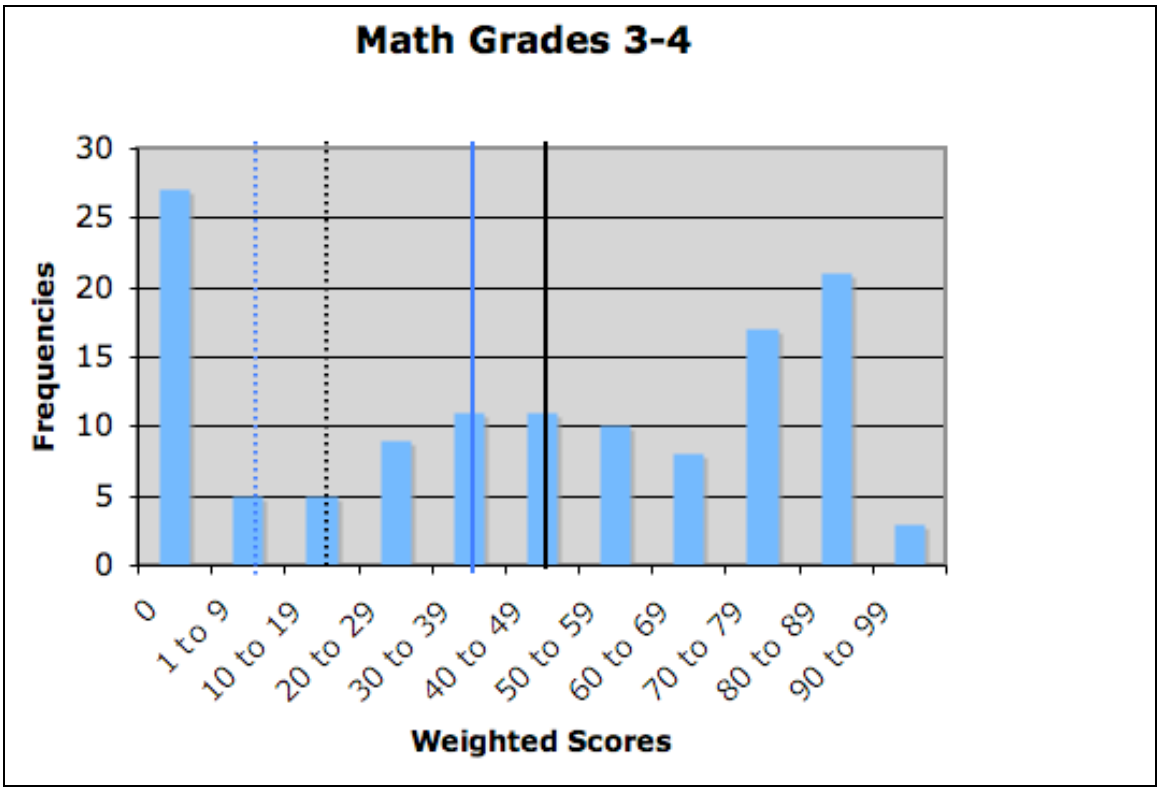


Figure 31. Math Grades 3-4: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

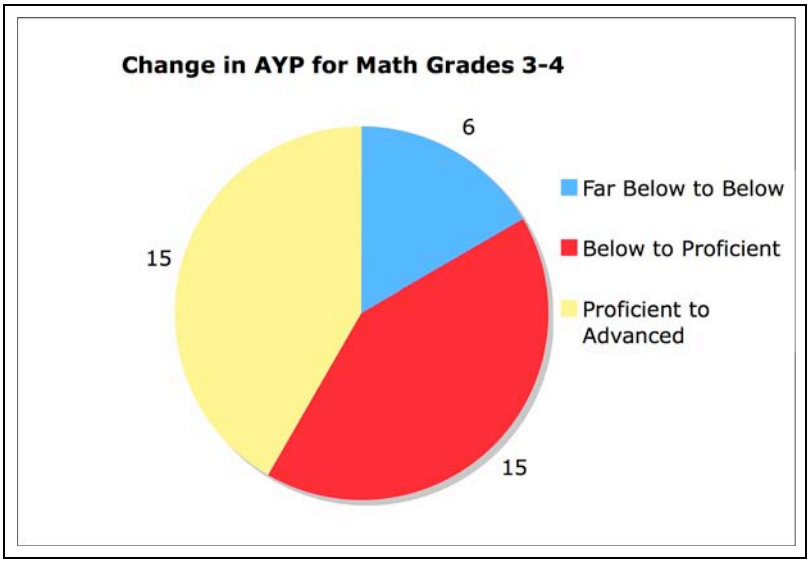


Figure 32. Math Grades 3-4: Number of Students Changing Adjacent Performance Categories

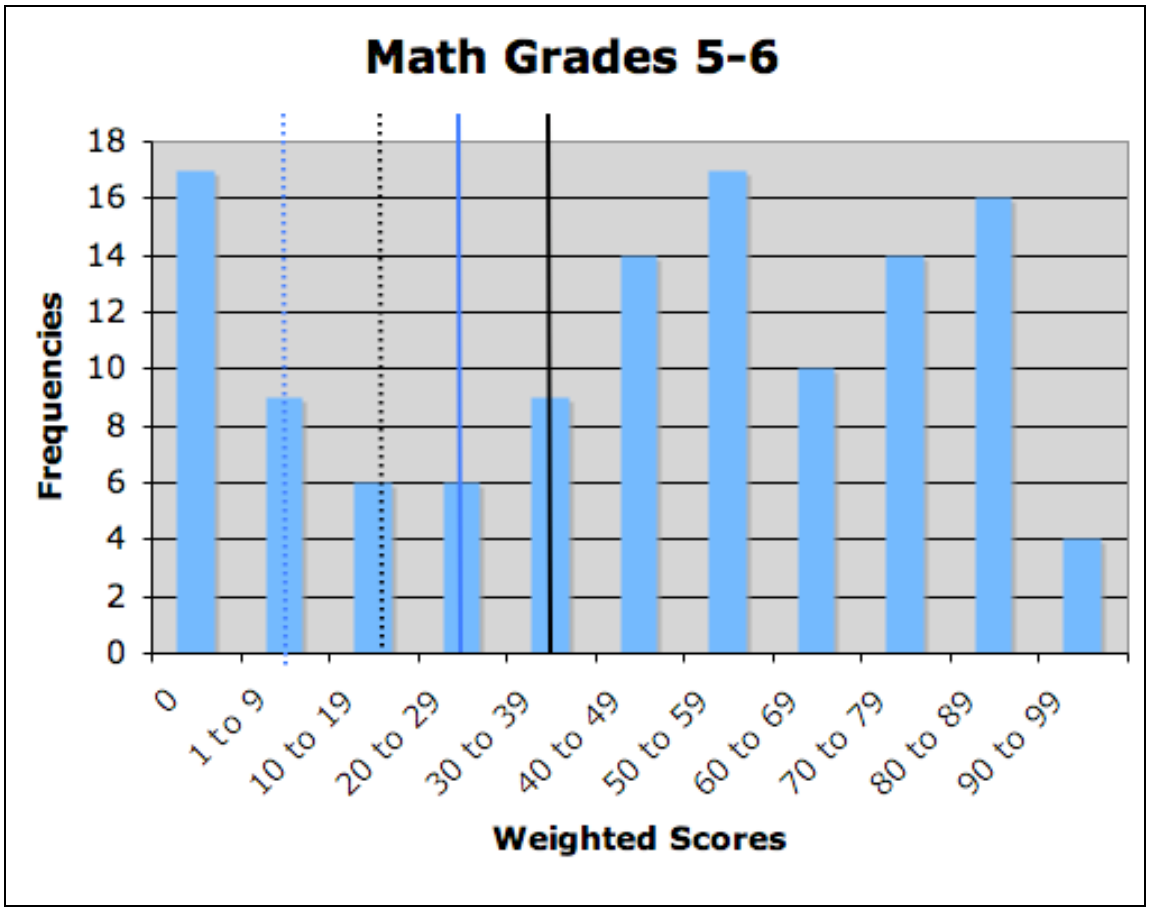


Figure 33. Math Grades 5-6: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

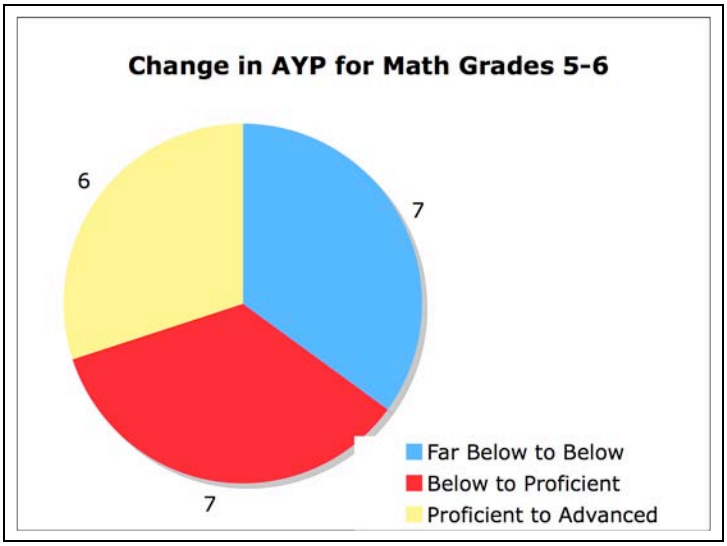


Figure 34. Math Grades 5-6: Number of Students Changing Adjacent Performance Categories

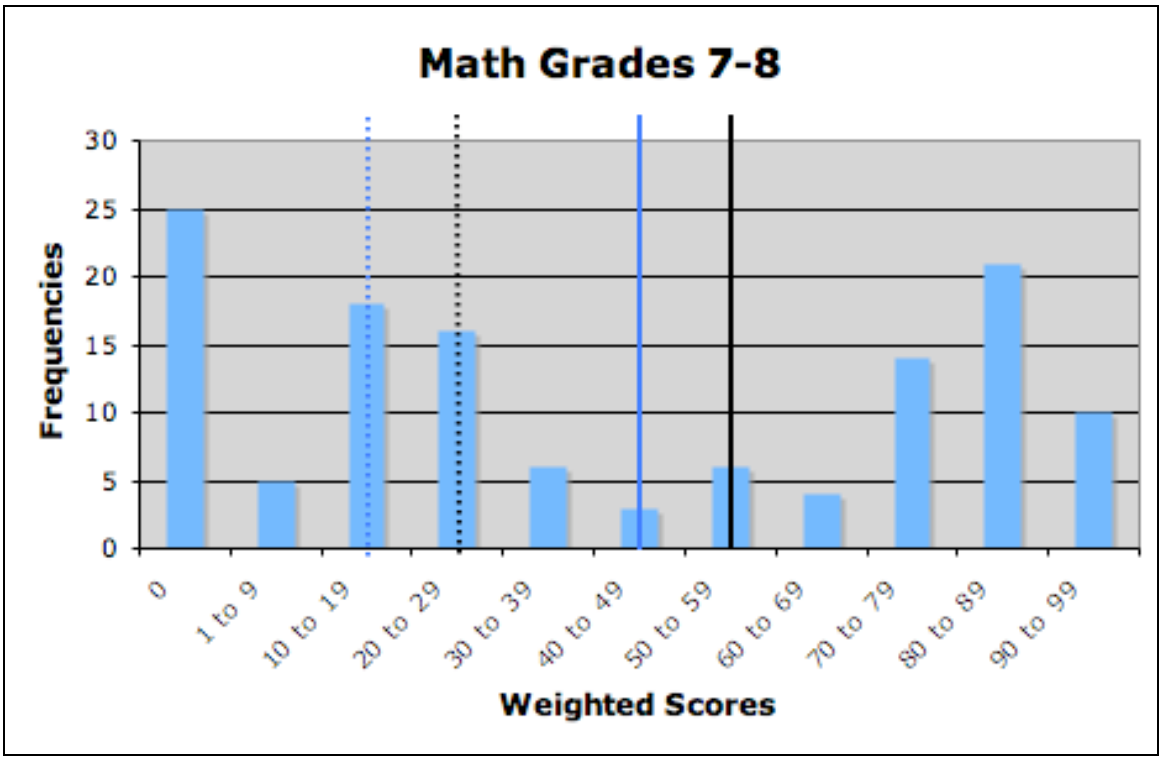


Figure 35. Math Grades 7-8: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

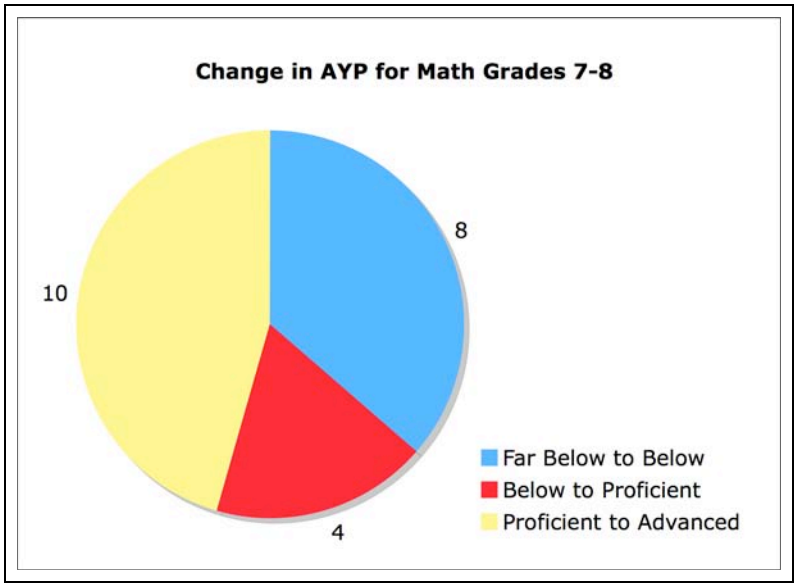


Figure 36. Math Grades 7-8: Number of Students Changing Adjacent Performance Categories

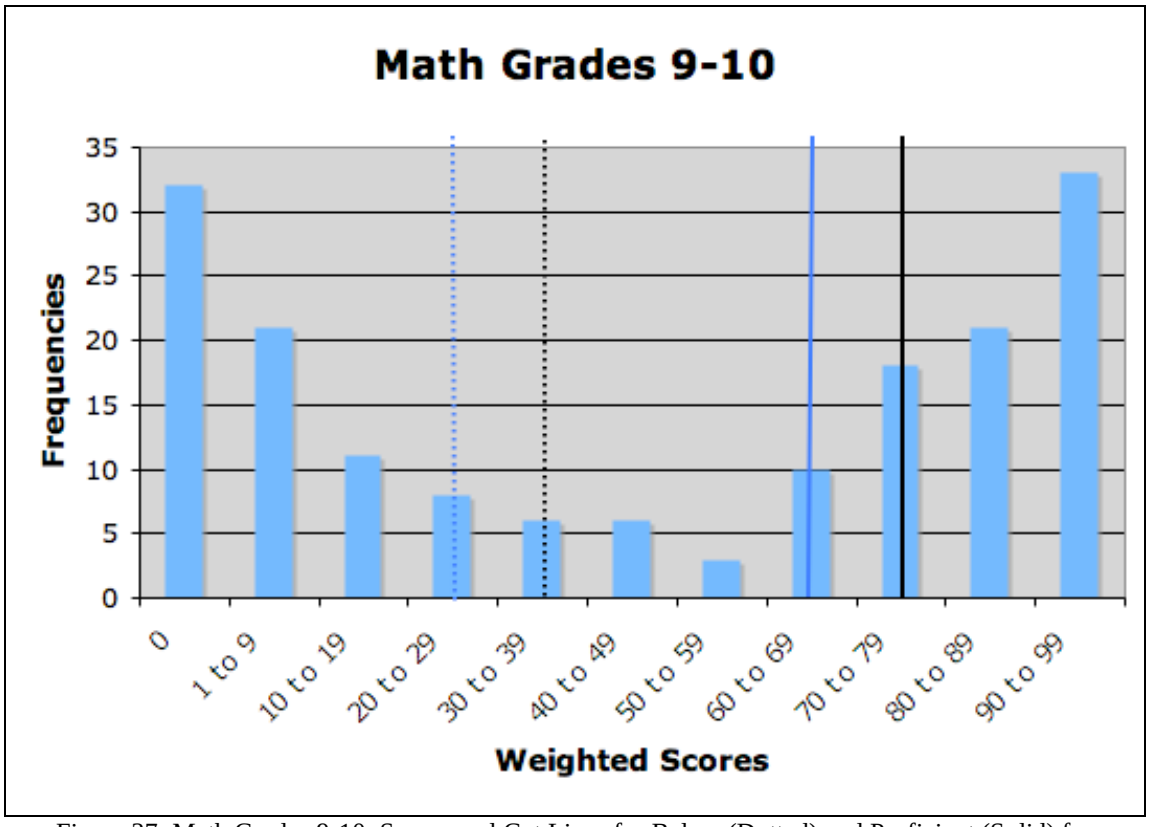


Figure 37. Math Grades 9-10: Scores and Cut Lines for Below (Dotted) and Proficient (Solid) for Weighted (Black) and Weighted -1 SEM (Gray)

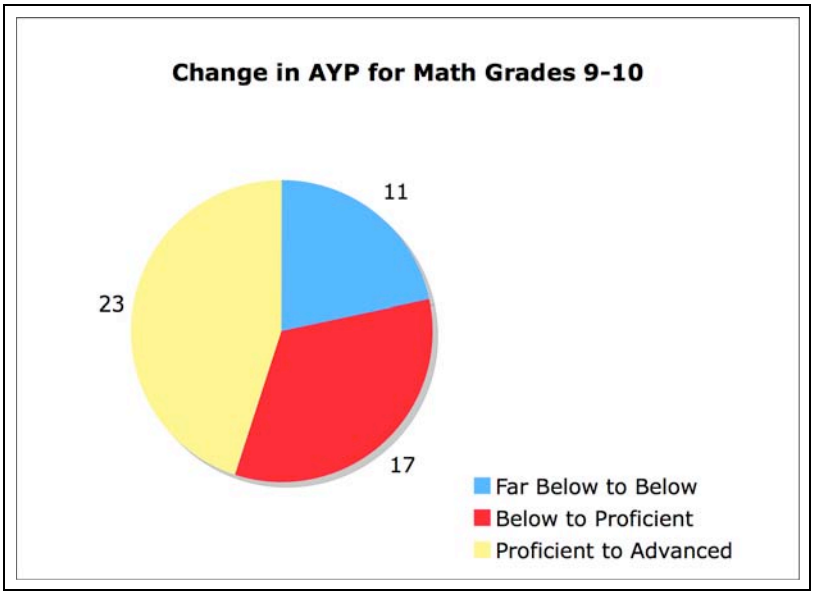


Figure 38. Math Grades 9-10: Number of Students Changing Adjacent Performance Categories

ATTRIBUTE/STRAND VALIDITY STUDY IN READING

Table 52. Grade 3-4 Reading Points and Weights

Strand-Attribute	Task Name	Points	Strand Equal	Strand Value	Strand Tot
Word Identification	1-Pictures	16			
	2-Signs	16			
	6-Sounds	8			
	7-Blends	25			
	Total	65	0.8		50
Form Gen Understand	4-Comp	12			
	Total	12	4.2		50
Total		77			100

Table 53. Grade 5-6 Reading Points and Weights

Strand-Attribute	Task Name	Points	Strand Equal	Strand Value	Strand Tot
Word Identification	8-Bgn words	16			
	9-Adv Words	16			
	10-Sentences	17			
	Total	49	0.7		33
Form Gen Understand	4-Comp	12			
	Total	12	2.8		33
Analysis of Cnt/Struc	4-Listen Comp Item 1*				
	12-Comp - 2 Passages (E & D)-1 item each	24			
	Total	24	1.4		33
Total		85			100

Table 54. Grade 7-8 Reading Points and Weights

Strand-Attribute	Task Name	Points	Strand Equal	Strand Value	Strand Tot
Word Identification	9-Adv Words	16			
	10-Sentences	17			
	11-Rdg (SP)	10			
	Total	43	1.2		50
Form Gen Understand	No ExGLEs				
Analysis of Cnt/Struc	4-Listen Comp	12			
	12-Comp - 2 Passages (E & D)	24			
	Total	36	1.4		50
Total		79			100

Table 55. Grade 9-10 Reading Points and Weights

Strand-Attribute	Task Name	Points	Strand Equal	Strand Value	Strand Tot
Word Identification					
	9-Adv Words	16			
	10-Sentences	17			
	11-Rdg (SP, L1, L2)	30			
	Total	63	0.8		50
Form Gen Understand	No ExGLEs				
Analysis of Cnt/Struc	4-Listen Comp-2 pass	12			
	12-Comp - 2 Passages (E & D)	24			
	Total	36	1.4		50
Total		99			100

Word Identification

Table 56. Descriptive Statistics: Grade 3-4

	N	Minimum	Maximum	Mean	Std. Deviation
XR Word Ident Total	128	.00	49.00	17.9922	18.42435
XR Word Ident Total Weighted	128	.00	39.20	14.3938	14.73948
Valid N (listwise)	128				

Table 57. Descriptive Statistics: Grade 5-6

	N	Minimum	Maximum	Mean	Std. Deviation
XR Word Ident Total	123	.00	49.00	24.1138	19.32602
XR Word Ident Total Weighted	123	.00	34.30	16.8797	13.52822
Valid N (listwise)	123				

Table 58. Descriptive Statistics: Grade 7-8

	N	Minimum	Maximum	Mean	Std. Deviation
XR Word Ident Total	60	1.00	44.00	31.1833	10.00592
XR Word Ident Total Weighted	60	1.20	52.80	37.4200	12.00710
Valid N (listwise)	60				

Table 59. Descriptive Statistics: Grade 9-10

	N	Minimum	Maximum	Mean	Std. Deviation
XR Word Ident Total	70	3.00	66.00	49.3571	14.00004
XR Word Ident Total Weighted	70	2.40	52.80	39.4857	11.20003
Valid N (listwise)	70				

Summary

Students improve in each attribute/strand across successive grade levels. The trend over the grades in positively accelerated with the raw scores (unweighted) and negatively accelerated with the weighted scores.

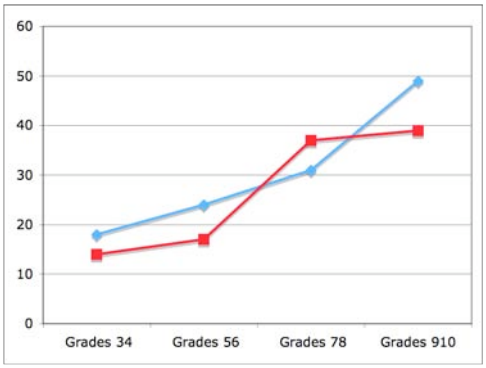


Figure 39. Change over Grades in Reading: Word Identification

Framing a General Understanding

Table 60. Descriptive Statistics: Grade 3-4

	N	Minimum	Maximum	Mean	Std. Deviation
Forming General Understanding	128	.00	12.00	2.7500	3.80220
Forming General Understanding Weighted	128	.00	50.40	11.5500	15.96922
Valid N (listwise)	128				

Table 61. Descriptive Statistics: Grade 5-6

	N	Minimum	Maximum	Mean	Std. Deviation
Forming General Understanding	123	.00	12.00	4.4390	4.22938
Forming General Understanding Weighted	123	.00	33.60	12.4293	11.84226
Valid N (listwise)	123				

Table 62. Descriptive Statistics(a): Grade 9-10

	N	Minimum	Maximum	Mean	Std. Deviation
Analysis of Content-Structure	87	.00	36.00	26.6552	8.35523
Analysis of Content-Structure Weighted	87	.00	50.40	37.3172	11.69733
Valid N (listwise)	87				

a Grade Band = 910

Summary

Students improve in each attribute/strand across successive grade levels. A nearly equivalent (small) amount of change occurs in the elementary grade bands with an equally (large) change in the last grade band.

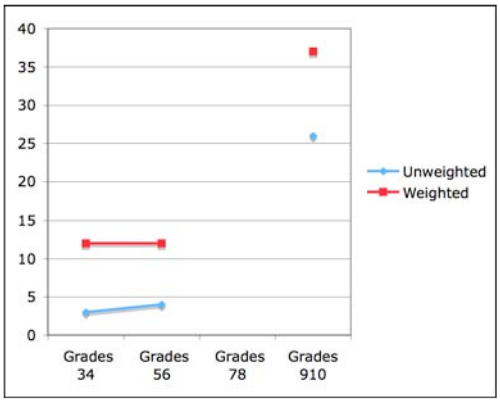


Figure 40. Change over Grades in Reading: Frame A General Understanding

Analyzing Content and Structures

Table 63. Descriptive Statistics(a): Grade 5-6

	N	Minimum	Maximum	Mean	Std. Deviation
Analysis of Content-Structure	48	.00	36.00	21.3542	10.30490
Analysis of Content-Structure Weighted	48	.00	50.40	29.8958	14.42687
Valid N (listwise)	48				

a Grade Band = 56

Table 64. Descriptive Statistics(a): Grade 7-8

	N	Minimum	Maximum	Mean	Std. Deviation
Analysis of Content-Structure	55	.00	36.00	23.0182	11.10971
Analysis of Content-Structure Weighted	55	.00	50.40	32.2255	15.55359
Valid N (listwise)	55				

a Grade Band = 78

Summary

Students improve in each attribute/strand across successive grade levels nearly the same with the weighted and the unweighted in the two grade levels.

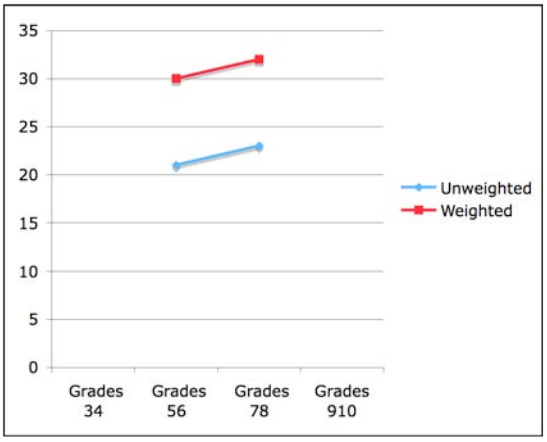


Figure 41. Change over Grades in Reading: Analyzing Content/Structure

SCALING ANALYSIS (INTERNAL STRUCTURES)

Using a 1-parameter item response theory analysis of the data, all items in reading reflected appropriate (item and person) fit statistics (near 1.0). Reliability coefficients were uniformly high. Differential item functioning was found for several items across various grade bands.

Table 65. Reading IRT and IPL Item Analysis Summaries by Task

Task	# of items	Item Fit Statistics			Person Fit Statistics			
		Mean	Min.	Max.	Reliability	Mean	Min.	Max.
1	8	1.01	.58	1.34	.88	.99	.13	9.9
2	8	1.09	.82	1.63	.81	1.03	.12	9.9
3	8	1.05	.72	1.53	.88	1.05	.35	3.27
4	6	.98	.66	1.46	.94	.98	.18	4.33
5	8	1.02	.70	1.73	.90	1.02	.11	8.22
6	8	.93	.53	1.34	.91	.94	.23	4.76
7	8	.93	.53	1.34	.91	.94	.23	4.76
8	8	1.01	.79	1.27	.93	1.01	.24	3.99
9	8	1.04	.74	1.75	.91	1.04	.08	3.25
10	5	1.53	.72	3.82	.91	1.25	.07	9.90
11	9	NA	NA	NA	NA	NA	NA	NA
12	6	1.0	.74	1.16	.92	1.01	.14	6.82
13	6	.99	.75	1.18	.93	.99	.27	5.45

Table 66. Reading Differential Item Functioning by Grade Level

Task	# of Items	Items with Significant Difference Across Grade Groups (sign indicates sign of *t*-score)					
		34 & 56	34 & 78	34 & 910	56 & 78	56 & 910	78 & 910
1	8		-2, 8		-2		2
2	8			4, -5		-5	
3	8		-4	-1			-1
4	6	1, -6		1, -2			1, -2
5	8						
6	8						
7	8						
8	8			6	5		
9	8	-8		7			
10	5	5	-2, 5	-3, 5			
11	9	NA	NA	NA	NA	NA	NA
12	6					1, 4	
13	6					-4	

Table 67. Writing IRT and IPL Item Analysis Summaries by Task

Using a 1-parameter item response theory analysis of the data, all items in reading reflected appropriate (item and person) fit statistics (near 1.0). Reliability coefficients were uniformly high. Differential item functioning was found for several items across various grade bands.

Task	# of items	Item Fit Statistics			Person Fit Statistics			
		Mean	Min.	Max.	Reliability	Mean	Min.	Max.
1	10	.99	.56	1.91	.93	.99	.11	2.7
2	8	.98	.85	1.08	.91	.98	.06	3.43
5	10	1.01	.71	1.84	.90	1.01	.12	4.78
7	8	.99	.72	1.29	.86	.99	.51	2.69

Table 68. Writing Differential Item Functioning by Grade Level

Task	# of Items	Items with Significant Difference Across Grade Groups (sign indicates sign of *t*-score)					
		34 & 56	34 & 78	34 & 910	56 & 78	56 & 910	78 & 910
1	10	7			1		
2	8			7	4, 8	6	4
5	10	-10	2		2		-4, 6
7	8	-5		2,6		6	2

Table 69. Math IRT and IPL Item Analysis Summaries by Task

Using a 1-parameter item response theory analysis of the data, all items in reading reflected appropriate (item and person) fit statistics (near 1.0). Reliability coefficients were uniformly high. Differential item functioning was found for several items across various grade bands.

Task	# of items	Item Fit Statistics			Person Fit Statistics			
		Mean	Min.	Max.	Reliability	Mean	Min.	Max.
1	8	1.01	.54	1.74	.89	1.04	.34	6.98
2	8	1.31	.35	2.70	.83	.95	.03	9.90
3	8	1.27	.76	2.69	.78	1.08	.08	9.90
4	10	.70	.40	1.02	.90	.72	.06	8.12
5	4	1.02	.84	1.17	.89	1.02	.4	3.32
6	4	1.06	.74	1.76	.90	1.06	.23	8.13
7	9	.99	.36	4.62	.79	.99	.04	2.93
8	3	.99	.88	1.12	.90	.99	.84	1.17
9	5	1.0	.69	1.47	.88	1.0	.14	3.64
10	4	1.55	1.09	2.56	.87	.98	.02	9.90
11	1	NA	NA	NA	NA	NA	NA	NA
12	2	1.0	1.0	1.0	.00	1.0	.82	1.22
13	5	.90	.71	1.07	.84	.89	.06	9.9
14	7	.98	.60	1.36	.83	.8	.09	9.9
15	4	1.11	.77	1.51	.74	1.03	.13	9.9
16	5	1.13	.57	1.81	.86	1.01	.06	9.9
17	5	1.13	.65	2.92	.94	1.13	.25	5.34
18	4	1.0	.89	1.22	.87	1.0	.31	3.19
19	3	1.05	.52	1.57	.00	.48	.00	6.28
20	2	NA	NA	NA	NA	NA	NA	NA
21	2	NA	NA	NA	NA	NA	NA	NA
22	1	NA	NA	NA	NA	NA	NA	NA

Table 70.Math Differential Item Functioning by Grade Level

Task	# of Items	Items with Significant Difference Across Grade Groups (sign indicates sign of *t*-score)					
		34 & 56	34 & 78	34 & 910	56 & 78	56 & 910	78 & 910
1	8	1				-5	-5, 7
2	8	-2, 7	2	-3, -5, 7		-3	
3	8		-2		-2		2
4	10			-2, 4, -7, 8		4	4
5	4		-3		-3		
6	4						
7	9		-2,9	-2, 9		9	
8	3						
9	5			4			
10	4						
11	1	NA	NA	NA	NA	NA	NA
12	2						
13	5						
14	7			-2			
15	4	2					
16	5			4			
17	5	-5	-5	-5			
18	4						
19	3	1					