

**Alaska Alternate Assessment
2009 - 2010
Technical Report**

Dillard Research Associates
August 23, 2010

Table of Contents – 2010 AK Alternate Assessment Technical Report

CHAPTER 1: SUMMARY OF TASK DEVELOPMENT AND PREVIOUS (2009) TECHNICAL REPORT	1
<i>Overview</i>	1
<i>Historical Information</i>	1
<i>Summary of New Items</i>	1
<i>Bias and Sensitivity Review</i>	2
<i>Reduction in Complexity, Depth, and Breadth</i>	2
<i>Item Development and Expanded Levels of Support (ELOS)</i>	3
<i>Description of Directory of Test Specifications (DOTS)</i>	3
CHAPTER 2: 2009 TRAINING	4
<i>Improvements to Web-based Training System)</i>	4
<i>New Mentor Training</i>	6
<i>Summary of Dates and Participants</i>	6
<i>Scoring Reliability Analysis</i>	7
<i>Online Proficiency</i>	7
<i>Scoring Reliability Analysis, Proficiency Test Attempts</i>	7
<i>Annual Mentor Training</i>	8
<i>Summary of Dates and Participants</i>	8
<i>Scoring Reliability Analysis, New Mentor and All Mentor Training</i>	8
<i>Scoring Reliability Analysis, Science Score-Behind Study</i>	8
<i>Webinar</i>	9
CHAPTER 3: TEST IMPLIMENTATION IN 2009-2010	10
<i>Updates on Procedures</i>	10
<i>Corrected errors on alternate assessment documents</i>	10
<i>Developed more detailed quality assurance schedule</i>	10
<i>Help Desk Summary and Frequently Asked Questions Summary</i>	11
<i>Quarterly Reports</i>	11
CHAPTER 4: CONTINUOUS IMPROVEMENT	12
<i>Review of Replacement Process for Operational and Field Test Items</i>	12
<i>Form C Test Development</i>	12
<i>Proficiency Testing and Practice Tests</i>	12
<i>Refresher Tests</i>	13
<i>Mentor Oversight and Database Development</i>	13
<i>Summary of Consequential Survey</i>	13
<i>Technological Improvements</i>	14
CHAPTER 5: INTERPRETING STATISTICAL ANALYSIS	15

CHAPTER 6: RELIABILITY	16
<i>Total Test</i>	16
<i>Strand Reliability</i>	17
<i>Task Reliability</i>	18
<i>Split Half Reliability</i>	19-21
<i>Participation in Extended Levels of Support (ELOS)</i>	22-27
CHAPTER 7: DESCRIPTIVE STATISTICS	28
CHAPTER 8: ADEQUATE YEARLY PROGRESS	29
<i>Reading</i>	29
<i>Writing</i>	29
<i>Mathematics</i>	29
<i>Science</i>	29

APPENDICES

<i>Appendix 2.1 – New Mentor Training Dates and Participants</i>	6
<i>Appendix 2.2 – New Mentor Training Scoring Reliability Analysis</i>	7
<i>Appendix 2.3 – Annual Mentor Training Dates and Participants</i>	7
<i>Appendix 2.4 – Annual Mentor Training Dates and Participants</i>	8
<i>Appendix 2.5 – Annual Mentor Training Accommodations and Focus Group</i>	8
<i>Appendix 2.6 – New Mentor and All Mentor Training Scoring Reliability Analysis</i>	8
<i>Appendix 2.7 – Science Score-Behind Study</i>	9
<i>Appendix 2.8 – Webinar</i>	9
<i>Appendix 3.1 – Quality Assurance</i>	10
<i>Appendix 3.2 – Helpdesk and Frequently Asked Questions</i>	11
<i>Appendix 3.3 – Quarterly Reports</i>	11
<i>Appendix 4.1 – Consequential Validity</i>	14
<i>Appendix 4.2 – Technological Improvements</i>	14
<i>Appendix 7.1 – Grade Level Descriptive Statistics</i>	27
<i>Appendix 7.2 - Grade Band Descriptive Statistics</i>	27
<i>Appendix 8 – Adequate Yearly Progress</i>	28

CHAPTER 1: SUMMARY OF TASK DEVELOPMENT AND PREVIOUS (2009) TECHNICAL REPORT

Overview

The 2009–2010 Alaska Alternate Assessment represented an equivalent form test to the 2008–2009 Alternate Assessment. Approximately one-half of the items from the 2008–2009 assessment were paired with equivalent items developed to match the content, construct, and point value of the replaced items. This version of the assessment is referred to as “Form A” and the equivalent items are referred to as “Cousin Items.” Sufficient cousin items were developed between 2008 and 2010 to permit as many as six versions of the Alaska Alternate Assessment. Within the Form A assessment, items that were carried forward from the 2008–2009 assessment are called “operational items,” while matched cousin items are called “field test items.” Operational items have at least two years’ of student performance data analysis on the items from 2007-2008 and 2008-2009 tests analysis.

Historical Information

In 2005, a Reliability and Validity study was conducted by Dr. Gerald Tindal which concluded in a need for revision in the State of Alaska’s Student Portfolio system in order to meet technical quality requirements set by the No Child Left Behind legislation. Following the department’s Request for Proposals, Dillard Research Associates was awarded a contract to secure a standardized performance-task assessment for students with significant disabilities. To provide greater reliability in administration and scoring of the assessment, an online administrator training program was developed. This online training program includes training and proficiency tests for each task. Secure tests were developed in accordance with the State of Alaska’s Extended Grade Level Expectations (ExGLEs). Proficiency Level Descriptors have been created by teams of content experts.

Summary of New Items

During the testing windows in 2005-2006, and 2006-2007 test items consisted of new and field test items. In 2007-2008, the secure tests were organized into grade bands and items were constructed by closely aligning the items to the Extended Grade-Level Expectations (ExGLEs). Crosswalks were created to map the 2006-2007 test items to the 2007-2008 test items for all subject areas. The test design and specifications were applied in four areas of (a) reading, (b) writing, (c) mathematics, and (d) science. Each test includes both teacher *administration and scoring protocols* and *student materials*.

Each test consists of tasks that are comprised of several items. These items closely align to the Extended Grade-Level Expectations (ExGLEs). The Alaska Alternate Assessment may be administered in a standard manner or with extended levels of support. The standard administration of items ascertains students’ knowledge and skills on extended grade level expectations. The extended levels of support administration engages students with pre-emergent or emergent communication skills in items that address early entry points to the ExGLEs. Specific rules related to student performance in the standard administration permit an assessor to

provide extended levels of support for a student. For the 2008-2009 testing window, items and test formatting were retained from the 2007-2008 secure test.

Bias and Sensitivity Review

A bias and sensitivity review of the new test items was conducted in November 2007. During this process, reviewers examined the bias of the assessment and if the format would affect student performance. A group of 12 participants from Alaska and two specialists with the deaf and blind community from Oregon were selected to review all items. All reviewers were given examples to focus on during the review and all held qualified assessor certificates and certification in special education. Items were updated based on the results of this review prior to the 2007-2008 testing window.

A second Bias and Content Committee was convened in September 2009 to analyze cousin items, a pool of new, related items to the existing items in the Alternate Assessment. The results of this analysis are indicated in the Directory of Tests Specifications document.

Reduction in Complexity, Depth, and Breadth

Due to the federal regulations provided in December 2003, steps were taken to increase the cognitive accessibility of items. This was done by analyzing and removing potential barriers for students with significant cognitive disabilities. This process was used in the development of items and for both administration and scoring and student materials. Simplified language was used in all text. Alignment was ensured between teacher scripted language and student materials. General test layout was considered from the view of readability and legibility. Specific administration directions were limited to a single page of the Scoring Protocol for ease of administration. Student materials were organized for ease of administration into ‘cards’ that either could be cut out or masked by the administrator. Pictures were constructed using primarily black and white for minimal complexity. All items were reviewed with administration and development steps toward reducing complexity.

Depth-of-knowledge (DOK) was judged in the analysis of the Alternate Assessment. An alignment study was conducted in 2007 by Karvenon and Almond; the information was used to guide item adaptations for the 2007-2008 secure test items. Categorical concurrence, range of knowledge, and balance of representation were defined originally by Webb, and adapted by Dr. Tindal for use with students with significant cognitive disabilities, and then defined based on operational use within the Alaska Alignment Study.

Items were developed based on a one-to-one correspondence with the Extended Grade-Level Expectations (ExGLEs). All strands and attributes were equally addressed in accordance to proportion of points for each task. The total points for each test was fixed at 100 points to allow proficiency standards from the first year to be comparable to the second year of testing. No weighting was needed and an algorithm was used to equalize the differential points across strands/attributes.

Item Development and Expanded Levels of Support (ELOS)

Items were developed within grade-bands allowing items to be grouped into tasks. Two to eight items were developed for each task with scoring developed for partial or full credit depending on the item. Two types of items were created: Standard items, and Expanded Level of Support (ELOS) items. The ELOS items were created to ensure participation and allow the assessor to ascertain a student's level of independence. These items also allow maximum participation for students with the most significant cognitive disabilities and provide information for assessors on what level of support is necessary for the student to interact with the assessment materials.

Description of Directory of Test Specifications (DOTS)

Descriptions of Test Specifications for the 2009–2010 Alternate Assessment are Excel spreadsheets which define all aspects of each item used in the Form A test materials. In addition to items used in the 2009–2010 assessments, information related to all items used in tests beginning with the 2007–2008 test materials are displayed. Information includes the strand name, the number of answer options, maximum score points, item degree of knowledge, whether the item was an operational or field test item, and statistical data for each item (mean, standard deviation), a statistical analysis of the difficulty of the item (the mean points for each item divided by the maximum points available), and the task weight.

The DOTS documents for reading, writing, mathematics and science contain confidential secure test information and are not available to the public.

CHAPTER 2: 2009 TRAINING

Only school personnel may administer the Alternate Assessment. This includes both teachers and paraprofessionals. In order to become a Qualified Assessor (QA), individuals must participate in online training, pass proficiency tests, and administer a practice assessment that is then reviewed by their Qualified Mentor-Trainer (QT). Each QT must go through this same training, as well as additional in-person training provided annually by the Alaska Department of Education and Early Development, in order to serve as a valuable resource to QAs. These individuals have been appointed by the Special Education Director or Superintendent to be the primary point of contact for the Alternate Assessment Program Manager.

During annual training, all participants are required to sign and return a test security agreement. This document reiterates the message from training: test security is of the utmost importance in obtaining valid and reliable scores. As such, QAs must keep all materials in a confidential location, and refrain from discussing specifics of the test with others. Following the close of the test administration window, all testing materials should be shredded. Teachers cannot access the secure test documents until they have passed the training requirements (passing all proficiency tests and, for assessors-in-training, administration and submission of a practice test). After completion of all requirements, they are granted access to the secure test materials.

Special education teachers who were selected by their districts to serve as Qualified Mentors in the Alaska Alternate Assessment attended a two-day New Mentor training on September 28 and 29, 2009. After Assessors-in-Training completed all training and proficiency tests successfully, they administered a practice test, which was reviewed by DRA. Once the Assessor-in-Training had completed these tasks, his or her account was updated to the status of QA. During training, these participants also scored a protégé's assessment protocols. After passing all these tasks, participants were upgraded to Qualified Mentor-Trainer, and were invited to attend the All Mentor Training in October 2009.

The additional responsibilities of a QT necessitated additional training, which was held October 26-27, 2009 in Anchorage. This training provided more in-depth information on the creation of and changes to the 2009-2010 Alternate Assessments and Secure website, including training tips to the Qualified Trainers.

Improvements to Web-based Training System

Requests from the field and from EED resulted in several improvements to the web-based training system were made prior to the start of the 2009-2010 testing window, including:

Navigation

A menu of training topics was added to the website as a locator page and as a table to the right side of each training page. These menus allowed users to see the training topics from each training page, and to choose to move to the next page, or to skip to a topic chosen by the user for review.

In addition, a checkbox at the bottom of each page (Mark this area complete) was added. Choosing the checkbox added a green check mark next to the topic in the menus, allowing users to see quickly which topics had been reviewed, and which had not.

Proficiency testing

The proficiency tests were rewritten to address areas of frequent concern or error. Two versions of each topic area (administration, reading, writing, math and science) were created. Previous years' training allowed users ten trials before requiring reset of the proficiency assessment; the 2009-2010 training site allowed only two trials. If a user was unsuccessful in reaching proficiency after two trials, the user's mentor was alerted and reset the proficiency tests to allow an additional two trials. Mentors were trained to review the reasons for not passing and to provide additional coaching as needed.

Enhanced supporting documents

Each training video and each proficiency assessment video were accompanied by supporting documents: A transcript of the video; a copy of the student scoring protocol and relevant student materials; and an explanation of the answers in the training. Only the mentors had access to the explanation of answers for the proficiency tests.

Student demographic data entry

The student demographic data entry page was reorganized. The "Date of Birth" field was altered to force users to fill this field in, rather than an automatic population of that field.

Score entry

The Score Entry main page was reorganized, and the "Submit Data" button moved to a more prominent location. In addition, the "submit data" button was also added to each score entry page for each assessment. After completing data entry for an assessment, users were prompted to submit the data to EED.

Accommodations data

In response to federal requirements to collect data on the accommodations that are offered to students with disabilities in state assessments, DRA added a data collection tool at the beginning of each assessment's score entry page. Users were asked to answer if they had employed accommodations (Yes or No). If "Yes" then a new text box became visible, asking them to describe the accommodations offered.

Enhancements for mentors

New tools were added to allow Qualified Mentors to play a more active role in the training and monitoring of their protégés:

Resetting proficiency attempts – Mentors were alerted when one of their protégés required more than two proficiency attempts, and tools were added to allow the Mentors to easily reset the proficiency tests for an individual protégé, or for all protégés requiring reset.

Changing protégé status – The website was enhanced to allow Mentors the ability to change a protégé's status from assessor-in-training to Qualified Assessor.

Certificates – All users were given the ability to print an official certificate which stated the user’s name, school, the year, and the user’s status. The status field is populated by the user’s current status (assessor-in-training, Qualified Assessor or Qualified Mentor-Trainer).

Reports – New reports were created and Mentors were given access to those reports. The reports included: User Contact; User Status History; User Training; User Proficiency Overview; User Proficiency Item Performance; User Data Entry; and Student Demographic Verification reports.

New Mentor Training

The purpose of the Alternate Assessment Mentor Program is to prepare district level trainers who train district personnel in correct test administration procedures for the Alternate Assessment. Mentors are available throughout the year to answer questions and assist district personnel. They are the first point of contact in the district for the state’s Alternate Assessment Program Manager. Additionally, mentors act as an advisory group for the Alternate Assessment. Mentors should be a certified teacher in the State of Alaska with a special education endorsement and have experience with low-incidence disabilities. The state encourages every district to have at least one Qualified Mentor Trainer and one Qualified Assessor.

The bulk of training occurs on the website <http://ak.k12test.com>. Assessors-in-Training (AIT) participate in a series of video vignettes designed to familiarize them with both appropriate testing and scoring techniques. These training vignettes familiarize Assessors-in-Training with the wide variety of tasks they will encounter on the Alternate Assessment, and demonstrate all the nuances needed in a proper administration. Following the training exercises, Assessors-in-Training must pass a series of brief proficiency tests related to the different tasks in each content area, as well as tests on general administration.

Summary of Dates and Participants

The New Mentor training was conducted in Juneau September 28 and 29, 2009. After a brief introduction, instruction in obtaining passwords and login identities, and navigating through the Alternate Assessment training and score entry website, participants completed online training and proficiency tests for each of five content areas: test administration, reading, writing, mathematics and science.

< See Appendix 2.1 >

Scoring Reliability Analysis

The second day of training was devoted to gaining proficiency in administering the test to a “protégé.” Eight of the participants had previously reached Mentor status, while eight were new to the alternate assessment system. The new participants administered and scored practice tests to the experienced participants. These materials were collected and rated against a master document for accurate administration and scoring. Participants administered assessments in all four subject areas. All participants earned Qualified Mentor status as a result of this two-day training regimen.

< See Appendix 2.2 >

Online Proficiency

In order to ensure that the Alternate Assessments are administered reliably throughout the state, and that valid and reliable test scores are being recorded, thorough training is required for all QAs. As described in the previous section, QAs must complete online training and proficiency tests, which focus on proper scoring in all of the different task types. Ample practice is provided through training vignettes and corresponding proficiency tests. Only after passing these tests does an individual become a QA and begin administering the test to students.

Scoring Reliability Analysis, Proficiency Test Attempts

DRA conducted an analysis of the number of attempts required before an assessor-in-training (AIT) successfully completed proficiency tests in each subject area. Each AIT is allotted two attempts to pass a proficiency test in each of five areas (administration, reading, writing, mathematics, and science). Each test consists of 20 questions; 80% correct is required for passing. If an AIT is unable to pass a test in two trials, that user's Mentor is able to reset the testing to permit an additional two attempts.

Nearly 300 teachers participated in all five proficiency testing subject areas.

Administration: Only one participant required 3 or more attempts to pass

Reading: Three participants required 4 attempts to pass the reading proficiency

Writing: Fifteen assessors required a fourth attempt to gain proficiency in writing; all but one passed the writing proficiency within five attempts. .

Math: Only three assessors required a fourth attempt to pass the math proficiency test.

Science: Only one assessor required a third attempt to pass the science proficiency test.

< See Appendix 2.2 above >

Annual Mentor Training

Summary of Dates and Participants

Annual Mentor training was held October 26 and 27 in Anchorage. All participants from the New Mentor Training in September attended as well as veteran Qualified Mentors and Trainers.

< See Appendix 2.3 >

On the first day, Mentors engaged in a review of the web training site and strategies for training protégés in the Alternate Assessment. In addition, Mentors participated in inter-rater reliability activities, discussed below.

< See Appendix 2.4 >

On the second day, Mentors engaged in training around using accommodations in instruction and assessment, led by Dan Kaasa, and a mentor focus group. In addition, Mentors offered suggested improvements to the writing scoring rubric for Ideas and Organization.

< See Appendix 2.5 >

Scoring Reliability Analysis, New Mentor and All Mentor Training

Mentors participated in scoring reliability events in writing, math and reading. All mentors worked from the same scoring protocols, while the leader posed as a student being assessed. Correct scoring was explained, and percent correct calculated for each mentor.

Reliability agreements for the math tasks were 93%, while the inter-rater reliability rate for the reading tasks was 92%. Mentors participated in scoring writing tasks: Correct word sequence and Ideas and Organization (I/O) scoring. The first task yielded 50 of 54 points in agreement, for a 93% inter-rater reliability. The second task included a Correct Word Sequence (CWS) score and a score for Ideas and Organization (I/O). Forty-nine of the 51 Qualified Assessors / Mentors agreed on the CWS (96% agreement), while 49 of the participants agreed on the I/O score (86%). The third task had 88% agreement on the CWS score and 55% agreement on the I/O score. However, when factoring in scores that were within 1 point of the correct I/O score, the inter-rater reliability for this task 96%.

< See Appendix 2.6 >

Scoring Reliability Analysis, Science Score-Behind Study

Observations of qualified assessors administering the eighth grade Alaska Alternate Science Assessment to eighth grade students was conducted by Aran Felix (EED) and Kim Sherman (DRA) on March 15, 16, and 17, 2010. In addition to these observations in Fairbanks, North Pole and Anchorage, volunteer Qualified Assessors in Lower Kuskokwim, Kenai Peninsula, Mat-su and Juneau school districts conducted similar observations.

Observers reviewed student IEP files to ensure that the student was qualified to participate in the Alternate Assessment, was in eighth grade, and had Alternate Assessment marked as the assessment choice in his or her IEP. In addition, observers noted accommodations or assistive technology devices indicated in the student's IEP for instruction and for assessment.

Observers sat near enough to the student and assessor to hear or see the student's response and to mark a scoring protocol. After the assessment, the observer collected a copy of the official scoring protocol, and completed an observers checklist. Finally, the observer rated the experience.

In general, assessors were found to deliver the assessment as instructed, and to record student responses and scores as the student moved through the assessment. Observers were also able to note issues in test layout that posed problems to assessors in terms of breaking the flow of delivery, or in one case, increasing the likelihood of a missed question. These issues were noted for incorporation in improvement to the 2010-2011 test documents.

< See Appendix 2.7 >

Webinar

The week prior to the opening of the Alternate Assessment testing window, DRA and EED hosted a web-based seminar (webinar) of the updates to the Alternate Assessment website. The webinar was repeated on January 29th and February 4th. The webinar served to update Qualified Mentors to improvements to the Alternate Assessment website (led by DRA) and to update them on procedural information related to the Alternate Assessment system and procedures (led by EED).

< See Appendix 2.8 >

CHAPTER 3: TEST IMPLEMENTATION IN 2009-2010

Updates on Procedures

The 2009-2010 Alternate Assessment represented a new form of the assessment, though the layout and administration were similar to previous years' assessments. Referred to in-house as "Form A", this year's assessment required extensive editing and online programming prior to being available to assessors in February. An editing work group, consisting of two representatives from EED and two from DRA, convened in Juneau in late August to work through the new version of the assessment to identify errors. Corrections were made and a final draft was sent to an external proofreader. Her corrections were returned to DRA in September. At this time, a switch in layout software at DRA allowed new errors to appear in the final draft, errors that were not caught by DRA internal quality assurance procedures.

Corrected errors on alternate assessment documents

After the window opened, errors were reported on the test that would have compromised results for those tasks/items. DRA and EED determined to make corrections to the test documents, and to alert the assessors of the errors and the corrections.

Writing, Grades 3/4, Scoring Protocol Task 1.34C: The task should have been for students to write their first name. The test incorrectly directed students to write their first and last name. The documents were corrected to require first name only, and assessors who had already administered this item were instructed to rescore this item to obtain a score based only on the student's first name.

Reading Grades 7/8, Scoring Protocols for Task 2.78B, Item 2): Item 2 should have read, "Tell/Show me the beginning of the story." The document incorrectly read "Tell/Show me the ending of the story." The document was corrected.

Math Grades 9/10, Scoring Protocols for Task 6.910A and Student Materials for Task 6.910A-D: The Math Task 6.910 Student Materials were mislabeled; they should have been labeled as:
6.910A - Describe and Compare Shapes / Shapes Greater Than, Less Than, Equal To
6.910B - Lines of Symmetry
6.910C - Identify Perimeter
The document was corrected.

Developed more detailed quality assurance schedule

To prevent the future need to correct test documents after the testing window opens, DRA developed a more detailed quality assurance plan, with a greater number of independent, outside proofreaders involved in document development.

< See Appendix 3.1 >

Help Desk Summary and Frequently Asked Questions Summary

The helpdesk was open from October 2, 2009 through April 17, 2010. During this time period, a helpdesk representative was available both by phone and by email to answer questions about the training, test administration, data entry, and any other questions from the field. Questions were logged into a database and categorized by question type. The most frequently asked questions and important issues to address were also added to a database that will be posted for all assessors to view prior to the 2010 testing window.

< See Appendix 3.2 >

Quarterly Reports

At the conclusion of each quarter (September, December, March and June), DRA submitted a summary report of accomplishments completed in the preceding four months. These quarterly reports described completion of contract deliverables.

< See Appendix 3.3 >

CHAPTER 4: CONTINUOUS IMPROVEMENT

Review of Replacement Process for Operational and Field Test Items

The 2008-2009 Alternate Assessment served as the baseline document; items included in this version are referred to as Operational Items. The Form A tests were developed by identifying the strand, task, and construct for each operational item and locating matching cousin items for each. Approximately one-half of the 2008-2009 test items (operational items) were replaced by matching cousin items (field test items). Operational items are items that have been used (without modification) in the 2007–2008 and 2008–2009 assessments, and thus had two years of statistical data collected for each item. Field test items were carefully created to match the operational items they would replace.

This system will allow DRA to conduct statistical analysis on the operational items, on the field test items, and equivalent test analysis.

At least six iterations of the Alternate Assessment are possible using the existing operational items and field test items:

- 1) AK AA Test (2007-2008, 2008-2009)
- 2) 1/2 (a) AK AA test + 1/2 (a) FT (FORM A, 2009-2010)
- 3) 1/2 (b) AK AA test + 1/2 (b) FT (FORM B, 2010-2011)
- 4) 1/2 (a) AK AA test + 1/2 (b) FT (FORM C, 2011-2012)**
- 5) 1/2 (b) AK AA test + 1/2 (a) FT
- 6) 1/2 (a) FT + 1/2 (b) FT

Form C Test Development

During 2010-2011, DRA will create Form C of the assessment. This form will use the same operational items used in Form A (1/2 (a) AK AA test) and the field test items used in Form B (1/2 (b) FT). Item selection will be guided by this format, and recorded in the Directory of Test Specifications.

Proficiency Testing and Practice Tests

The practice tests used in training Qualified Assessors and Qualified Mentor-Trainers (and available for practice with students) were edited to accurately reflect the current Alaska Alternate Assessment items. The proficiency tests for 2009-2010 were newly developed, with new video segments, scoring protocols and student materials; DRA will review these documents in 2010-2011 for currency and effectiveness. Issues to consider include:

- Do the tests address complicated issues likely to cause confusion (and requiring greater training)?
- Are the test questions worded in non-ambiguous ways, allowing only one clearly correct answer? Or, if nuanced, is the answer completely explained?
- Is a representative sampling of strands, scoring and administration strategies included in the tests?

Refresher Tests

EED has determined a four-year cycle for proficiency / refresher test requirements:

Year 1: New Assessor-in-training, or returning after absence; take all proficiency tests

Year 2: Take reduced number of refresher tests

Year 3: Take reduced number of refresher tests

Year 4: Take all proficiency tests, start cycle over

DRA will develop a system to track an assessor's status, so that each may be directed to the appropriate assessment type (proficiency or refresher). That system is in development currently, and will be refined in the coming years.

Mentor Oversight and Database Development

Mentors have access to online reports that will allow them to track their protégés' progress through training, update their status to Qualified Assessor when appropriate, track progress toward entering student demographic information, progress toward completion of assessment administrations, and to track any assessors who have not completed student assessments during the last week(s) of the testing window.

Detailed training will be delivered in the 2010-2011 All Mentor Training session to help mentors understand these reports and their use.

Summary of Consequential Survey

Virtually all teachers agreed that ...

After completing all training, I feel fully capable of administering the Alternate Assessment.
Decision-making is clear for administration of standard (STD) or extended levels of support (ELOS) test items.
Scoring criteria for tasks are clear.

Most teachers agreed that...

The results accurately depict what students are capable of doing on the Extended Grade Level Expectations (ExGLEs).

A split group of teachers agreed that...

The content covered on Alaska's Alternate Assessment is closely related to what I teach.
I use the ExGLEs to guide instruction.

The IEP team uses the ExGLEs to guide writing academic goals and objectives.

Giving the Alternate Assessment increases the kinds of accommodations, assistive technologies, supports I provide to my students.

Giving the Alternate Assessment increases my academic expectations for students with significant cognitive disabilities.

With the alternate assessments, I learned new information about my students' knowledge and skill in academic content.

Through administering the alternate assessment, I personally learned new skills for assessing student academic performance.

It is important to include students with significant cognitive disabilities in the statewide assessment and accountability system.

How many hours per week do your students who take the Alternate Assessment spend learning academics (reading, writing, math, science)? Most teachers spent 4 to 7 or more hours.

A vast majority of teachers agreed that...

Providing instruction in language arts (reading and writing) aligned to content standards and alternate assessments.

Providing instruction in mathematics aligned to Extended Grade Level Expectations and alternate assessments.

Providing instruction in science aligned to Extended Grade Level Expectations and alternate assessments.

Appropriately balancing academic content aligned to standards and functional or daily living skills.

Using appropriate accommodations that assist students in learning and showing what they know and can do.

Using Alternative or Augmented Communication (AAC) systems that assist students in learning and showing what they know and can do.

Explaining alternate assessment results to parents.

< See Appendix 4.1 >

Technological Improvements

Continuous improvement of the web-based training, monitoring, and data entry/scoring system is agreed to annually. The 2010-2011 system improvements include:

- Secure encryption of all data entry on all sections of the web-based system
- Reduction in proficiency testing requirement for returning Qualified Assessors and Qualified Mentors
- Require user increased demographic input (years in teaching, certifications held)
- Collect data regarding timing/scheduling of the alternate assessments
- Reorganize secure tests for download
- Restrict download of tests to Qualified Assessors or higher, while making available all other materials to all users, including Assessors-in-Training
- Make available Unofficial Student Reports after the testing window closes

<See Appendix 4.2>

CHAPTER 5: INTERPRETING STATISTICAL ANALYSIS OF 2009-2010

A number of tables are presented in the appendix, displaying various statistics for use in interpreting the AK Alternate Assessments.

Frequency counts are used to display the number and percentage of students at various grade bands. The number and percentage of students at each score value are also displayed in the appendix. An important statistic in every table is the valid N or the number of students represented in the statistic for any given measure. Means and standard deviations are used to describe the distributions at various grade bands. These two statistics should be interpreted relative to each other; ideally, the SD is less than (even half) the amount mean, which can be interpreted as reflecting an appropriate amount of variation. When the SD is close to or greater than the mean, then the distribution is difficult to describe as there appears to be as much variation as there is centeredness. Minimums and maximums reflect the smallest and largest scores obtained on the test, respectively.

Many tables have a total that simply reflects the sum of any frequency count across all categories (e.g. grade level or score value). System missing refers to the number of students who are not in that statistical calculation (either frequency or mean).

When reliability coefficients are displayed, a value is presented that varies from a low moderate decimal (in the .30-.50 range), a moderate range (.51 to .79) or a relatively high value (in the .80 to .97 range). These values represent the degree to which two variables (e.g. forms of the test or items within the test) are related. Generally, higher is better, as the information from one measure (item or form) can be used to predict another item or form. In some cases, however, the values should not be too high (e.g., when reflecting the relations among different items in the test), because it would mean that, essentially, they are duplicating the information.

This statistic, however, is a function of the number of values (in the test) that are counted (as well as the number students behind any of these values). For example, at the total test level, many items are used to calculate the coefficient; at the strand level, sufficient items are present. However, at the task level, the number of items is so few that the values are likely to be low because there simply is not enough variation present to reflect a high coefficient.

CHAPTER 6: RELIABILITY

The data file was analyzed for reliability at several levels. First, at the total test level, which is the most important because Adequate Yearly Progress (AYP) is established on the basis of this score, reliability coefficients are reported for every grade band and subject area. Second, at the strand level, coefficients are reported for every grade band and subject area. The test was designed to reflect scores at this level to ensure adequate representation across the entire range of Extended Grade Level Expectations; in the official student reports, scores for every strand are reported so that parents and teachers can follow the performance and progress of students. Third, and perhaps least important, are the scores at the task level; though we report these coefficients, they are primarily directed toward the continuous improvement of the test as EED develops new field tests and integrates them into the operational test.

In the tables for total test and strands, the reliability coefficients are reported for both the entire population (ALL students) and the students who took the complete Standard administration with students who participated in the Expanded Levels of Support (ELOS) removed (with NO ELOS). This population includes students with extremely low levels of functioning with little to no interactivity or means of communication. The reason for removing this group was to investigate the influence of missing data and its potential to spuriously inflate reliability coefficients. The first step in removing this group was to integrate the ELOS data file with the Standard Administration file. The second step involved splitting the file on ELOS participation and removing them so that all reliability coefficients could be recomputed at each level (total test, strand, and task). This re-analysis was done for each subject area and at all grade bands.

In general, the findings indicate that the test is very reliable for decision-making (of AYP) at the total test level. Scores were quite reliable at the strand level (with only a few strands reflecting moderate coefficients, which was primarily a function of the few number of tasks involved). Finally, as expected, scores were moderately reliable at the task level, primarily because of the few items involved. Another general (and expected) finding is that the coefficients are somewhat lower when the ELOS students scores are removed from the Standard administration file although the reduction is not large, as only 9-11% of the students took the ELOS option.

Total Test

In this analysis, all items used in a grade band **test** were entered for computing the reliability of the entire test: Cronbach's Alpha based on Standardized Items.

Total Test Reliability with ALL Students

	<i>Reading</i>	<i>Writing</i>	<i>Math</i>	<i>Science</i>
Grade 3-4	.912	.953	.934	.855
Grade 5-6	.912	.936	.933	NA
Grade 7-8	.914	.930	.918	.836
Grade 9-10	.868	.876	.916	.850

Total Test Scores with NO ELOS Students

	<i>Reading</i>	<i>Writing</i>	<i>Math</i>	<i>Science</i>
Grade 3-4	.848	.896	.882	.744
Grade 5-6	.825	.894	.870	NA
Grade 7-8	.893	.913	.918	.836
Grade 9-10	.838	.815	.908	.842

Strand Reliability

In this analysis, all items used in a grade band **strand** were entered for computing the reliability of the strand: Cronbach's Alpha based on Standardized Items.

Reading with ALL Students

<i>Strands</i>	<i>Word ID</i>	<i>Form Gen Understand</i>	<i>Analysis Content-Structure</i>
Grade 3-4	.912	.814	NA
Grade 5-6	.863	.887	.731
Grade 7-8	.876	.877	.872
Grade 9-10	.822	.616	.782

Reading with NO ELOS Students

<i>Strands</i>	<i>Word ID</i>	<i>Form Gen Understand</i>	<i>Analysis Content-Structure</i>
Grade 3-4	.882	.670	NA
Grade 5-6	.759	.839	.634
Grade 7-8	.835	.852	.851
Grade 9-10	.752	.582	.753

*Writing with ALL Students**

<i>Strands</i>	<i>Variety Forms</i>
Grade 3-4	.953
Grade 5-6	.937
Grade 7-8	.890
Grade 9-10	.874

**Not all strands were represented*

*Writing with NO ELOS Students**

<i>Strands</i>	<i>Variety Forms</i>
Grade 3-4	.896
Grade 5-6	.902
Grade 7-8	.868
Grade 9-10	.811

**Not all strands were represented*

*Mathematics with ALL Students**

	Num.	Geom.	Est.Comp.	Func.Rel.	Msmt.	Stats.Prob.
Grade 3-4	.926	.857				
Grade 5-6	.914	.812	.932	.846	.745	
Grade 7-8	.893	.686	.842	.937	.805	
Grade 9-10	.846	.712	.783	.680	.747	.837

*Not all strands were represented in the test

*Mathematics with NO ELOS Students**

	Num.	Geom.	Est.Comp.	Func.Rel.	Msmt.	Stats.Prob.
Grade 3-4	.878	.745				
Grade 5-6	.749	.646	.916	.743	.699	
Grade 7-8	.829	.686	.833	.933	.805	
Grade 9-10	.665	.700	.748	.522	.744	.824

*Not all strands were represented in the test

Science with ALL Students

<i>Strands</i>	<i>Physical</i>	<i>Life</i>	<i>Earth</i>	<i>SciTech</i>
Grade 4	.847	.636	.826	.661
Grade 8	.435	.596	.832	.571
Grade 10	.804	.776	.450	.616

Science with NO ELOS Students

<i>Strands</i>	<i>Physical</i>	<i>Life</i>	<i>Earth</i>	<i>SciTech</i>
Grade 4	.582	.537	.709	.480
Grade 8	.435	.599	.746	NA
Grade 10	.703	.736	.416	.554

Task Reliability

In this analysis, all items used in a grade band **task** were entered for computing the reliability of the task: Cronbach's Alpha based on Standardized Items.

<i>Reading*</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>T7</i>
Grade 3-4	.679	.882	.937		.814		
Grade 5-6	.882		.882				
Grade 7-8		.740		.874			
Grade 9-10	.947	.665		.815			

*Note. Reliability analyses conducted for all subtests with 4 or more items.

<i>Writing*</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>T7</i>
Grade 3-4	.910	.932					
Grade 5-6			.937				
Grade 7-8			.890				
Grade 9-10			.874				

*Note. Reliability analyses conducted for all subtests with 3 or more items.

<i>Mathematics*</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>T7</i>	<i>T8</i>	<i>T9</i>	<i>T10</i>
Grade 3-4	.961			.768						
Grade 5-6					.932	.846			.745	.826
Grade 7-8						.842		.937		
Grade 9-10								.837		

For grades 7-8, the reliability for task 13 was .777 and for task 14 was .703

**Note. Reliability analyses conducted for all subtests with 4 or more items.*

Split Half Reliability of Alternate Forms for Subject Areas and Grade Bands

Reading Grade 3-4

Task	No. Items	Split half
1	4	.72
2	5	.91
3	4	.95
5	4	.76
6	4	.76
7	4	.88

Reading Grade 5-6

Task	No. Items	Split half
1	4	.93
2	3 and 2	.93
3	6	.91
4	6	.88
5	6	.88

Reading Grade 7-8

Task	No. Items	Split half
1	2	.92
2	4 and 3	.75
3	2	.92
4	6	.88
5	6	.87
6	6	.87

Reading Grade 9-10

Task	No. Items	Split half
1	4	.95
2	3	.82
3	6	.89
4	6	.82

Mathematics Grade 3-4

Task	No. Items	Split half
1	4	.96
2	4	.78

Mathematics Grade 5-6

Task	No. Items	Split half
1	3 and 2	.86
5	4 and 3	.93
6	4 and 3	.91
7	5	.94
9	2	.68
10	4 and 3	.89

Mathematics Grade 7-8

Task	No. Items	Split half
3	3 and 2	.89
6	4 and 3	.94
8	2	.93
10	5	.74
11	3	.79
13	2	.85
14	4 and 3	.67

Mathematics Grade 9-10

Task	No. Items	Split half
1	3 and 2	.88
2	2	.79
5	3	.86
6	3	.80
7	2	.81
8	5	.80
9	3	.70
11	4	.73

Split Half Reliability – Operational Versus Field Test Items
Reading

GBand	<i>n</i> items		Valid <i>n</i>		Excluded		Cronbach's Alpha	
	O	F	O	F	O	F	O	F
34	21	19	119	117	509	511	.901	.859
56	19	16	131	127	497	501	.904	.857
78	20	15	128	124	500	504	.895	.914
910	22	19	103	101	525	527	.863	.844

Note. O = Operational item, F = Field test item

Writing

GBand	<i>n</i> items		Valid <i>n</i>		Excluded		Cronbach's Alpha	
	O	F	O	F	O	F	O	F
34	19	11	128	122	499	506	.934	.742
56	30	9	130	134	497	494	.933	.745
78	31	12	125	133	502	495	.918	.871
910	31	9	112	119	515	509	.916	.859

Note. O = Operational item, F = Field test item

Mathematics

GBand	<i>n</i> items		Valid <i>n</i>		Excluded		Cronbach's Alpha	
	O	F	O	F	O	F	O	F
34	19	7	128	129	499	498	.911	.783
56	30	21	130	128	497	499	.907	.901
78	31	33	125	128	502	499	.869	.918
910	31	33	112	109	515	518	.897	.909

Note. O = Operational item, F = Field test item

Science

GBand	<i>n</i> items		Valid <i>n</i>		Excluded		Cronbach's Alpha	
	O	F	O	F	O	F	O	F
4	12	12	89	87	176	178	.836	.718
8	12	12	74	74	191	191	.837	.768
10	14	14	61	61	204	204	.840	.735

Note. O = Operational item, F = Field test item

Participation in Extended Levels of Support

Because the reliability coefficients are influenced by scores of 0 (students taking ELOS tests), several analyses were conducted on participation in the ELOS.

ELOS Students by Grade-level: Reading

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	15	11.8	11.8	11.8
	4	19	15.0	15.0	26.8
	5	13	10.2	10.2	37.0
	6	14	11.0	11.0	48.0
	7	16	12.6	12.6	60.6
	8	11	8.7	8.7	69.3
	9	16	12.6	12.6	81.9
	10	23	18.1	18.1	100.0
	Total	127	100.0	100.0	

Grade 3 – At grade 3, all students fulfill the 3 x 3 rule in the first 3 tasks.

Grade 4 – At grade 4, all but two students fulfill the 3 x 3 rule in the first 3 tasks. Stu_sid 378 and 384 receive three consecutive zeros on tasks 5 and 6, at which point the assessment concludes.

Grade 5 – At grade 5, all but two students fulfilled the 3 x 3 rule in the first 3 tasks. Stu_sids 295 fulfilled the rule in tasks 2-4 and stu_sid 469 appears to finish the assessment.

Grade 6 – At grade 6, all but four students fulfilled the 3 x 3 rule in the first 3 tasks. Stu_sid 382 receives zeros on the first three items of tasks 3 and 4. Stu_sid 28 and 202 do not complete task 4. Stu_sid 450 appears to complete the test

Grade 7 – At grade 7, all but three students fulfilled the 3 x 3 rule in the first 3 tasks. Stu_sid 96 and 582 fulfilled the rule during tasks 2-4. Stu_sid 583 appears to complete the test.

Grade 8 – At grade 8, all but two students fulfilled the 3 x 3 rule in the first 3 tasks. Stu_sid 26 fulfilled the rule during tasks 2-4 while stu_sid 685 completed all but task 4.

Grade 9 – At grade 9, all but one student fulfilled the 3 x 3 rule in the first 3 tasks. Stu_sid 49 fulfilled the rule during tasks 2-4.

Grade 10 – At grade 10, all but one student fulfilled the 3 x 3 rule in the first 3 tasks. Stu_sid 282 did not complete tasks 4 or 5, at which point the test concludes.

ELOS Students by Grade-level: Writing

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	17	12.9	12.9	12.9
	4	20	15.2	15.2	28.0
	5	13	9.8	9.8	37.9
	6	12	9.1	9.1	47.0
	7	19	14.4	14.4	61.4
	8	11	8.3	8.3	69.7
	9	15	11.4	11.4	81.1
	10	25	18.9	18.9	100.0
	Total	132	100.0	100.0	

Grade 3 – At grade 3, all students fulfill the 3 x 3 rule in the first 4 tasks (task 3 has only 1 item).

Grade 4 – At grade 4, all students fulfill the 3 x 3 rule in the first 4 tasks (task 3 has only 1 item).

Grade 5 – At grade 5, all but three students fulfill the 3 x 3 rule in the first 4 tasks (task 2 has only 1 item). Stu_sid 170 received zeros during tasks 2-4, at which point the assessment concludes. Stu_sid 295 scored a 5 on item 1 of task 4 and has zeros everywhere else. Stu_sid 469 appears to finish the test.

Grade 6 – At grade 6, all but one student fulfills the 3 x 3 rule in the first 4 tasks (task 2 has only 1 item). Stu_sid 202 received scores on all three items of task 4, and had all zeros on all other tasks.

Grade 7 – At grade 7, all but five students fulfill the 3 x 3 rule in the first 3 tasks. Stu_sid 108 fulfills the rule during tasks 3, 4, and 5. Stu_sid's 157, 313, 368 and 583 appear to complete the test.

Grade 8 – At grade 8, all but one student fulfill the 3 x 3 rule during the first 3 tasks. Stu_sid 685 fulfills the rule during tasks 4, 5, and 6.

Grade 9 – At grade 9, all but three students fulfill the 3 x 3 rule during the first 3 tasks. Stu_sid's 49 and 50 score points during the first task, then receive zeros on all items on the next two tasks, at which point the assessment concludes. Stu_sid 387 appears to complete the test.

Grade 10 – At grade 10, all but four students fulfill the 3 x 3 rule during the first 3 tasks. All four of these students (stu_sid 46, 51, 552, and 699) receive points during the first task, then receive zeros on all items on the next two tasks, at which point the assessment concludes.

ELOS Students by Grade-level: Math

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	14	12.6	12.6	12.6
	4	16	14.4	14.4	27.0
	5	11	9.9	9.9	36.9
	6	12	10.8	10.8	47.7
	7	11	9.9	9.9	57.7
	8	10	9.0	9.0	66.7
	9	13	11.7	11.7	78.4
	10	24	21.6	21.6	100.0
	Total	111	100.0	100.0	

Grade 3 – At grade 3, all but one student fulfilled the 3 x 3 rule on tasks 1-4 (task 3 has only 1 item). Stu_sid 182 was the exception and was able to complete the test.

Grade 4 – At grade 4, all but three students fulfilled the 3 x 3 rule on tasks 1-4 (task 3 has only 1 item). Stu_sids 384, 173, and 181 fulfill 3 x 3 rule on tasks 4, 5, and 6, at which point the assessment concludes.

Grade 5 – At grade 5, all but one student fulfilled the 3 x 3 rule on tasks 1-5 (tasks 3 and 4 have only 1 item each). Stu_sid 295 fulfills the 3 x 3 on tasks 2-6.

Grade 6 – At grade 6, all but two students fulfill the 3 x 3 rule on tasks 1-5 (tasks 3 and 4 have only 1 item each). Stu_sids 74 and 202 fulfill the 3 x 3 rule on tasks 2-6.

Grade 7 – At grade 7, all but four students fulfill the 3 x 3 rule on tasks 1-6 (task 2 has only 2 items and tasks 4 and 5 have only 1 item each). Stu_sid 25 received zeros on tasks 4, 5, and 6, at which point the assessment was stopped. Stu_sid 29 received zeros on all but the first item on task 6, then received all zeros on tasks 7 and 8 at which point the assessment was stopped. Stu_sid 96 received zeros on tasks 4, 5, and 6, at which point the assessment was stopped. Stu_sid 582 does not have data for all tasks, but does not appear to ever have any task with three zeros.

Grade 8 – Stu_sid 230, 243, 323, 324, 400 all receive zeros on the first three tasks. Stu_sid 155, 685 received zeros on tasks 3-7. Stu_sid 26 received a zero on tasks 4-8. Stu_sid 156 received zeros on tasks 6, 7, and 8.

Grade 9 – At grade 9, all but one student fulfilled the 3 x 3 rule in the first three tasks. Stu_sid 387 fulfilled the 3 x 3 rule in tasks 4, 5, and 6.

Grade 10 – At grade 10, all but 3 students fulfilled the 3 x 3 rule in the first three tasks. Stu_sids 46 and 552 fulfilled the 3 x 3 rule in tasks 4, 5, and 6. Stu_sid 359 completed the test.

Descriptive Statistics - Math

	N		Minimum		Maximum		Mean		Standard Deviation	
	Non-ELOS	All	Non-ELOS	All	Non-ELOS	All	Non-ELOS	All	Non-ELOS	All
Nm34	124	154	.00	.00	25.00	25.00	20.2097	16.4610	6.34672	9.69073
Geom34	124	154	.00	.00	16.00	16.00	13.1613	10.6883	3.89159	6.18296
Nm56	131	152	.00	.00	10.00	10.00	8.5954	7.3816	1.97203	3.46589
EC56	131	152	.00	.00	8.00	8.00	5.6260	4.7961	3.28009	3.62540
FR56	130	144	.00	.00	8.00	8.00	6.6769	5.9861	2.00083	2.76336
Msmt56	129	138	.00	.00	8.00	8.00	5.7209	5.3188	2.59500	2.87447
Geom56	129	138	4.00	.00	14.00	14.00	9.9612	9.2681	2.47299	3.47785
Nm78	143	164	1.00	.00	10.00	10.00	8.4755	7.5366	2.45475	3.44170
EC78	143	164	.00	.00	8.00	8.00	4.9371	4.3354	2.99111	3.22048
FR78	142	149	.00	.00	4.00	4.00	3.4507	3.3221	1.26366	1.39627
Msmt78	142	145	1.00	.00	11.00	11.00	8.0000	7.8414	2.96552	3.13288
Geom78	142	146	3.00	.00	13.00	13.00	9.7042	9.4589	2.35137	2.74710
Nm910	118	155	.00	.00	8.00	8.00	6.0424	4.7097	2.16181	3.15604
EC910	118	155	.00	.00	9.00	9.00	4.8644	3.7032	3.04360	3.37133
FR910	117	130	.00	.00	5.00	5.00	3.7009	3.3462	1.43998	1.74192
SP910	115	120	.00	.00	14.00	14.00	9.5913	9.2833	3.68233	4.00039
Msmt910	114	118	.00	.00	9.00	9.00	5.3158	5.1780	3.30008	3.35029
Geom910	114	119	.00	.00	5.00	5.00	3.5614	3.4370	1.45757	1.56573

CHAPTER 7: DESCRIPTIVE STATISTICS

Descriptive statistics were calculated for each task, in every subject area, and in both grade bands and grade levels. The upper right header of each page refers the reader to the type of descriptive statistics displayed. For instance, "Grade Band Total Test Descriptive Statistics" refers to the descriptive statistics at the total test level for each subject, while "Writing Task Descriptive Statistics (Grade Band 3/4)" refers to the descriptive statistics for writing at the task level, in grade band 34. The following statistics are reported in the tables in Appendix 7 (leftmost column to rightmost column).

In each appendix, (grade band appendix and grade level appendix) the documentation is presented in the following order:

- Participation descriptive statistics at the total test level for each grade or grade band
- Score descriptives for total tests for each subject, including the frequencies of each score (some tables were overly lengthy and were thus not included).
- Task descriptives for each subject in the lowest grade or grade band (i.e, grade 3 or grade band 34).
- Strand descriptives for each subject in the lowest grade or grade band (i.e, grade 3 or grade band 34).
- Task descriptives for each subject in the next lowest grade or grade band (i.e, grade 4 or grade band 56).
- Strand descriptives for each subject in the next lowest grade or grade band (i.e, grade 4 or grade band 56).

This pattern continues until the highest grade or grade band (10 or 910). The subjects always appear in the following order: reading, writing, math, and science.

In these tables, the initial tables depict percentages of students participating, then the number of students at each score value for the total sum in a subject area. The following codes apply to grade levels and bands:

w=34, x=56, y=78, and z=910

r=reading, w=writing, m=mathematics, and s=science.

As an example, wtotr = Grade band 3-4 writing total score in reading.

Note: When a number follows the letter 't', it is the task number. When a number follows an underscore (_), it refers to the item number.

As an example, w3tot = Grade 3-4 writing task 3 total score and xt1_1 refers to grade 56, task 1, item 1.

< See Appendix 7.1 and 7.2 >

CHAPTER 8: ADEQUATE YEARLY PROGRESS

Adequate Yearly Progress (AYP) results are displayed with each attained score value presented in two different ways for depicting proficiency:

1. Four categories with 1 = Far Below, 2 = Below, 3 = Proficient, and 4 = Advanced
2. Two categories with 0 = Below (with Far Below and Below collapsed) and 1 = Above (with Proficient and Advanced collapsed)

For each table, the data present: (a) the frequency of the score value (Frequency), reflecting the number of students at that score value, (b) the percentage of students (Percent), reflecting the number of students in the grade band with a score value divided by all students taking the alternate assessment, including those with missing score values or in a different grade, (c) the percentage of students (Valid Percent), reflecting the number of students who actually had values divided by only those students with a score value in that grade band, and (d) the percentage of students with score values (Cumulative Percent), reflecting a running accumulation of percentages at/below that specific score value using only students in the grade band. The ‘Frequency’ and ‘Valid Percent’ need to be the focus of interpretations.

Reading

Grade 3-4: 63% of the students were proficient (46.8%) or advanced (16.2%).

Grade 5-6: Over 40% of students (43.8%) were proficient (41.2%) or advanced (2.6%).

Grade 7-8: About half of the students were proficient (48.8%); none were advanced.

Grade 9-10: A total of 45.2% were proficient (31.6%) or advanced (13.5%).

Writing

Grade 3-4: Over 2/3 of the students (67.5%) were proficient (26.6%) or advanced (40.9%).

Grade 5-6: Well over half (60.5%) of the students were proficient; none were advanced.

Grade 7-8: Over half of the students (52.1%) were proficient (42.4%) or advanced (9.7%).

Grade 9-10: About half (49.7%) of the students were proficient (41.9%) or advanced (7.7%).

Mathematics

Grade 3-4: Over 2/3 of the students (68.8%) were proficient (21.4%) or advanced (47.4%).

Grade 5-6: Over 3/4 of the students (77%) were proficient (34.2%) or advanced (42.8%).

Grade 7-8: Over half of the students (54.9%) were proficient (45.1%) or advanced (9.8%).

Grade 9-10: About 1/2 of the students (45.2%) were proficient (20.6%) or advanced (24.5%).

Science

Grade 4: Well over 3/4 of the students (84.2%) were proficient (25.7%) or advanced (58.4%).

Grade 8: Almost 3/4 of the students (72.8%) were proficient (16%) or advanced (56.8%).

Grade 10: Over half of the students (55.4%) were proficient (37.3%) or advanced (18.1%).

< See Appendix 8 >