

District Alternate Assessments

**A Guidance document prepared by the
Exceptional Students Leadership Unit (ESLU) at
the Colorado Department of Education (CDE)**

**By
Sri Srinivasan,
Colorado Department of Education
Exceptional Student Leadership Unit**

Table of Contents

- I. Introduction
- II. Processes for Developing a District Alternate Assessment
 - A. Overview
 - B. District-wide Assessments that Mandate Alternate Assessments
 - C. Colorado Basic Literacy Act and District Alternate Assessments
 - D. Broad Stakeholder Involvement
 - E. Alignment with District-wide Regular Assessment
 - F. Developing Accommodations Guidelines
 - G. Developing Eligibility Guidelines
 - H. Assuring Universal Design
 - I. Informing Parents and IEP teams
 - J. Understanding Validity and Reliability
- III. Choosing an Alternate Assessment Format
 - A. The Performance Based Assessment
 - B. The Body of Evidence Assessment (BOE)
 - C. The Three Step Guide to developing an Alternate Assessment System
 - D. Recommendations while purchasing or acquiring a commercially available test
 - E. Progress Monitoring
- IV. Data Management and Reporting
 - A. Data Collection, Storage and Analysis
 - B. Reporting of Results (legal requirements)
- V. Conclusion
 - A. Resources for developing and administering alternate assessments
 - B. District Alternate Assessment Development and Administration Considerations
 - C. CDE monitoring and compliance procedures for the implementation of the district alternate assessment
- VI. Contact Information
- VII. References
- VIII. Appendices

I. Introduction

Since the 1997 inception and the 2004 reauthorization of the Individuals with Disabilities Education Act (IDEA), the inclusion of students with disabilities in state and district-wide assessments has been a legal mandate. It is both expected practice and a state legal requirement (CRS 22-7-409) in Colorado that *all* students are included in state assessments by participating in the state's general academic test (the Colorado Student Assessment Program (CSAP) with or without accommodations or participating in the state alternate test - the Colorado Student Assessment Program Alternate (CSAPA)). The purpose of this document is to provide non-regulatory guidance to school districts on ways to align their existing district assessments with a district alternate assessment that meets the requirements of IDEA 2004.

Federal statute 20USC1411 provides the regulations on including children with disabilities in both state and district level assessments (including alternate assessments) with appropriate accommodations as indicated on their IEPs.

In relation to district assessments, IDEA states that districts must:

- Establish accommodation guidelines for their regular district assessments.
- Establish participation guidelines for the District Alternate Assessments (DAA) for those students who cannot participate in the regular district assessment even with accommodations.
- Report results for students with disabilities (on both the district assessment and district alternate assessment) with the same frequency and detail as students who do not have disabilities.
- Use Universal design principles in the development and administration of the alternate assessment.
- Must provide IEP Teams with a clear explanation of the differences between assessments based on grade-level academic achievement standards and those based on modified or alternate academic achievement standards, including any effects of State or local policies on the student's education resulting from taking an alternate assessment based on alternate or modified academic achievement standards (such as whether only satisfactory performance on a regular assessment would qualify a student for a regular high school diploma).
- Must ensure that parents of students selected to be assessed based on alternate or modified academic achievement standards are informed that their child's achievement will be measured based on alternate or modified academic achievement standards.

II. Processes for Developing a District Alternate Assessment

A. Overview

This document is intended to provide non-regulatory guidance to district personnel, teachers, researchers and test publishers on parameters in developing district alternate assessments as well as requirements in meeting the provisions of the law. This document is written by the Colorado Department of Education – Exceptional Students Leadership Unit (CDE-ESLU). Any information from this document could be copied only for educational not-for-profit purposes provided CDE-ESLU is cited appropriately.

B. District-wide Assessments that Mandate Alternate Assessments

IDEA reauthorization requires that if a school district has a district-wide assessment, then that school district must also have a district alternate assessment for those students who cannot participate in the regular assessment even with accommodations.

CDE has provided clarification related to district alternate assessments by defining the district-wide assessments as *assessments (in any content area) that are given to all students enrolled in a district in a grade for the purposes of assessing student's academic growth and achievement. A district-wide assessment does not include an assessment that is given district-wide for a specific group of students in a specialized program or class, for example, identification of giftedness.* These district-wide assessments should measure student growth and achievement in the district's academic curriculum. If the district assessment covers several different content areas e.g. Reading, Writing, Math, Science, etc., then the district alternate assessment must also be available in the same content areas. Also, any assessment that is diagnostic in its purpose and does not measure academic growth or achievement does not require an alternate assessment in place. Thus, districts must ensure that there is an alternate assessment in place for every **grade** and **content** area where a district-wide assessment is available.

In general, formative assessments do not mandate a district alternate assessment, while summative assessments do mandate a district alternate assessment. It is the point at which the assessment occurs in a program and the purpose of the assessment that distinguishes these two categories of assessment. These two forms of assessments are described below.

Summative Assessments

A Summative assessment for the purpose of the DAA guidance has been described under two sub headings:

1. The requirements
2. The guide to best practice

1. The requirements

CDE defines a summative assessment as an assessment that meets some or all of the following criteria.

A summative assessment:

- is an assessment whose purpose is to define district-wide performance or school-wide performance
- has high stakes for students or the schools (like accreditation)
- is an “end-of-course exam”
- is an exam which needs to be passed to graduate

- informs program evaluation at a school level
- is district-wide, point-in-time for all students in a given grade to take at the same time
- has a standardized administration with allowed accommodations
- measures performance in relation to grade-level standards
- measures grade-level expectations that are aligned to state and district content standards
- measures accountability
- is intended to provide a complete picture of student performance across the district
- is one shot, once a year, typically towards the end of the year or mid year.

The above mentioned list describes summative assessments for the purpose of DAAs. *While each of the above descriptors for a summative assessment will warrant an alternate assessment, it is still the responsibility of each district or administrative unit to identify the purpose of a district-wide assessment and decide on an assessment by assessment basis the need for an alternate assessment.*

The guide to best practice

Rationale: All district-wide assessments which ensure students are monitored in critical content areas mandate an alternate assessment. Another important reason for alternate assessments is for districts to develop a systematic plan to track the academic progress of students who are eligible for them. Measuring progress for all students will inform the quality of instruction all students receive. Given this, a district has an obligation to include all students when planning for assessments and budgeting for individual schools and programs, and eventually, for better quality instruction. An example is when an assessment is used to evaluate budget to schools or buildings within a district. In this light CDE expects district personnel to use their best judgment in identifying the district-wide summative assessments currently in place (or planned for in the future) that will mandate alternate assessments.

In most cases, the foundation for an assessment plan is to collect summative assessment data; this type of data can many times be stand-alone. Data from a summative assessment will inform student growth and academic achievement. Formative assessment data, however, can contribute to a comprehensive assessment plan by enabling faculty to identify particular points in a program to assess learning (i.e., entry into a program, before or after an internship experience, impact of specific courses, etc.) and monitor the progress being made towards achieving learning outcomes¹.

Hence, summative assessments in a district or an administrative unit that inform academic growth and achievement mandate that alternate assessments be in place to meet the requirements of the law.

Some examples of summative assessments currently in place in school districts and administrative units in Colorado include:

	Name of the Assessment	Grade	Purpose
1.	EXPLORE	8	To provide standardized data and interest/career exploration
2.	PLAN	10	To provide standardized data and interest/career exploration
3.	NWEA/MAP	3-9	To measure academic achievement
4.	CSAP and CSAPA	3-10	To measure student progress on Colorado state content standards
5.	McGraw Hill Unit tests	Various	Measure student progress

Note: The list above includes a few assessments used for summative purposes in some districts. These assessments may be used in other districts for other purposes. Therefore, it is still the responsibility of each district and administrative unit to ensure the intention (formative or summative) of every district-wide assessment that is currently in place and make decisions on the requirement of alternate assessments.

Formative Assessments

For the purpose of this guidance document, a formative assessment can be described as an assessment done at the beginning or during a program that provides immediate evidence of student learning in a particular course or at a particular point in a program. The purpose of a formative assessment is to improve the quality of student learning and determine where to begin instruction. The intention behind a formative assessment should be to inform curricular modifications when specific courses / curricula / teaching methodologies have not met student learning outcomes. Formative assessments also provide important program information when multiple sections of a course are taught, because it enables programs to examine if the learning goals and objectives are met in all sections of the course. Formative assessments also can improve instructional quality by engaging the faculty in the design and practice of the course goals and objectives and the impact of the course on the program¹.

Since formative assessments in a district or an administrative unit do not inform academic growth or achievement, a district or administrative unit is not required to have

an alternate assessment in place (for district-wide formative assessments) to meet the specifics of the law. However, it is highly recommended that similar formative assessment procedures be in place for all students who cannot participate in the district-wide formative assessments due to a disability.

Some examples of formative assessments currently in place in school districts and administrative units in Colorado include:

	Name of the Assessment	Grade	Purpose
1.	MAP	K-8	Instructional guidance, benchmarks
2.	Nelson Denny	9	Grade level equivalent, placement
3.	DRA	Elementary	Diagnostic
4.	PALS	Elementary	Diagnostic
5.	QRI	Elementary	Diagnostic
6.	Explore, Plan	8, 10	To provide standardized data and interest/career exploration

Note: The above mentioned list includes a few assessments used for formative purposes in some districts. These assessments may be used in other districts for other purposes. It is the responsibility of each district and administrative unit to ensure the intention (formative or summative) of every district-wide assessment that is currently in place and make decisions on the requirement of alternate assessments.

C. Colorado Basic Literacy Act (CBLA) and District Alternate Assessments

Is a district required to have alternate assessments for CBLA?

The Individuals with Disabilities Education Improvement Act 2004 (IDEA) requires districts to provide alternate assessments for students with disabilities who cannot participate in the district assessments with appropriate accommodations. All students, including students with disabilities, participate in the district chosen CBLA end-of-year assessments. Any student who is not proficient on the end-of-year assessment(s) is administered individualized diagnostic assessments to determine the individual student's reading proficiency levels and develop an individual literacy plan if necessary. CBLA assessments are inclusive of all students; therefore no alternate assessments are required for students with disabilities to participate in the district's assessments. Further, CBLA assessments are intended for diagnostic purposes and do not inform academic achievement or growth. Given the two above mentioned rationales, **CBLA assessments do not require alternate assessments.**

D. Broad Stakeholder Involvement

The development of a district alternate assessment should mirror the development and pilot phases of a large-scale assessment happening over a period of time. This process should occur at the district level in most cases, unless the district chooses to purchase a published test that aligns with the district-wide regular assessment and the district's standards. Administrative Units (AUs) are encouraged to use a regional or inter-district collaborative alternate assessment development process that will reduce the burden on any one district in developing a set of alternate assessments. A collaborative effort will prove effective in terms of utilization of resources (including personnel and material costs) and expansion of professional networks and will also provide for an increased awareness of the complexities of alternate assessment. Districts should verify that the district-level alternate assessment developed by the collaborative aligns with the district-wide-regular assessment already in place in their district. The alignment of the district alternate assessment with district-wide assessment and the general education curriculum is described in Section IV.E.

CDE emphasizes that the development of a district-level alternate assessment should involve special education experts, general education experts, content specialists and assessment specialists. Such membership from various disciplines in a district's assessment development process will result in a better opportunity to learn about the ongoing issues around alternate assessments and will pave ways for better alignment of the district-level alternate assessment to the district-wide regular assessment. Hence, CDE recommends to both groups (special education and assessment) to facilitate a united effort in the development of district alternate assessments.

An example of a collaborative process is outlined in the appendix A1.

E. Alignment with District-Wide Regular Assessment

The district alternate assessments should be developed on grade level expectations and the standards against which the district-wide assessment is developed. The district is responsible for creating guidelines that will help determine which assessment a student should take. However, it is still the responsibility of the IEP team to make the final determination as to which assessment the student should take. Eligibility guidelines for DAAs are discussed in section IV. G.

It is important for each district and administrative unit to undertake a crosswalk process to compare the district's regular assessment and district alternate assessment. Participation in this process should include those with expertise in special education, content (subject matter), assessment development and psychometric expertise. Using a blueprint (rubric) or framework to compare the district's regular assessment, the district curriculum, and the alternate achievement standards is an ideal process for a district alternate assessment to emerge. In 2003-05 CDE was involved in a multiple state collaborative effort (Enhanced Assessment Grant) to develop and expand various alternate assessments. The EAG project was aimed at improving the processes involved in assessing students with complex cognitive disabilities in the content areas of English language arts, mathematics, and science. Through that process a structure emerged for conducting a crosswalk process. This example of a crosswalk process from the EAG is provided in appendix A2. CDE cautions readers that the EAG was a seven-state collaborative and the time required to develop these assessments was extensive. A district would not need the amount of time required by the EAG to develop their DAAs.

Districts are encouraged to investigate the extent to which the district alternate assessment is aligned with the district standards, curriculum and the district-wide regular assessment. To determine if such an alignment exists, an alignment study is recommended concomitantly with the assessment development process or immediately after the development of the district alternate.

F. Developing Guidelines for the Use of Accommodations in Assessment

IDEA requires States to develop guidelines for the provision of appropriate accommodations for students with disabilities to access the state-wide assessment. Similarly, for a district-wide assessment IDEA requires the administrative unit or district to develop guidelines for the provision of appropriate accommodations for students with disabilities to access the district-wide assessment.

Guidelines for Accommodations should include

- A list of approved accommodations that can be allowed for students to take during the district-wide regular assessments
- Guidance on who is allowed to provide these accommodations
- Guidance on eligibility of students who will qualify for these accommodations
- A process that allows teachers and parents to clarify their questions on accommodations. This process should include procedures (with contact information of the personnel responsible) for approval or disapproval of requests for use of accommodations not identified in the guidelines. At the state level there is a similar process called the [request for consideration of a nonstandard accommodation](#) to the Colorado Department of Education which outlines such a process.
- The frequency and detail on the application of allowed accommodations
- The types of accommodations that can be used for item presentation and accommodations allowed to expand student responses
- Reference to the existence of eligibility guidelines (described on page 12) that would address participation in alternate assessments for those children who cannot participate in regular assessments even with these accommodations as indicated in their IEPs.

The accommodations document should be made available to parents, teachers, administrators and IEP team members.

Use of the CDE website that describes allowable CSAP accommodations and the [CSAP accommodations manual](#) is recommended to districts and AUs to use as a template to consider while developing the accommodations guidelines. Finally, the district-approved accommodations should not interfere with the reliability of the district-wide assessment.

G. Developing Eligibility Guidelines

IDEA 2004 requires districts to maintain eligibility guidelines for students who will take part in the district alternate assessment. These guidelines should be used to determine eligibility for students who cannot participate in the district-wide regular assessments even with the accommodations allowed by the district (according to a district's accommodations guidelines document) and as indicated in their IEPs.

Factors to consider while developing the eligibility document:

- For the statewide alternate assessment in Colorado (CSAPA), CDE provides documents entitled [CSAPA Eligibility Criteria](#) and [eligibility checklists](#) for each of the content areas in which CSAPA is administered. In accordance with IDEA and NCLB, CSAPA eligibility criteria state that only students with a severe cognitive disability are eligible for CSAPA. However, for a district alternate assessment the eligibility document should focus on participation in alternate assessments for those children who cannot participate in regular assessments as indicated in their IEPs even with the district allowed accommodations as described in the districts accommodations guidelines
- While the district is responsible for creating guidelines as to which assessment the student might take, it is still the responsibility of the IEP team to make the final determination on a student-by-student basis as to which assessment (the general district assessment or the district alternate assessment) the student should take
- The use of the [CSAPA Eligibility Criteria](#) document and the [eligibility checklists](#) as guiding documents may assist districts and AUs in developing their eligibility documents
- The eligibility guidelines should refer to the existence of guidelines for the use of accommodations (described on page 11) that the IEP team considers for allowing student participation in the regular assessments
- The eligibility guidelines should be made available to parents, teachers, administrators and IEP team members
- IN GENERAL – ALL STUDENTS with disabilities should be included in all State and district-wide assessment programs with appropriate accommodations and alternate assessments where necessary and as indicated in their respective Individualized Education Programs.

NOTE: The decision about the student's participation in the assessment should be made by the IEP team and should be based upon documented outcomes of previous assessment efforts and the student's current curriculum (general education or alternate achievement). Assumed performance levels should not be used in the determination of eligibility for any assessment.

H. Assuring Universal Design

IDEA states that districts must observe principles of Universal design in the development and administration of the alternate assessment. Universal Design for Learning (UDL) is described as a framework for designing curricula that enable all individuals to gain knowledge, skills, and enthusiasm for learning. UDL provides rich supports for learning and reduces barriers to the curriculum while maintaining high achievement standards for all.

The Center for Applied Special Technology (CAST) is a nonprofit organization that works to expand learning opportunities for all individuals, especially those with disabilities. CAST describes Universal Design for Learning (UDL) as a “blueprint for creating flexible goals, methods, materials, and assessments that accommodate learner differences”. UDL is an approach to teaching, learning and the development of curriculum and assessment that draws on current brain research and new media technologies to respond to individual learner differences. UDL curricula, teaching practices, and policies are inherently flexible and therefore may reduce the demand on educators to develop and implement modifications and accommodations to meet individual differences within general education learning environments. This framework is premised on the following principles;

1. Provide alternative formats for presenting information (multiple or transformable accessible media).
2. Provide alternative means for action and expression (write, draw, speak, switch, graphic org., etc.).
3. Provide alternative means for engagement (background knowledge, options, challenge and support)”.²

CDE mandates districts to incorporate Universal Design in all phases involved in the development of the district alternate assessments by including personnel with expertise in Universal Design in the various phases of the alternate assessment development cycle.

I. Informing Parents and IEP teams

Amendments to IDEA regulations mandate a district to provide IEP Teams with a clear explanation of the differences between assessments based on grade-level academic achievement standards and those based on modified or alternate academic achievement standards, including any effects of State or local policies on the student's education resulting from taking an alternate assessments based on alternate or modified academic achievement standards (such as whether only satisfactory performance on a regular assessment would qualify a student for a regular high school diploma). CDE recommends districts to include this information in the district's professional development efforts for IEP teams. Districts should keep all the IEP team members updated should policy changes take effect that influence the above mentioned requirements.

Idea regulations also mandate districts to ensure that parents of students selected to be assessed based on alternate or modified academic achievement standards are informed that their child's achievement will be measured based on alternate or modified academic achievement standards. The 2008 revised [Colorado IEP forms](#) allow for this provision under section 12. Districts need to be sure that parents are informed about the choice of assessment for their child and the following item on the IEP be checked.

If student is taking an alternate assessment, parents have been informed about the differences between regular and the alternate assessments (both state and district) and the effects of these, if any (including that, for students taking alternate assessments, their achievement will be measured based on alternate achievement standards).

J. Understanding Validity and Reliability

NCLB requires all assessments provided to students for achievement purposes to be valid and reliable. Authors³ of the [Writing Guide](#) at the Colorado State University define reliability and validity as mentioned below

- A. Reliability is “the extent to which an experiment, test, or any measuring procedure yields the same result on repeated trials”. Four key types of reliability, mentioned below, play a significant role in assessments.
 - 1. Equivalency Reliability is the extent to which two items measure identical concepts at an identical level of difficulty.
 - 2. Stability Reliability (sometimes called test, re-test reliability) is the agreement of measuring instruments over time.
 - 3. Internal Consistency is the extent to which tests or procedures assess the same characteristic, skill or quality.
 - 4. Inter-rater Reliability is the extent to which two or more individuals (coders or raters) agree.

- B. Validity refers to the degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure. While reliability is concerned with the accuracy of the actual measuring instrument or procedure, validity is concerned with the study's success at measuring what the researchers set out to measure. There are four key types of validity mentioned below that play a significant role in assessments.
 - 1. Face validity is concerned with how a measure or procedure appears. Does it seem like a reasonable way to gain the information the researchers are attempting to obtain? Does it seem well-designed? Does it seem as though it will work reliably?
 - 2. Criterion-related validity, also referred to as instrumental validity, is used to demonstrate the accuracy of a measure or procedure by comparing it with another measure or procedure which has been demonstrated to be valid.
 - 3. Construct validity seeks agreement between a theoretical concept and a specific measuring device or procedure.
 - 4. Content validity is based on the extent to which a measurement reflects the specific intended domain of content (Carmines & Zeller, 1991, p.20)”³.

In the case of district alternate assessments, a **Reliable** district alternate assessment would be an assessment that yields the same results on repeated trials for any given student. For the district alternate assessment to be a **Valid** assessment, it should accurately reflect or assess the specific standards against which the district-wide regular assessment is constructed to a high degree.

It is recommended that districts incorporate psychometric expertise as early as the planning phase of the development of district alternate assessments to account for the reliability and validity of the assessments they develop. Two forms of assessment approaches to meet the requirement of a district alternate assessment are described in this document. The reliability and validity involved in these two forms of assessment, along with means of assuring reliability and validity on those two forms of assessment, are described in Appendices A3 and A4.

III. Choosing an Alternate Assessment Format

Two forms of alternate assessment can be used, performance-based assessments and a body of evidence method.

A. The Performance Based Assessment

1. Overview

A Performance Based Assessment is a point-in-time measure in which the student is scored on her or his ability to demonstrate skills required to complete a task (McBrien & Bernadt, 1997).

An example of a performance based alternate assessment is the **Colorado State Assessment Program Alternate (CSAPA)**. Its score is based both on the level of support needed to perform each item and on the response provided by the student. This assessment measures how independently a student can perform activities related to Colorado Model Content Standards. The information obtained from a performance based assessment is helpful for evaluating curricular and instructional programs and can be used to determine future training and support needs within districts. At the state level, information from this assessment is also used in state calculations of Adequate Yearly Progress under federal law (No Child Left Behind)⁴.

Another example for a performance based assessment: The Enhanced Assessment Grant EAG Performance Based Alternate Assessment

One of the two assessments created by the EAG is a performance based assessment (The Performance Task –PT). These assessments were developed in collaboration with teams representing general and special education teachers. Tasks were created based on nationally recognized curricula within each content area and grade level-appropriate instructional activities implemented within general education classrooms at that grade level.

Scaffolding and Scoring

The CSAPA and the EAG performance based assessment score student performance on the basis of the level of independence the student needs in performing the activity. If the student is unable to independently perform the skill required, scaffolding procedures are outlined for administrators to follow. Scaffolding, the use of levels of prompting, is a systematic approach to providing the student with increasing levels of information about an expected response. This scaffolding/scoring methodology has been derived from the pedagogy of Vygotsky and the principal of partial participation (Brown, et. al., 1979)⁴.

2. Advantages of a Performance Based Assessment

Both the older version of the CSAPA (2001-2006) and the EAG performance based assessments allow for scripted scaffolding procedures. These levels of scaffolding are intended to support the student in reaching the correct response by increasing the specificity of information about the skill or by limiting the number of choices for

response options. Additionally, providing levels of prompting allows teachers to identify a student's instructional readiness to learn specific skills. Scaffolding also allows students to meaningfully engage in an activity and to demonstrate the required skills with the least amount of support. These skills can then be described and measured, resulting in an accurate picture of what students can do⁴. This information is intended to be helpful to educators and families to evaluate the student's current progress toward expanded benchmarks in the content areas. The process to identify the essential concepts students will need to learn and the alternate ways he/she can demonstrate learning is defined as the ***Expanded Benchmark Process***. The ***Expanded Benchmark Process*** allows IEP teams to individualize benchmarks and/or assessments to focus on the *key components of the Standards, related access skills, or any combination* necessary for a particular student to progress toward the Standard.

A logistical strength of the Performance Based Test is that in many ways it mirrors components of a general assessment. For example, it uses a set of ordered questions for each student and yields quantitative data that can be analyzed in much the same way as the general assessment. In addition, since this is the test form used for the state alternate assessment in Colorado (the CSAPA) teachers should be familiar with administration and scoring this type of assessment. Further, resource materials like the content standards, expanded benchmarks and assessment frameworks from CSAPA are readily available from CDE.

3. Limitations and Concerns

Making Performance Based Tests valid can be challenging for this population of students. Threats to validity result in part because the specific content of the test is more prescribed. That is, the test defines how a student can demonstrate Reading or Math skills, etc., in a more standardized way. Typically this would be an advantage in a testing system but for this population of students it may limit how they can demonstrate what they know. Further, the demonstration of the academic knowledge/skill is at one point in time, instead of across a longer period. The skill/s demonstrated by each student is relevant only to that particular point in time, and may or may not be generalized to daily routine (as is typical with other large scale assessments).

Reliability on performance based assessments is easily achievable in comparison to validity. The threats to reliability on performance based assessments stem from inconsistent scoring procedures, inadequate training of test providers and non-standardized testing environments.

Methods to assure validity and reliability of Performance based alternate assessments are presented in Appendix A3. A process outlining the development of the EAG performance based assessment is described in Appendix A5.

B. The Body of Evidence Assessment (BOE)

1. Overview

A *body of evidence*, as related to an alternate assessment process, is defined as a collection of information about a student's progress in the *general curriculum*. A body of evidence incorporates data from multiple sources and assessment methods, many of which are already administered, developed or gathered on an ongoing basis. Examples include IEP goals and objectives, personal communications, structured observation, student self-assessment, interviews and record reviews, performance assessments, progress reports, report cards and other sources that document student performance.

2. Advantages of a Body of Evidence Assessment

A BOE shows growth over time. It also offers other school personnel and families to have a better understanding of a student's educational progress. The purpose of the body of evidence is to provide data that document the student's individualized progress toward performance goals and standards. In addition, a body of evidence, when analyzed at the classroom, building and/or district level can:

1. Set up the process to ensure access to the general curriculum
2. Ensure that IEP goals and objectives are aligned with alternate, district, and/or state standard
3. Determine who will participate in alternate and district assessments
4. Record student progress towards goals and objectives

Further, for the population of students who will take the district alternate assessment, using this form of assessment may be more valid by letting students demonstrate their knowledge through various means over an extended period of time.

3. Limitations and Concerns

The biggest limitation to the body of evidence method is ensuring reliability in the administration of the assessment. All teachers administering the body of evidence method should be trained to follow a protocol that allows for a consistent administration method through out the district. Further, all teachers in a district should collect the same kind of evidence from students to ensure that results across students are comparable.

NOTE: CDE has provided a few examples of the various evidences that could be collected to create a body of evidence (IEP goals and objectives, personal communications, structured observations, student self-assessments, interviews and record reviews, performance assessments, progress reports, report cards and other sources that document student performance). This list is a sample; a collection of one or two examples from this list will not constitute for a valid body of evidence. For example, **IEP goals and personal communication alone will not be valid forms of evidences to collect for the BOE method because they do not measure student's growth and cannot be considered as a district alternate assessment.**

Methods to assure validity and reliability for the body of evidence method are presented in Appendix A4. A process outlining the development of the EAG – BOE based assessment, called the Instructionally Embedded Assessment, is described in Appendix A5.

C. The Three-step Guide to Developing an Alternate Assessment:

CDE recommends the following three step process to create an alternate assessment process using either the performance based method or the body of evidence method (or a combination of both) to develop a process that could measure student progress.

The Alternate Assessment Process:

An alternate assessment process can use performance based activities, a collection of evidences of student performance or a combination of both. Districts and AUs need to report on at least 10 indicators per content area when creating an alternate assessment process or an alternate assessment.

Comment [t1]: This doesn't make sense. Either "should" or "need to".

The three steps:

1. Districts/AUs develop a common assessment format utilizing 10 grade level indicators per grade and content area that will measure student performance towards content standards. The selected 10 indicators should align with the indicators assessed on the district-wide assessment. The [grade level indicators for CSAPA](#) in the areas of Reading, Writing, Math and Science from grades three through ten are available on the CDE website. When using this method, a minimum of 10 indicators should be selected from this list for each grade and content area. These indicators need to align to the district-wide assessment/s for which an alternate assessment process is created. With the selected indicators, districts and AUs can come up with a combination of performance based activities and evidences that need to be collected in a standardized manner. A district or AU could adhere to one of the methodologies (performance based or BOE) and collect data on the selected indicators.

Comment [t2]: Redundant from above, but maybe there's something in this sentence that is necessary to stay in addition to the first sentence in the paragraph.

2. A standardized scoring rubric to score student's performance on the 10 indicators should be created. If the district or AU chooses to use performance based activities, then the scoring rubric from EAG is available on the [EAG website](#) and could be used as a template. As mentioned on pages 16 and 17 of this document, one of the advantages of choosing a performance based alternate assessment system is that the teachers in the state of Colorado who typically would administer the district-alternate assessment are already trained to use a rubric because of their experience with CSAPA.

3. The district's expectations of the alternate assessment process should be clear and include reporting procedures. The district / AU needs to define to the stake holders involved in this assessment how the results from the alternate system can be used to improve the quality of instruction students receive. If a district chooses to do an alternate assessment process, then the results of all the components of the alternate assessment process need to be aggregated and reported.

D. Recommendations to Purchase or Acquire a Commercially Available Test

Purchasing a commercially available test

It is possible for districts and AUs to purchase a commercially available test to meet the requirements of the law on implementing district alternate assessments. In doing so, it is the responsibility of the district or AU to make sure that the commercially available test aligns with the district-wide regular assessment for which it serves as an alternate. The commercially available test should be available in the same grade and content areas as the district-wide regular assessment and should test on the same content standards tested by the district-wide assessments. It is also the responsibility of the district or the administrative unit to ensure that the commercially available test will meet the eligibility guidelines set forth by the district or the administrative unit.

Using the EAG Assessments

As mentioned in the previous sections of this document CDE was involved in a seven-state collaborative, federally funded grant – the EAG. Two forms of assessments (Performance task and Body of evidence) were developed in three content areas (English Language Arts, Math, and Science) in the elementary, middle school and high school level. All products of this grant, like the assessment materials, administration manual, training materials, content standards and expanded benchmarks are owned by CDE, and are available through the [CDE- EAG](#) website to the public for non-profit educational purposes. Please refer to the contact information available on page 29 to gather more information on the use of EAG products.

Pikes Peak Region DAA

In a regional effort to satisfy the IDEA District Alternate Assessment requirement, the Directors of Special Education of the Pikes Peak Region opted to collaborate in the development of an alternate assessment. In August of 2003, the region hosted a training which included Special Education Directors, District Assessment Coordinators, special educators and general educators. From this group emerged the Pikes Peak Alternate Assessment Team (PPAAT) which in the course of three years with two different region-wide pilots has developed assessments in Reading, Writing, Social Studies, Science and Math at the elementary, middle school and high school level. Statistical analyses on these assessments occurred to ensure validity and reliability of the assessments.

The PPAAT is conducting a three year study phase of their product. Participating districts received training, assessment protocols, reporting protocols and instructional decision-making protocols in the summer of 2007. Administration of alternate assessments (AA) will occur in the fall 2007. Participating districts will be required to report results to PPAAT each fall for three years. Training consists of one half day session to be offered in regional sites. For further information and view the PPAAT Newsletter www.cmsd12.org, [link to special education](#), [link to Pikes Peak Alternate Assessment](#)⁵.

The Pikes Peak Alternate Assessment process development is outlined in Appendix A6.

E. Progress Monitoring

In districts where progress monitoring is the assessment tool of choice, alternate assessments are not required as long as the tools of measurements are not standardized throughout the district and the results are not used for summative purposes. For example, if a district has various indicators and measures for eighth graders to assess reading proficiency, and if these indicators and measures are *not standardized* through out the district, then a district-wide alternate assessment is not required. However, CDE recommends that if all students of a certain grade are monitored for progress and academic achievement using a uniform tool, then this process should provide accommodations for all students needing accommodations or provide for an alternate assessment process that monitors progress for those students who cannot take part in the standardized progress monitoring tool, even with accommodations. Further, the process of progress monitoring is generally formative in nature and will not require an alternate assessment as mentioned on pages four and five. In the case where a very small district does progress monitoring, it is highly likely that the tools of measurement for progress monitoring are uniform for a given grade and content area. In this case CDE encourages the inclusion of all students through the use of accommodations.

IV. Data Management and Reporting

A. Data Collection, Storage and Analysis

The data collected on alternate assessments should be reported in a manner and format consistent with other district-wide assessments. The district has the responsibility of assuring that the data collected and reported on any district assessment (regular and alternate) is reported consistently across all of the assessments. Districts should identify the system they use to collect and analyze district-wide assessment data and ensure that secure procedures are in place to avoid any misuse or loss of data.

B. Reporting of Results (Legal Requirements)

IDEA requires districts to report results for students with disabilities (on both the district assessment and district alternate) with the same frequency and detail as students who do not have disabilities. If the districts report the results from a district-wide regular assessment on the web or through the media, then a similar process should be adopted for a district-wide alternate assessment with the consideration of FERPA regulations.

The Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99) is a Federal law that protects the privacy of student education records. The law applies to all schools that receive funds under an applicable program of the U.S. Department of Education. FERPA gives parents certain rights with respect to their children's education records⁶. The policy at the Colorado Department of Education has been to not report results when there are 15 or less students in any sub category. In situations like this CDE reports all results but places a 'X' mark in those categories where there are 15 or fewer students. Please refer to the examples from [CSAPA](#) state summary reports. It is recommended that the availability of results be acknowledged, but that the results not be reported publicly if a subcategory of students has 15 or fewer members.

If districts or AUs are using an alternate assessment process, then the results of the individual measures should be aggregated. These aggregated results need to be reported only if the results of the district wide assessments get reported.

V. Conclusion

A. Resources for developing and administering alternate assessments

Designing alternate assessments

1. The OSEP tool kit available at the National Center on Educational Outcome (NCEO) website provides an abundance of resources on the development of alternate assessments and the issues surrounding it. “These materials are designed to be used with a variety of stakeholder groups at the state and local level to engage in the construction of a coherent and effective system of instruction and assessment for students with the most significant cognitive disabilities”⁷.

http://www.osepideasthatwork.org/toolkit/ground_floor.asp

<http://www.osepideasthatwork.org/toolkit/index.asp>

2. The Enhanced Assessment Grant developed and piloted two sets of alternate assessments. One of the assessments, the performance task, is a performance based assessment, while the other is a body of evidence based assessment. The development of these two assessments and all the products of this project, including the consensus framework, expanded benchmarks, the assessment materials and training materials are available on the CDE website at <http://www.cde.state.co.us/cdesped/EAG.asp#Prod>

3. The CDE website on CSAPA provides information on administration procedure of the CSAPA along with expanded benchmarks and eligibility checklists

<http://www.cde.state.co.us/cdesped/StuDis-Sub2.asp#CSAPA>

4. The National Center on Educational Outcomes also provides an abundance of resources on alternate assessments that include but are not limited to state policies on alternate assessments, technical manuals on alternate assessments and other resources on alternate assessments.

http://education.umn.edu/NCEO/TopicAreas/AlternateAssessments/alt_assess_topic.htm

Reliability and Validity

1. The OSEP tool kit which addresses assessment reliability and validity concern and emphasizes the adequate care needed in the design and administration of assessments for students with disabilities.

http://www.osepideasthatwork.org/toolkit/tk_reliability.asp

http://www.osepideasthatwork.org/toolkit/tk_validityvidnce.asp

http://www.osepideasthatwork.org/toolkit/tk_validityargue.asp

2. The writing guide available through the Colorado State University's website provides definitions and examples of technical terms that describe the validity and reliability of alternate assessments.

<http://writing.colostate.edu/guides/research/relval/>.

Universal Design

1. As mentioned in the Assuring Universal Design Section of this manual, CAST is the national leader in Universal Design and their website is information rich in both principles and application of Universal Design.

<http://www.cast.org/>

2. NCEO and the OSEP tool kit also provide guidance on the development of Universally Designed Assessments, and describe Universal Design as Applied to Large Scale Assessments. The NCEO website offers an online manual on Universal Design.

<http://www.education.umn.edu/NCEO/OnlinePubs/StateGuideUD/default.htm>

<http://www.education.umn.edu/NCEO/UDmanual/default.html>

http://www.osepideasthatwork.org/toolkit/ta_uni_design.asp

Accommodations

1. The OSEP tool kit (Accommodations for Students with Disabilities) and the NCEO website (Online Accommodations Bibliography) offer an abundance of information on appropriate accommodations for assessment purposes.

http://education.umn.edu/NCEO/TopicAreas/Accommodations/Accom_topic.htm

http://www.osepideasthatwork.org/toolkit/ta_bibliography.asp

2. The CDE information on accommodations for the state assessment is available at http://www.cde.state.co.us/cdeassess/documents/csap/manuals/2007/CO_%20ACCOMMANUAL_0708_090707.pdf

Professional Development

1. The following websites offer resources on training teachers and administrators on alternate assessments.

http://www.osepideasthatwork.org/toolkit/tk_development.asp

2. **CSAP**

http://www.cde.state.co.us/cdeassess/index_assess.html

3. **CSAPA**

<http://www.cde.state.co.us/cdesped/StuDis-Sub2.asp>

4. *Colorado District Alternate Assessments*
<http://www.cde.state.co.us/cdesped/Distassess.asp>

5. *EAG*
<http://www.cde.state.co.us/cdesped/EAG.asp>

B. District Alternate Assessment Development and Administration Considerations

CDE has produced the following questions to provide guidance for districts and AUs who are considering the development and implementation of a district alternate assessment process. These questions will help ensure that districts comply with all requirements of the law around district alternate assessments.

Please note: These questions are not intended to be used for state monitoring purposes in their current form.

- What is the purpose of your district's large-scale assessment?
- Is the district alternate assessment aligned with the district standards, curriculum and the district-wide regular assessment?
- Does your district have a guidelines document on the use of accommodations that provides appropriate and allowable accommodations for students with disabilities to access the district-wide regular assessment?
- Does your district have a guidelines document in place to determine eligibility for students who cannot participate in regular assessments even with the accommodations allowed by the district (according to a district's accommodations guidelines document) and as indicated in their IEPs?
- Is the accommodations document made available to parents, teachers, administrators, and IEP team members?
- Is the eligibility document made available to parents, teachers, administrators and IEP team members?
- Were principles of universal design utilized in the development and administration of the alternate assessment?
- Does the alternate assessment provide the same results over repeated administrations?(i.e., is the district alternate assessment reliable?)
- Will your alternate process measure what is intended to be measured? (i.e., is the district alternate assessment valid?)
- Will your system be able to demonstrate academic growth for all students who participate in the assessments?
- Will your system provide meaningful information to parents and teachers?
- How are the district-wide regular assessment data reported?
- How will the results of the alternate be reported?
- Does your district report results for students with disabilities (on both the district assessment and district alternate) with the same frequency and detail as students who do not have disabilities?
- Did both groups (special education and assessment) come together to develop the district alternate assessment?
- In developing the assessment, were representatives with content expertise from both general and special education involved?

C. CDE Monitoring and Compliance Procedures for the Implementation of the District Alternate Assessment

Upon analysis of Colorado's FFY2005-2010 State Performance Plan, the federal Office of Special Education Programs (OSEP) identified a potential non-compliance issue related to the provision of district-wide assessments in the State of Colorado. As part of that determination, the Colorado Department of Education is required to collect and submit to OSEP data indicating the extent to which districts throughout the State are in compliance with federal statute [IDEA 2004 Section 612 (16)] regarding district-wide assessments. In order to provide the necessary information to OSEP, CDE-ESLU conducted a survey that requested for preliminary information. This survey is available in Appendix A7.

Currently, CDE monitors the provision of DAA through the Continuous Improvement Monitoring Process (CIMP). This entails a self assessment from the special education director and the gathering of data through focus groups and interviews that determines the presence or lack of compliance for the provision of DAAs. Beginning in 2008 academic year, CDE will continue this process for a limited number of districts and AUs through the CIMP process. Districts and AUs will be chosen based on the determination given to them through the ESLU's determination process. Districts and AUs in the 'needs intervention' category for two consecutive years or in the 'needs substantial intervention' category for one year will be required to have a comprehensive monitoring process. Additionally, some districts and AUs will be randomly selected to have a focused CIMP process. A monitoring process is currently under development at CDE and will be in place for the 2008 academic year. This process will also inform the presence or lack of compliance for the provision of DAA.

If it is determined that a district or AU is out of compliance it would be necessary to develop an improvement plan with strategies, proposed evidence of change, targets and timelines designed to ensure correction. Correction is required within one year of notification of non-compliance. At the state level, data submitted could be used by OSEP in making its state-level determinations with regard to compliance with IDEA and may affect future funding for Special Education Programs in Colorado.

VI. Contact Information

Sri Srinivasan, Senior Consultant
Colorado Department of Education
Exceptional Student Leadership Unit
201 East Colfax Avenue, Suite 308
Denver, CO 80203-1799
Phone: (303) 866-6732
Fax: (303) 866- 6811
E-mail: srinivasan_s@cde.state.co.us
TTY (Deaf & Hard of Hearing): (303) 860-7060

VII. References

1. A. Angelo, T. A., & Cross, K. P. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers* (2nd ed.). San Francisco : Jossey-Bass.

B. Bardes, B. & Denton, J. (2001, June). *Using the Grading Process for Departmental and Program Assessment*. Paper presented at the American Association for Higher Education Conference; Denver , CO.

C. <http://www.provost.cmich.edu/assessment/toolkit/formativesummative.htm>
2. Retrieved [05/24/06] from <http://www.cast.org/>
3. Jonathan Howell, Paul Miller, Hyun Hee Park, Deborah Sattler, Todd Schack, Eric Sperry, Shelley Widhalm, and Mike Palmquist. (2005). Reliability and Validity. Writing@CSU. Colorado State University Department of English. Retrieved [05/24/06] from <http://writing.colostate.edu/guides/research/relval/>.
4. CSAPA Administration Manual 2001 - 2006
5. Email 9/06/07 Extract from Dr, Carolena Steen, Special Education Director, ***District 12 Cheyenne Mountain School District***.
Ph: (719) 327-2823 Email: csteen@cmsd12.org
6. Retrieved [05/24/06] from FERPA
<http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html>
7. Retrieved [05/24/06] from <http://www.osepideasthatwork.org/toolkit/index.asp>
8. Enhanced Assessment Grant. <http://www.cde.state.co.us/cdesped/EAG.asp>

VII. Appendices

A1: Enhanced Assessment Grant (EAG) - The Alternate Assessment Collaborative

The Alternate Assessment Collaborative for the Enhanced Assessment Grant was composed of states and not-for-profit organizations who desire to improve how students with complex disabilities are alternately assessed in the content areas of English language arts, Mathematics, and Science. In early 2003, Colorado was awarded a \$1.7 million enhanced assessment grant to work as the lead state for the collaborative to improve a variety of alternate assessment strategies, by increasing their technical soundness, accessibility, efficiency, and feasibility for measuring adequate yearly progress. The collaborative also examined the nature of the information derived from the multiple measures. The assessments were developed, pilot tested, and analyzed during the course of the project.

States and organizations in the collaborative include: Colorado, Iowa, New Hampshire, Missouri, South Carolina, West Virginia, Wyoming, the Center for Applied Special Technologies, Inclusive Large Scale Standards and Assessment, Measured Progress, Inc., and the Research and Development Center for the Advancement of Student Learning at Colorado State University. Illinois and Oregon were involved as members of the Collaborative during various phases of the project.

Funding to support this project has come from the U.S. Department of Education Student Achievement and School Accountability Programs Enhanced Assessment Grant S368A0001 and Alternate Assessment Collaborative member states.

Development of the EAG Assessments

Each state participating in the Collaborative dedicated the content and special education expertise for constructing the elementary Science, middle school Language Arts, and high school Math Performance Tasks (PT) and the Instructionally Embedded Assessments (IEA). Six to seven general and special education teachers and at least one curriculum content expert for each of the six assessment were brought together twice during the course of the assessment development for a two day work session. The purpose of the first work session was to assist the writers in defining activities that would incorporate the concepts identified by the PLT. The second session was dedicated to revising the preliminary work of the writers, defining indicators for each step and linking these to the appropriate Expanded Benchmarks. Teams working on the PT met in Colorado and those assisting with the IEA met in Kentucky. All teams received information about the project and foundations of each assessment, consensus frameworks and expanded benchmarks, and priority concepts for each content area at the first meeting. Center for Applied Special Technology (CAST) also provided training related to the methodology of Universal Design and a set of guidelines to facilitate this thinking during the creation of each assessment. IEA team members were provided additional training on the use of the system of least prompts, an effective instructional strategy for

students with significant cognitive disabilities and an instructional methodology integrated into IEA.

After the whole group informational session, each content team divided up into separate rooms. The first meeting of these teams initially focused on an in-depth review of the concepts and related expanded benchmarks linked to the assessments that were to be developed. Teams then exchanged ideas and accessed websites associated with nationally recognized content associations for identifying grade-appropriate activities related to the concepts to be assessed. Consideration was given to tasks that could be made accessible to any student, regardless of their disability. In addition, members were asked to give thought to the standardization of the materials. Both the IEA and the PT were subjected to numerous reviews prior to finalization. Once the activities were revised based on the feedback from the teams, they were sent to state PLT members for review with the SDT. CAST reviewed the wording, materials, and adaptations of each step for fidelity to universal design. A modified alignment study based on Webb's analysis of the categorical concurrence between items and standards and the grade level expectations evident in the assessment was also conducted. Assessments were sent to content experts, identified by PLT members, who provided input regarding the alignment of the assessment indicators/steps to the expanded benchmarks and consensus frameworks.

A bias review was also conducted for both assessments. Using the format developed by Colorado, individuals representing community and disability representatives reviewed each assessment for possible areas of language, physical access, racial/ethnic, cultural, or gender bias. The combined feedback gathered was organized and given to the writers for further editing of the tasks. A beta test administration was videotaped by willing volunteer teachers from participating states and the results reviewed by writers for cohesiveness and flow of the steps, comprehensiveness of directions, manageability of materials, and item and scaffolding wording. Final drafts for all six assessments were edited by Measured Progress (test publisher) staff and finalized for publication.

A2: Development of Consensus Frameworks and Concepts

Each assessment component for the EAG was created based on a cross-state consensus framework, and this deliverable was used as the input for the entire project. Measured Progress collected the states' current academic content standards documents and listed all of the key standards by essential concepts. A comprehensive list of the essential concepts in each document was developed and sorted by the similarities and uniqueness of each framework.

A consensus framework for state academic standard was created to show all common and important unique English Language Arts, Mathematics, and Science standards among the states. This document contained a hierarchy of content standards, essential concepts within standards and grade-level benchmarks/alternate expanded benchmarks tied to these standards, with performance indicators (or alternate performance indicators) associated with each grade-level benchmark.

Measured Progress curriculum staff reviewed each framework, and found commonalities, as well as unique aspects of the collective group of standards documents.

Source Documents for the Crosswalk and Consensus Frameworks

The most recent versions of each state's content standards and Grade Level Expectations (GLEs) were gathered. Generally, these documents were found on the state education agencies' websites. In some circumstances, recent revisions were not yet posted and were sent by the Collaborate state participant. In all cases the State Educational Agency (SEA) participant verified the appropriate version to use.

Development of the Collaborative content comparison crosswalk document

In order to compare content standards and grade level expectations from the participating states, a content comparison document was built for each of the content areas of Reading, Writing, Mathematics and Science. This document was then populated with state specific GLE data. Organizational frameworks from national content organizations were used for the initial matrix that described the topics addressed by grade level. These standards frameworks were: The National Council of Teachers of English/International Reading Association Standards for English Language Arts, National Council of Teachers of Mathematics (NCTM) as revised in 2000, and the National Science Education Standards. This resulted in a Crosswalk describing where the consensus standards/topics appeared in each state's standards frameworks. The Consensus Frameworks were developed after review and reorganization of the crosswalks.

Development of Expanded Benchmarks

Once the Consensus Frameworks were organized and reviewed by the state Central and State Development Teams, CAST provided feedback to ensure that the expectations for students were described in language that allowed universal access and demonstration.

Source Documents

Collaborative states contributed their documentation on the learning expectations aligned to state standards for students with significant cognitive disabilities. These

documents included resource guides, alternate performance indicators, and expanded benchmarks. Since expanded benchmarks and frameworks were already produced for the Colorado Student Assessment Program Alternate (CSAPA) in Reading, Writing, Mathematics and Science, they served as a primary source to begin the development of the expanded benchmarks aligned to the consensus frameworks.

A3: Assuring Validity for Performance based alternate assessments

To address concerns about validity, the content of the test must be closely aligned with academic standards. For the CSAPA we call these “expanded benchmarks”, which are the foundational/fundamental concepts which lead to state model content standards. These “expanded benchmarks” are what the CSAPA assesses. Also, because of the variety of challenges this population of students present in accessing and responding to assessments, significant accommodations and modifications should be allowed on the Performance Based Test. Some examples of accommodations include the extensive use of assistive technology devices, flexibility in the length of testing sessions, scheduling, and changes to the materials presented to the student so that they may better understand, manipulate, and respond to questions and activities.

The four important measures of validity described earlier are outlined below in the light of a performance task based assessment to ensure that the district alternate assessment developers consider these measures and result in a valid performance based assessment.

A. Face Validity: Is the performance based assessment a reasonable way to collect information on student’s academic growth in the given grade and content area? Is it well designed to collect student’s academic growth?

B. Criterion Related Validity: Is the performance based assessment accurate? Does this assessment compare to another assessment that has been demonstrated to be valid to ensure accuracy?

C. Construct Validity: Do the levels of scores in a performance based method reflect an agreement between the theoretical concept of the scores and the test instrument?

D. Content Validity: To what extent does the performance based alternate assessment reflect the specific grade level content being measured by the assessment?

Assuring Reliability for Performance based alternate assessments

Reliability on a Performance Based Test is typically strong, if those administering the test give the same questions to all students taking it in the same manner (standardized administration procedure) and if there is clear guidance provided on scoring the levels of independence in performing each task. Further, test providers should be adequately trained on administering and scoring a performance based test. Standardized test administration protocols that outline the testing environment should be made available.

Four important measures of reliability are outlined below to ensure that the district alternate assessment developers consider these measures and result in a reliable performance based assessment.

A. Internal consistency: The *Internal consistency* for a performance based assessment can be described as a measure of the precision between two observers on *the same*

scoring procedure. This will help researchers interpret data from performance based tests and predict the value of scores and the limits of the relationship among variables

B. Stability Reliability: To give a measure of stability reliability when using performance based alternate assessments the assessment is repeated on the same student at a future date and the results are compared and correlated with the initial test.

C. Equivalency Reliability: In performance based alternate assessment equivalency reliability can be determined by relating two sets of test scores (from questions measuring the same concept with the identical levels of difficulty) to one another to highlight the degree of relationship or association.

D. Inter-rater Reliability: In performance based assessments inter rater reliability could be checked by video taping a teacher administering the test and then having a separate rater scoring the same student performance from observing the video tape. The extent of agreement between the teacher and the rater will provide for Inter-rater reliability. This measure will then address the consistency of the administration procedure and inform training needs on the performance based alternate assessments.

A4: Assuring Validity for the BOE method

The four important measures of validity described earlier in this chapter are outlined below in the light of a body of evidence assessment method to ensure that the district alternate assessment developers consider these measures and result in a valid body of evidence based assessment.

A. Face Validity: Is the body of evidence assessment a reasonable way to collect information on student's academic growth in the given grade and content area? Is it well-designed to collect student's academic growth?

B. Criterion Related Validity: Is the body of evidence assessment method designed accurately? Did the district compare this method with another method has been demonstrated to be valid to ensure accuracy?

C. Construct Validity: Do the levels of scores obtained from body of evidence assessment method reflect an agreement on the theoretical concept of the scores and the evidences collected?

D. Content Validity: To what extent does the collection of data through a body of evidence assessment method reflect the specific grade-level content being measured by the assessment?

If the district chooses to use a body of evidence method, then the district alternate assessment development team should decide the frequency at which the data will be collected, analyzed, and reported. In addition, to minimize subjectivity when determining a student's growth towards the standard and progress toward goals and objectives, these teams should:

1. Agree upon a student performance that would demonstrate achievement of goals and corresponding objectives.
2. Select methods to collect data on the student's performance that can be included in a body of evidence.
3. Ensure that the data collected aligns with the data collected over the district-wide regular assessment.
4. Ensure that the data collected adequately represents the academic growth for this population of students.

Assuring Reliability for the BOE method

In order to ensure the reliability of a body of evidence method the district should ensure that a test administration protocol detailing consistent methods of collection of evidence is available for any given grade and content area where the test is administered. The evidences collected should be uniform through out the district and the procedures to collect evidences should be outlined very specifically in the test administration protocol. Teachers and other test providers should be trained on these procedures. These

procedures should be available to parents, teachers and IEP team members. Periodic evaluation of the procedure and the training provided to test providers will further enhance the reliability of the body of evidence based alternate assessment method.

Four important measures of reliability are outlined below to ensure that the district alternate assessment developers consider these measures and result in a reliable body of evidence based assessment.

A. *Internal consistency:* The *internal consistency* for a body of evidence method used for assessment can be described as a measure of the precision between two or more observers *on rating the evidences*. This will help researchers interpret data from the body of evidence based tests and predict the value of scores and the limits of the relationship among variables.

B. *Stability Reliability:* When using a body of evidence approach to alternate assessments the assessment is repeated on the same student at a future date and the results are compared and correlated with the initial test to give a measure of stability.

C. *Equivalency Reliability:* When using the body of evidence approach, equivalency reliability can be determined by relating two sets of test scores (from evidences measuring the same concept with the identical levels of difficulty) to one another to highlight the degree of relationship or association.

D. *Inter-rater Reliability:* When using the body of evidence approach, inter rater reliability could be checked by video taping a teacher collecting a certain type of evidence and then having a separate rater scoring the same student performance (evidence) from observing the video tape. The extent of agreement between the teacher and the rater will provide for Inter-rater reliability. This measure will then address the consistency of the administration procedure and inform training needs on the body of evidence based alternate assessments.

A5: Development of the PT Assessment

The PT elementary Science team set about designing an assessment that was centered on the student collecting living and nonliving items, classifying these items based on characteristics and needs, and conducting a simple experiment related to the needs and characteristics of living objects. The English Language Arts (ELA) team suggested an activity where students would read a newspaper article about cooking and carry out the recipe. Ensuing discussions with the EAG management teams redirected the ELA writers toward a more grade relevant activity, since cooking is a life-skill activity that most typical middle school students do not participate in during the course of the school year. In the resulting ELA activity, students would be read a newspaper article that reported facts about the opening of a new recreational park with related maps and articles in other sections of the newspaper. Students were asked comprehension questions, directed to locate the key on the front page to find additional information and to develop a poster that would provoke the interest of their peers to attend the park.

The Math PT team considered a task of quilt-making or bridge construction that could easily get at a student's understanding of patterns. This team finally settled on the bridge construction after feedback from the whole group. Writers from the Colorado Department of Education and Measured Progress worked over the course of six weeks to construct a framework of steps for the activity. The procedures, directions and possible adaptations were drafted for the content teams to provide feedback. The following meeting was dedicated solely to reviewing and revising the steps within the activity and identifying the indicators and associated benchmarks for each step. CAST personnel were available to guide adaptations and to provide input regarding the accessibility of each step. Measured Progress personnel were able to give teams information about materials so that adjustments could be made as needed.

Development of the IEA

The prototype for the IEA was developed by staff from the University of Kentucky's Inclusive Large-scale Standards and Assessment Group (ILSSA). ILSSA staff have been involved in the development of alternate assessments particularly portfolio assessments, since 1992. The Content Development Teams from seven participating states, facilitated by ILSSA staff, developed the instructional units and lessons that are aligned to a set of consensus-framework standards approved by the participating states. These teams included general education content experts and special educators from the participating states. Standards were targeted for three content areas including elementary school Science, middle school Language Arts, and high school Mathematics. Three instructional units with assessment strategies were developed along the same grade and unit focus as the PT. A unit for elementary Science was developed to teach the concepts of living and non living using darkling beetles and mealworms. A middle school English Language Arts unit was developed to teach the features of biography and finally a high school Mathematics unit was developed to teach patterns by building fences.

The three units (Science, Language Arts, and Mathematics) embedded within the instructional assessment process have been designed within the framework of the

Universal Design for Learning (CAST, 2002) theoretical model and an instructional assessment process described by Wiggins, & McTighe (1998) in Glatthorn (1998) and adapted for use with alternate assessments by ILSSA using principles of assessment for students with moderate and severe disabilities. The theory of Universal Design for Learning incorporates multiple and flexible means of presentation, expression and apprenticeship, and multiple, flexible options for engagement.

Keeping within the Universal Design framework, multiple authentic assessment methodologies were combined to complete the “picture of performance” for a particular student. These assessment methodologies are: systematic instructional data collected by the teacher/assessor on a daily basis prior to instruction, constructed student responses from instructional activities, a comprehensive student product reflective of the entire unit coupled with a video-taped performance of the student demonstrating concepts and skills from the unit. The multiple assessment formats provide an internal consistency or triangulation of student assessment data.

To address the needs of learners who may need more specialized accommodations or extensive/pervasive supports throughout the units, both the process for determining appropriate materials and sample materials were provided. This is in recognition of the high variability among this population of students that makes it virtually impossible to identify all the types of accommodations that might be required. In an effort to assist teachers/assessors with this process, a tool “Pathways to Learning” (Denham, 2004) was included to identify support strategies for a wide variety of learning modalities for three essential response modes – Reading, Writing, and presenting information (speaking). This tool was developed as part of another grant and was used in the EAG project with permission. Assistive technology is embedded within this tool. In addition, an electronic glossary was provided with each unit. The electronic glossary provides a picture, definition, and a symbol for each essential word. Examples and non-examples of the word were also provided.

Each unit contains approximately 10 one hour lessons and mirrors the typical instructional delivery seen in general education classes. Although the lessons were designed to be taught in one hour units, they can be adapted to three 20 minute segments or six 10 minute segments or whatever time requirement meets the student’s needs. The units of study were designed so they can be taught across the continuum of service delivery models whether general education classroom or special education classroom. In addition, the lessons could be taught to individual, small groups or whole classes of students. They may even be used as content for a learning center if the content does not align with what the lessons taught concurrently in the general education classroom. It was recommended that peers participate in lessons with the assessed student(s) to the greatest extent possible.

A6: Pikes Peak Alternate Assessment Development Process

Special Education Directors from Pikes Peak region decide on a collaborative effort to develop a district alternate assessment.



A regional group of interested teachers and directors assemble to determine first steps.



Special Education and *Content Area Expert* teachers gather to construct 3 assessments (elementary, middle, and high school) for reading and math.



Volunteer teachers pilot Reading and Math assessments in the Pikes Peak region. Each teacher administers assessments and fills out a demographic and teacher feedback survey for each assessment given.



Data is collected and analyzed from pilot. Data is analyzed for item range, disability category, age, ethnic background, and participation in CSAPA.



Using data from the pilot, special education and content area experts gather again to revise Reading and Math and to develop Writing, Science, and Social Studies.



Focus on length, performance assessment item construction, and alignment with Colorado State Standards characterizes this phase of development.



Fifteen (Reading, Writing, Math, Science, and Social Studies for elementary, middle school, and high school) are piloted across the Pikes Peak region. Teachers again complete demographic survey and teacher feedback survey for each assessment given.

Comment [t3]: Fifteen what?



Qualitative and quantitative data from the pilot assessments are gathered and analyzed.



Assessments and/or items are revised to enhance their reliability and validity with special attention to scoring rubrics (for consistency) and administration directions.



Data analysis does not yield a large enough “n” for validity purposes. The decision is made to undergo a 3-year study phase with the second revision assessments to increase the “n” and address other validity issues.



All assessments undergo revision based on data analysis for fall administration to initiate the study phase.



Training phase for teachers and administrators participating in the study. Training includes history, philosophy, and practical application of the assessments along with training on administering the particular assessments. Training and administration also addresses inter-rater reliability issues.

A7: Survey used by CDE to collect Preliminary Information on District-Wide Assessments

District Name: _____

1. Does your district administer district-wide assessment(s) at any grade level K-12?

- No (*Please answer question #3 only*)
- Yes (*Please answer questions #2 - #6*)

2. Is the district-wide assessment(s) used for screening purposes or measuring academic achievement?

- Screening ~ Example: Dibels (*Please answer question #3 only*)
- Academic achievement ~ Example: Stanford 10 (*Please answer questions #4 - #6*)
- Both screening & academic achievement (*Please answer questions #4 - #6*)

3. What are your district's future plans with regard to district-wide assessment(s) that measure academic achievement?

- Do not plan on using a district-wide assessment(s) in the foreseeable future
- Considering the use of a district-wide assessment(s) in the future
- Developing a district-wide assessment(s) for future use
- Will implement district-wide assessment(s) in FY2007-2008

4. For assessments used to measure academic achievement, has the district identified allowable accommodations and made accommodations available to students as indicated on an IEP?

- No
- Under development
- Will implement in FY2007-2008
- Yes

5. For assessments used to measure academic achievement, has the district identified and made available an appropriate alternate assessment, which measures the same subjects, to students identified as needing an alternate assessment as indicated on an IEP?

- No
- Under development
- Will implement in FY2007-2008
- Yes

6. Has your district developed guidelines for the provision of alternate district-wide assessments?

- No
- Under development
- Will implement in FY2007-2008
- Yes