Alternate Assessment Design–
English Language Arts

Technical Report 2:

Current State of Alternate Assessments in English Language Arts

March 28, 2012

SRI International
Katherine M Nagle, Ph.D.
Renée Cameto, Ph.D.
Geneva Haertel, Ph.D.
Angela Haydel DeBarger, Ph.D.
Kathryn Morrison

SRI Project No.: P19474

The contents of this report were developed under a grant from the U.S. Department of Education. However, those contents do not necessarily represent the policy of the U.S. Department of Education and you should not assume endorsement by the Federal Government.

Contents

Summary 1
Introduction and Background 2
Challenges to Developing AA-AAS 4
Accounting for the Characteristics of the Tested Population 4
Overcoming a Lack of Research 10
Demonstrating Alignment between Alternate Assessments, Academic Content Standards, and Instruction 11
Future Directions in the Development of AA-AAS 15
Conclusion: The Role of ECD in Future Alternate Assessment Development 18
References 19

Summary

Alternate assessments are the means by which some students with disabilities can be represented in a broad assessment and accountability system, such as those developed by states in response to standards-based accountability reform. It is now more than 10 years ago that the Individuals with Disabilities Education Act Amendments (IDEA) of 1997 (PL 105-17) required states to develop alternate assessments for those students with disabilities who could not participate in general education assessments even with accommodations. The 2001 reauthorization of the Elementary and Secondary Education Act (PL 107-110) and the Individuals with Disabilities Education Improvement Act of 2004 (PL 108-446) reinforced IDEA'97's requirement that states develop alternate assessments, specified that such assessments were to be based on alternate achievement standards, and clarified that they were for students with significant cognitive disabilities (SCD).

Developing alternate assessments that had the technical rigor of traditional large-scale assessments was and remains a significant undertaking for many states, researchers, and assessment developers. Four specific challenges to the development and implementation of valid and reliable alternate assessments based on alternate achievement standards were identified: (1) accounting for the characteristics of the students to be tested, (2) overcoming a lack of research, (3) demonstrating alignment between alternate assessments and academic content standards, and (4) aligning classroom instruction with grade-level academic content standards. In this technical report, we discuss the challenges encountered in the development of technically adequate assessments, describe the progress made, and consider the possible trajectory of alternate assessment development in relation to the Common Core Standards in English Language Arts.

Introduction and Background

It is now more than 10 years ago since the Individuals with Disabilities Education Act Amendments (IDEA) of 1997 (PL 105-17) required states and districts to develop alternate assessments for those students with disabilities who could not participate in their general education assessments (Kleinert, Quenemoen, & Thurlow, 2010). Until this reauthorization, many students with significant cognitive disabilities (SCD) and sometimes all students with disabilities were routinely excluded from state accountability systems as a matter of policy (Kohl, McLaughlin, & Nagle, 2006; Rigney, 2009; Thompson & Thurlow, 2003). Although most students with disabilities can participate in general assessments with or without accommodations, alternate assessments were seen as a way to include students with disabilities who could not do so. Interestingly in light of subsequent policy developments, IDEA'97 did not define the population for whom an alternate assessment could be designed; nevertheless most states restricted their use to a small percentage of their students with individualized education programs (IEPs; Kearns, Towles-Reeves, Kleinert, & Kleinert, 2009). Generally the cognitive functioning of these students was thought to be below that needed for instruction and assessment using the content and achievement standards and/or formats expected of students who received the general instructional program (Kleinert, Quenemoen, et al., 2010; Quenemoen, 2008).

IDEA'97 gave states until July 1, 2000, to develop and begin conducting alternate assessments. These were no small tasks given that at the time there were few models on which to base these new assessments, and subsequently each state charted its own course. The result was a medley of alternate assessments using widely different formats that attempted to measure a mix of functional academics and life skills thought at the time to be appropriate for students with significant cognitive disabilities (Quenemoen, 2009; Tindal, 2009). These early alternate assessments varied both from each other and more decidedly from the general education assessments typically used in large-scale systems (Kearns, 2010). Tindal points out that, given the nature and scope of the problems these students faced in accessing large-scale assessments, portfolios were identified as a viable strategy. Portfolios allowed the teacher leeway in how and when to administer the assessments and also allowed for greater flexibility in how the student interacted with the assessment items and tasks. However, the flexibility and teacher control of the assessment process common in the early alternate assessments came into conflict with the need to establish reliability—a key requirement if scores were to be included in accountability systems. Overtime, state alternate assessments have included more standardized components, such as the content strands to be assessed and the amount and type of data collected (Wakeman, Browder, Flowers, & Karvonen, 2011).

The 2001 reauthorization of the Elementary and Secondary Education Act (ESEA) (PL 107-110) and the Individuals with Disabilities Education Improvement Act of 2004 (PL 108-446) caused many states to rethink the content of their alternate assessments. ESEA required states to develop alternate assessments separately in reading/English language arts (ELA), mathematics, and science; specified that they were for students with significant cognitive disabilities; and allowed states to measure the performance of these students on alternate achievement standards (Rigney, 2009). ESEA reinforced IDEA'97 and required involvement in the general curriculum by stating that all assessment options, even if they differed in respect to their academic achievement standards, were to be aligned to grade-level content standards. Finally, ESEA

required that alternate assessments, like their general assessment counterparts, be "valid and reliable for the purposes for which the assessment system is used and be consistent with relevant, nationally recognized professional and technical standards" (U.S. Department of Education, 2007, p. 8) and established standards and assessment peer review panels to confirm the technical quality of each state's assessment systems.

The Standards and Assessments Peer Review process, used by the U.S. Department of Education (USDE) to determine state compliance began under the 1994 reauthorization of ESEA. It was an ongoing process that states went through whenever they developed new assessments (Rigney, 2009). The reviewers were state assessment directors, researchers, and others selected for their expertise in assessments. According to the Government Accountability Office (GAO; Government Accountability Office, 2009), the peer review process included multiple steps. In the first step of the peer review process, a group of at least three experts known as peer reviewers examined the boxes of evidence submitted by the state to demonstrate compliance with statutory requirements, identified areas for which additional state evidence is needed, and summarized their comments. After the peer reviewers completed their review, their notes and the state's evidence was reviewed internally by an official assigned to the state, and, using the same guidelines as the peer reviewers, a recommendation on whether the state met, partially met, or did not meet each critical element and on whether the state's assessment system should be approved was made. A group of Education officials from relevant offices—including a representative from the Office of the Assistant Secretary of Elementary and Secondary Education—met as a panel to discuss the findings. The panel made a recommendation about whether to approve the state and the Assistant Secretary made the final approval decision. A state's system of standards and assessments received one of the following categories of approval: Full Approval, Full Approval with Recommendations, Approval Expected, Approval Pending, or Non-Approval (U.S. Department of Education, 2009). Afterwards a letter was sent to the state notifying it of the decision and, if the state was not approved, identifying why. States also received a copy of the peer reviewers' written comments as a technical assistance tool to support improvement.

Challenges to Developing AA-AAS

For a field described recently as "still in its infancy," (Towles-Reeves, Kearns, Kleinert, & Kleinert, 2009) the goal of developing technically sound alternate assessments and achieving USDE approval was hard to accomplish (Kettler et al., 2010). During the initial peer review panels, it quickly became apparent that many states' alternate assessments were unable to meet the necessary technical rigor, and the U.S. Department of Education extended the 2005–06 deadline to permit states to revise or rebuild their alternate assessments based on alternate achievement standards (Rigney, 2009). In a survey conducted by the GAO (2009), nearly two-thirds of the states reported that demonstrating the validity and reliability of alternate assessments based on alternate achievement standards was either moderately or very difficult. In contrast, few states reported that demonstrating either the validity or reliability of general assessments was moderately or very difficult. Furthermore, although most states met peer review expectations for validity and reliability of their general assessments, as of July 15, 2009, alternate assessments based on alternate achievement standards were a factor preventing full approval in 11 out of 12 states (GAO). Moreover, in the four states where alternate assessments based on alternate achievement standards were the only issues preventing full approval, technical quality or alignment were at the heart of the Peer's concerns.

Four specific challenges to the development of valid and reliable alternate assessments based on alternate achievement standards were identified: (1) accounting for the characteristics of the students to be tested, (2) overcoming a lack of research on developing technically strong alternate assessments, (3) demonstrating alignment between alternate assessments and academic content standards, and (4) aligning classroom instruction with grade-level academic content standards (Cameto et al., 2009; Government Accountability Office, 2009; Kearns, 2010; Kettler et al., 2010; Towles-Reeves, Kleinert, & Muhomba, 2009). These four issues provide the framework for this technical report and are discussed in relation to ELA in the following three sections. In the final section we consider the implications of the widespread adoption of the Common Core State Standards Initiative on alternate assessments in ELA and discuss a possible role for evidence-centered design (ECD) in the development of the next generation of alternate assessments. For a review of their impact on alternate assessments in mathematics, see the report by Nagle and colleagues (2011).

Accounting for the Characteristics of the Tested Population

Who participates on AA-AAS? Ensuring the validity and reliability of alternate assessments has been challenging because of the highly diverse population of students eligible to be assessed on an alternate assessment on alternate achievement standards (AA-AAS) (Browder et al., 2003; Browder, Wakeman, & Flowers, 2009; Cameto et al., 2009; Government Accountability Office, 2009; Towles-Reeves, Kearns, et al., 2009; Towles-Reeves, Kleinert, et al., 2009). Students with significant cognitive disabilities may come from any of the 13 regulatory categories included in the IDEA. Cameto, Bergland, and colleagues (2010), in findings similar to those reported by Kearns et al. (2009), found that students primarily clustered into three disability categories: mental retardation, autism, and multiple disabilities. Teachers in three states were asked to indicate the primary disability category of a randomly selected student with SCD taking the AA-AAS in their classroom or on their caseload (known as the "target student"). Forty-four percent of teachers indicated the primary disability of the student was mental retardation, 19 percent

indicated the student's disability was autism, and 18 percent indicated that the students primary disability category was multiple disabilities. An additional finding was that a majority of these students have multiple disabling conditions with 69 percent of teachers indicating that their target student had at least one other disability (Cameto, Bergland, et al., 2010). Of these teachers, 25 percent indicated that their target student had three or more additional disabilities. Over half (54 percent) indicated that the additional disability was speech/language impairment. Most teachers responded that their target student's visual, auditory, and motor functioning were normal (90 percent, 93 percent, and 79 percent respectively) and that their target student attended school over 75 percent of the time.

Students with significant cognitive disabilities (SCD) demonstrate a range of learning characteristics and response modalities that needed to be considered in developing AA-AAS (Almond & Bechard, 2005; Kearns et al., 2009; Kleinert, Kearns, & Kleinert, 2010; Towles-Reeves, Kearns, et al., 2009). For example, students with SCD may have serious and complex medical conditions that interfere with their learning and performance for others the only method of communication is eye movements or blinking (Government Accountability Office, 2009). Towles-Reeves, Kearns, and colleagues (2009) used the Learner Characteristics Inventory (LCI) to identify three distinctive subgroups of students with SCD based on their levels of communication (see Figure 1).

Figure 1. Communication Levels of Students with Significant Cognitive Disabilities

Level 1: Pre-Symbolic. Has not yet acquired the skills to discriminate between pictures or other symbols (and does not use symbols to communicate). May or not use objects to communicate. May or not use idiosyncratic gestures, sounds/vocalizations, and movements/touch to communicate with others. A direct and immediate relationship between a routine activity and the student's response may or may not be apparent. The student may have the capacity to sort very different objects, may be trial and error. Mouthing and manipulation leads to knowledge of how objects are used. May combine objects (e.g., place one block on another).

Level 2: Early symbolic. May use some symbols to communicate (e.g., pictures, logos, objects). Beginning to acquire symbols as part of a communication system. May have limited emerging functional academic skills. Representations probably need to be related to the student's immediate environment and needs.

Level 3: Symbolic. Communicates with symbols (e.g., pictures) or words (e.g., spoken words, assistive technology, ASL, home signs). May have emerging or basic functional academic skills. Emerging writing or graphic representation for the purpose of conveying meaning through writing, drawing, or computer keying.

Using the LCI, Cameto, Bergland, and colleagues (2010) reported that 68 percent of teachers indicated that their target student used symbolic language to communicate including verbal or written words, signs, Braille, or augmentative systems to conduct a range of communicative transactions and a further 20 percent of teachers reported that their student used intentional, though not symbolic, communication. A much smaller percentage of teachers (12 percent) reported that their students had no reliable or even clear communication system. In the area of receptive communication, Cameto, Bergland, et al. reported that students with SCD fell into two primary groups: those who independently followed one- or two-step directions presented through words without any additional cues (46 percent of teachers so reporting) and those that required additional cues to follow one- or two-step directions (42 percent of teachers). A smaller

percentage of teachers indicated that their target students alerted to sensory input but required physical assistance to follow simple directions (9 percent) or had uncertain responses to sensory stimuli (3 percent). These findings are consistent with earlier findings (Almond & Bechard, 2005; Towles-Reeves, Kearns, et al., 2009).

Regarding variability across skills in ELA, Cameto, Bergland and colleagues (2010) reported that 22percent of teachers stated that their target student read fluently with basic understanding from paragraphs and short passages in print or Braille. Thirty-eight percent of teachers reported that their target student read basic sight words, simple sentences and directions, bullets, or lists in print or Braille. A further 15 percent of teachers reported that their target student was aware of text or Braille, followed directionality, made letter distinctions, or told a story from pictures that were not linked to the text. However, 24 percent of teachers reported that their target student had no observable awareness of print or Braille.

Increasing test accessibility for students with SCD. Considerations of the varied learning characteristics described above will continue to influence the development of AA-AAS in ELA. Although 60 percent of students with significant disabilities have at least basic functional reading skills and at a minimum can read familiar sight words, basic sentences, directions, or bulleted lists, research has repeatedly identified a small percentage of students who have insufficient communicative resources and tools and lack even the most basic understanding print or Braille (Kearns et al., 2009; Kleinert, Kearns, et al., 2010). Academic content and subsequently its assessment, by its very nature, reside in a symbolic world, one that is difficult for students, especially those who are functioning at a presymbolic or awareness level, to enter (Kleinert, Kearns, et al., 2010). Few researchers harbor the illusion that developing AA-AAS that both accurately capture the academic skills that all students with SCD have acquired and are clearly related to grade level academic content will be easy. Schafer (2005) argues that the breadth of learning domain coverage must be sufficient to place all alternate-assessment eligible students within the content and process domain that may appear on the assessments. Until this is empirically demonstrated, a concern lingers that alternate assessments may not cover the full range of the learning domain as it must be realized for all students. The solutions are unlikely to come from a single source and will require the collaboration of individuals from several fields including psychometricians, special educators, content experts, and experts and practitioners in universal design for learning (UDL) and assistive and adaptive technology.

Test accessibility can be defined as "the extent to which a test and its constituent item set permit the test-taker to demonstrate his or her knowledge of the target construct" (Beddow, 2011, p. 381). Three avenues for supporting access to assessments for students with unique learning needs have typically been followed: the provision of individual accommodations for students with disabilities, the creation of accessible tests through the principles of UDL, and more recently the integration of evidence-centered design (ECD) and UDL principles as demonstrated in the Utah Alternate Assessment Design-Mathematics (AAD–M) project (Cameto, Haertel, DeBarger, & Morrison, 2010) and the Idaho Alternate Assessment Design-English Language Arts (AAD–ELA) project.

Provision of individual accommodations and technology based accommodations. Bolt (2011) argues that test accommodation decisions must be made carefully and with due consideration of each student's individual access needs in relation to the content and constructs being assessed by the assessment. In determining the need for test accommodations, Bolt (p. 8) recommends that

IEP teams address the following five issues: (1) the skills intended to be measured by the assessment, (2) the nature and format of test items, (3) the range of possible accommodations that might promote access to the test for students with unique difficulties, (4) the individual skills and needs of the student at the given point in time, and (5) various potential challenges related to effective administration of the selected accommodations. Bolt points out that accommodation decisions must be made at the individual student level rather than at the disability level. Students with SCD have a range of unique learning challenges that may need to be accommodated, such as visual, auditory, and motor functioning, level and mode of communication, and health and attendance issues. Consider a student who is at the early symbolic level of communication. This student may use some symbols to communicate including pictures or objects to demonstrate his or her knowledge or skills. In an assessment, this student may demonstrate his or her understanding of how characters develop over the course of a text and interact with each other by selecting the correct picture cards that show how the main characters change and interact as the plot develops. Consider the same student who also has low vision. To access the task, this student may require that the pictures be enlarged, may require additional lighting, and may even require that the pictures be accompanied by a verbal description of the picture cards.

An emerging method of providing accommodations to individual students is through the use of technology-based accommodations (Christensen, Lazarus, & Shyyan, 2011). Technology-based accommodations include a computer, communication device, speech-to-text device, and assistive technology. While the increasing use of technology has implications for how students receive accommodations in testing situations, it is of crucial importance that students learn how to make the best use of such accommodations before the testing event.

Application of UDL. A second method of increasing access to assessments is the application of UDL principles as the test is being developed (Beddow, 2011; Johnstone, Thurlow, Moore, & Altman, 2006; Russell, 2011; Thurlow et al., 2009). Principles of universal design (Mace, Hardie, & Plaice, 1991) were used to address the challenge of designing and delivering tests that are accessible to and valid for a wide range of students with SCD. Dolan and Hall (2001, 2007) proposed applying UDL so that tests would minimize potential sources of construct irrelevant variance (CIV) by supporting the ways that students with a diverse set of characteristics interact with the assessment process. Thompson, Johnstone, and Thurlow (2002) adapted Mace's original elements from architecture to derive seven elements of accessible and fair tests: "(1) inclusive assessment population; (2) precisely defined constructs; (3) accessible, nonbiased items; (4) items amendable to accommodations; (5) simple, clear, and intuitive instructions and procedures; (6) maximum readability and comprehensibility; and (7) maximum legibility" (p. 1).The concept of developing assessments to be accessible from the ground up has received encouragement in a succession of federal laws. For example, the Assistive Technology Act of 2004 ("Assistive Technology Act of 2004," 2004) and the Individuals with Disabilities Education Act of 2004 (IDEA) both require that products and services be useable by people with the widest possible range of functional capabilities. Speaking specifically to test design, the Final Regulations for ESEA 2001 (U.S. Department of Education, 2003) require that state assessments be "designed from the beginning to be valid and accessible with respect to the widest possible range of students, including students with disabilities and students with limited English proficiency."(34 C.F.R Sec. 200.2[b][2]). Although UDL has been used primarily to differentiate instruction (Van Garderen & Whittaker, 2006) researchers are incorporating its principles into testing and especially testing via the use of technology (Ketterlin-Geller, Yovanoff, & Tindal,

2007). For example, Salend (Salend, 2009, p. 42) addresses how UDL principles relate to testing and presents examples of how to implement the principles of universal design via technology-based testing. Salend argues that integrating UDL principles into technology-based testing can potentially allow for the development and administration of assessments that are accessible to students with a wide range of ability levels.

Integrating ECD and UDL principles. ECD (e.g., Mislevy, Steinberg, & Almond, 2003) uses a rigorous and replicable assessment design process that carefully considers the interaction between content, task, and learner characteristics in the creation of assessment tasks. In this approach, co-design teams bring together the expertise of assessment specialists, special educators, and content-area specialists to create design patterns, task templates, task specifications, and exemplar tasks. During the co-design process, the guidelines on UDL developed by CAST (CAST, 2011) are integrated into the task templates. The UDL guidelines, which are based on neuroscience research, address multiple means of representation (the what of learning), expression (the how of learning), and engagement (the why of learning) in development of performance tasks (see Table 1). The Idaho Alternate Assessment Design–English Language Arts (AAD–ELA) project supported by an Enhanced Assessment Grant (EAG) and the earlier Utah Alternate Assessment Design–Mathematics (AAD–M) project (Cameto, Haertel, et al., 2010) use this integrated approach.

UDL Principles	Guidelines	Implementation Examples
Principle I. Provide Multiple Means of Representation	1. Provide options for perception	• Display information in a flexible format • Provide American Sign Language (ASL) for spoken English • Use touch equivalents (tactile graphics or objects of reference) • Provide physical objects and spatial models to convey perspective or interaction
	2. Provide options for language and symbols	• Embed support for unfamiliar references within the text (e.g., idioms, academic language, figurative language, jargon, archaic language, colloquialism, and dialect) • Embed visual, nonlinguistic supports for vocabulary clarification (pictures, videos) • Provide text-to-speech software (voice recognition), human dictation, recording
	3. Provide options for comprehension	• Pre-teach critical prerequisite concepts through demonstration or models • Use cues and prompts to draw attention to critical features • Highlight previously learned skills that can be used to solve unfamiliar problems
Principle II. Provide Multiple Means of Action and Expression	4. Provide options for physical action	• Provide alternatives for physically responding or indicating • Provide alternate keyboard commands for mouse action • Build switch and scanning options for increased independent access and keyboard alternatives
	5. Provide options for expressive skills and fluency	• Provide sentence starters or sentence strips • Provide spellcheckers, grammar checkers, word prediction software
	6. Provide options for executive functions	• Embed coaches or mentors that model think-alouds of the process • Provide models or examples of the process and product of goal-setting • Embed prompts to “stop and think” before acting as well as adequate space
Principle III. Provide Multiple Means of Engagement	7. Provide options for recruiting interest	• Vary activities and sources of information so that they can be culturally relevant and responsive, age and ability appropriate, appropriate for different racial, cultural, ethnic, and gender groups • Provide tasks that allow for active participation, exploration, and experimentation
	8. Provide options for sustaining effort and persistence	• Prompt or require learners to explicitly formulate or restate goal • Display the goal in multiple ways • Encourage division of long-term goals into short-term objectives
	9. Provide options for self-regulation	• Provide differentiated models, scaffolds, and feedback for managing frustration and developing internal controls and coping skills

By embedding supports into the infrastructure and delivery mechanism of assessments it is hoped that the need for test accommodations will be reduced (Russell & Kavanaugh, 2011). However, it is unlikely that accommodated test administrations for individual students will be entirely eliminated given the need to improve the measurement of skills on large-scale assessments (Bolt, 2011). To develop assessments that are accessible for the widest population, experts from a variety of disciplines must work together to pool their respective expertise and knowledge bases. In the next section, we discuss how an initial lack of research and existing knowledge created obstacles to development of technically adequate assessments

Overcoming a Lack of Research

According to state officials and assessment experts the field assessment development initially lacked the knowledge and expertise to develop valid and reliable alternate assessments (Government Accountability Office, 2009; Tindal, 2009; Towles-Reeves, Kleinert, et al., 2009). Work by Elliott et al. (2007) and Kettler et al. (2010) has begun to fill in the gaps in the knowledge base, but considerable work remains to be done in this area. Traditionally, large-scale assessments have relied on large sample sizes and quantitative data analyses to form their validity argument—that is, that the tests measured what they purported to measure and supported the types of decisions they were intended to support. Assessment developers did not have to consider deeply the ways in which content, design, or task characteristics influence the ability of students to perform, especially students at the margins of the achievement distribution until required to develop and defend an AA-AAS.

Cameto and colleagues (2009) found that states struggled initially to present the evidence at the Department of Education's Peer Review panels that satisfactorily demonstrated that their alternate assessments were consistent with the relevant, nationally recognized professional and technical standards against which large-scale assessments were typically measured. Unsurprisingly, the technical requirements in some areas were more difficult to address than others, generally because when states had first developed their alternate assessments based on alternate assessments, they were developed by special education experts and not by assessment and content area experts. Areas related to validity that were problematic included documentation of scoring and reporting structures consistent with the subdomain structures of the state content standards, documentation of test and item scores and their relationship to internal and external variables, documentation of intended and unintended consequences, and documentation of grade-level equating (Cameto et al., 2009). However, findings from Cameto and colleagues indicated that states had an easier time documenting the validity of their alternate assessments in terms of construct relevance. Areas related to reliability that were problematic included documentation of the reliability of the alternate assessment in terms of variability across groups, in terms of internal consistency of item responses, and in reporting standard errors of measurement. However, states had an easier time documenting the reliability of their alternate assessments in terms of interrater consistency in scoring.

To help states develop and implement alternate assessments that are psychometrically appropriate, USDE provided both technical and financial assistance to states (Government Accountability Office, 2009; Rigney, 2009). For example, the USDE provided technical assistance to states in a variety of ways, including written guidance, user guides, contact with staff, and assistance from its Comprehensive Centers and Clearinghouses. Relevant program offices within the USDE provided additional assistance as needed. For example, the Office of

Special Education Programs provided assistance to states in developing alternate assessments for students with disabilities and funded the National Alternate Assessment Center. In addition, the office of Elementary and Secondary Education awarded competitive Enhanced Assessment Grants to state consortia, including the Utah Alternate Assessment Design–Mathematics project (Cameto, Haertel, et al., 2010) and the Idaho Alternate Assessment Design–ELA project, to work on a variety of assessment topics such as developing valid and reliable assessments for students with disabilities. Projects are required to share the outcomes of their projects with other states at national conferences. One promising outcome of these investments is the development of a validity-based approach for evaluating the technical quality of alternate assessments (Marion & Pellegrino, 2006; 2009). This model places consequential validity as the central validity benchmark for this type of assessment while also addressing other essential elements of technical quality related to AA-AAS.

Although much valuable research has been conducted and progress has certainly been made in the area of technical adequacy of alternate assessments (Kettler et al., 2010; Towles-Reeves, Kleinert, et al., 2009) misconceptions about the evidence for the reliability and validity of alternate assessments and the inferences based on AA-AAS results remain (Center on Education Policy, 2009; Quenemoen, Kearns, Quenemoen, Flowers, & Kleinert, 2010). Clearly, much remains to be learned.

In the next section, we describe why demonstrating alignment between alternate assessments and academic content standards and classroom instruction has been problematic. In addition, we discuss the current methods of demonstrating alignment of AA-AAS to academic content and AA-AAS to classroom instruction.

Demonstrating Alignment between Alternate Assessments, Academic Content Standards, and Instruction

Federal regulations required that AA-AAS be aligned to academic content standards, promote access to the general education curriculum, and reflect the highest achievement possible (U.S. Department of Education, 2003). In this section, we discuss the challenges to demonstrating alignment between alternate assessments and academic content standards and present some approaches used currently. In addition, we address challenges to aligning alternate assessments to classroom instruction.

Challenges to demonstrating alignment between AA-AAS and academic content standards. AA-AAS were designed to assess grade-level academic content, but with less depth, breadth, and complexity than assessments based on grade-level achievement standards. This emphasis on academic content proved problematic and results from early alignment studies indicated that many states had initially created alternate assessments based on functional curricular and life skills thought to be important for students with SCD (Lehr & Thurlow, 2003; Roach, Elliott, & Webb, 2005) rather than academic content standards.

To differing degrees, many states therefore had to transition from functional curricula to academic curricula based on state content standards and then align their alternate assessments. To do this, many states extended, prioritized, or clarified their state grade-level content standards to describe how students with SCD accessed the academic content standards (Cameto et al., 2009; Flowers, Wakeman, & Browder, 2009; Wakeman, Browder, Jimenez, & Mims, 2010). Cameto and colleagues found that 88 percent of states reported they had extended or clarified

content standards for students with SCD. States that extended their content standards needed to develop a rationale for their approach to extending, prioritizing, or clarifying their content standards (Flowers et al., 2009). States that developed content extensions needed to provide evidence demonstrating how the alternate assessment aligned with the extended content standards *and* how well the extended content standards aligned with the state's grade-level content standards.

Methods of demonstrating alignment between AA-AAS and academic content standards. The Standards for Educational and Psychological Testing, (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) known as the Standards warn, "Valid interpretation of the results in light of the standards entails assessment of the degree of fit between the test domain and contents and the descriptive statements of standards and goals" (p. 140). Assessing this degree of fit generally involves mapping the content and skills targeted by the test items to the academic content standards and strands to check that no gaps or unevenness are present. Assessing the degree of fit is usually achieved by conducting an alignment study (Browder, Wakeman, et al., 2009). There are a number of methods to demonstrate alignment between large-scale educational assessments and academic content standards, including the model developed by Norman Webb (1997), the Surveys of Enacted Curriculum (SEC) model (Porter, 2002), and the Achieve Model (Achieve, 2002). Of these three approaches, the most frequently used in K–12 was the Webb model (Karvonen, Wakeman, & Flowers, 2006; Tindal, 2009). However, these methods have had mixed results in demonstrating alignment between alternate assessments and academic content standards, and, while some researchers (Almond & Bechard, 2005; Flowers, Browder, & Ahlgrim-Delzell, 2006; Roach et al., 2005) have used approaches similar to Webb's, others have developed new alignment approaches to demonstrate alignment of alternate assessments. One such new approach is the Links for Academic Learning (LAL) model, developed by researchers from the National Alternate Assessment Center (Browder, Gibbs, et al., 2009; Flowers, Wakeman, Browder, & Karvonen, 2007; Flowers et al., 2009). The LAL model (Figure 2) has eight alignment criteria based in part on the work of Webb and Achieve, but also on best practices for students with significant cognitive disabilities and guidelines from the U.S. Department of Education (Flowers, 2009).

Figure 2. LAL Alignment Criteria

1. The content is academic and includes major domains/strands of the content area as reflected in state and national standards.
2. The content is referenced to the student's assigned grade level based on chronological age.
3. The focus of achievement maintains fidelity with the content of the original grade-level standards and, when possible, with the specified performance.
4. The content differs from grade level in range, balance, and depth of knowledge, but matches high expectations set for students with significant cognitive disabilities.
5. There is some differentiation in content across grade levels or grade bands.
6. The expected achievement for students is for students to show learning of grade-referenced academic content.
7. The potential barriers to demonstrating what students know and can do are minimized in the assessment.
8. The instructional program promotes learning in the general curriculum.

According to its developers, the LAL addresses not only the relationship between the alternate assessment and content standards, but also how state content standards are extended for students with significant disabilities and whether opportunity to learn is provided to students with SCD.

Challenges to aligning classroom instruction to academic content. Following the lead of the *Standards* (American Educational Research Association et al., 1999), which emphasized the need for alignment between what was taught in the classroom and what was tested in the assessment, ESEA required that AA-AAS promote access to the general education curriculum to afford students with SCD the opportunity to learn the content and skills tested on the AA-AAS (U.S. Department of Education, 2005). This mandate initially led to a suggestion that literacy be defined more broadly so that any individualized and nonstandard behaviors of students with SCD could be described as literate behaviors (Downing, 2005). While these behaviors serve as a useful bridge to the development of literacy, by defining them as literacy there is a concern that these emergent literacy behaviors become a goal in themselves and that students with SCD will not have the opportunity to further develop emergent and then conventional literacy skills and understandings (Erickson, Hanser, Hatch, & Sanders, 2009; Koppenhaver, 2000).

Until recently, the instructional programs for students with SCD have had little focus on literacy (Browder, Wakeman, et al., 2009; Katims, 2000; Kliewer, 1998). First, it was thought that this population need not be provided literacy instruction because they lacked the cognitive capacity to become literate (Kliewer, Biklen, & Kasa-Hendrickson, 2006). Yet, numerous case study reports provide evidence that students with SCD who are provided with comprehensive literacy instruction do demonstrate progress in learning to read, write, and communicate (e.g., Erickson, Koppenhaver, & Cunningham, 2006; Ryndak, Morrison, & Sommerstein, 1999), although much more slowly than their peers without disabilities. However, research also indicates that to make progress most students with SCD need systematic, consistent, and explicit instruction, including systematic prompting with an explicit fading procedure (Al Otaiba, 2004; Allor, Mathes, Roberts, Jones, & Champlin, 2010; Browder, Spooner, Ahlgrim-Delzell, Harris, & Wakeman, 2008; Browder, Wakeman, Spooner, Ahlgrim-Delzell, & Algozzine, 2006; Erickson et al., 2009; Land, Pugalee, Denham, & Kleinert, 2010; Wiener, 2005). For example, findings from Allor and colleagues demonstrated that students with moderate intellectual disabilities can learn basic reading skills given consistent, explicit, and comprehensive reading instruction across an extended period of time.

Second, there was some resistance on the part of the education profession to teaching academic content to students with SCD, especially when balanced against the need to learn the functional and social skills thought to be essential for this population (Agran, Alper, & Wehmeyer, 2002; Flowers, Ahlgrim-Delzell, Browder, & Spooner, 2005; Quenemoen, 2008). In several recent and comprehensive reviews of research on reading for this population, it was found that the vast majority of studies focused on functional literacy such as sight word identification (see Browder et al., 2006; Erickson et al., 2009) to the exclusion of other areas. Only a limited number related to the other components of reading identified by the 2000 National Reading Panel (National Institute of Child Health and Human Development, 2000; e.g., phonemic awareness, phonics, fluency, comprehension). The instructional programs typically provided to students with SCD concentrated on the domains of functional curricula such as community, recreation, independent living, and self-determination (Orelove & Sobsey, 1991). Any reading instruction provided was typically focused on a list of specific sight words

encountered in daily living. A typical activity-based goal that included reading was: "Tony will use tactile signing and his communication flip-card book to order lunch on each of 5 different consecutive opportunities in the school cafeteria, Wendy's, and eat-a-burger" (McDonnell, Hardman, McDonnell, & Kiefer-O'Donnell, p. 207).

There is little research on what general education content teachers of students with SCD are currently teaching (Karvonen, Wakeman, Browder, Rogers, & Flowers, 2011). Cameto, Bergland, and colleagues (2010) asked teachers in three states to describe the frequency of instruction in reading/language arts received by a target student over a 30-day period. Teachers were asked to respond at the strand level: reading and literature, writing, communication, and research. The percentage of teachers reporting instruction three or more times per week in specific reading/language arts strands ranged from 78 percent for reading and literature to 13 percent for research. Ninety-five percent of teachers reported that the target student received instruction in reading and literature at least one or two times a week; 83 percent of teachers said that the target student received instruction in writing at least one or two times a week; 92 percent of teachers said that the target student received instruction in communication at least one or two times a week; and 37 percent of teachers said that the target student received instruction in research at least one or two times a week. While there is evidence that the instructional programs of students with SCD (Rigney, 2009), Karvonen and colleagues concluded that the academic content receiving the greatest emphasis is still based on a functional or developmental early childhood approach. For example, in reading language arts the focus remains on beginning reading skills, discussion, and questioning and listening skills.

A final reason for the limited opportunity to learn academic content is the insufficient training opportunities provided to teachers on how to teach grade-level academic standards to students with SCD (Burdge, Clayton, Denham, & Hess, 2010; Flowers et al., 2005; Flowers et al., 2009; Wehman, 2001; Wehmeyer, Agran, & Hughes, 1998). The special education teacher is usually responsible for planning and providing instruction to students with SCD, so access to academic content is dependent on his or her skills and knowledge (Cameto, Bergland, et al., 2010). While teachers may have become more receptive to teaching academic content, they may not have become more capable of delivering such instruction. Cameto, Bergland and colleagues (2010) reported that over 90 percent of teachers strongly agreed or agreed that it was important that students with significant cognitive disabilities receive instruction in academic content, the same percentage reported that teaching academic standards versus other skill areas was a large or moderate challenge (Cameto, Bergland, et al., 2010). Furthermore, 94 percent of teachers indicated that meeting the instructional needs of individual students versus meeting state expectations for academic instruction was a large or moderate challenge.

Cameto, Bergland, and colleagues (2010) found that although 80 percent of teachers reported that they had taught ELA for 2 years or more, few of the teachers (11 percent) surveyed for the National Study on Alternate Assessments had any formal qualifications (such as a teaching certificate or concentration/endorsement) in teaching reading language arts.

The change from teaching functional academics to teaching grade-level academic content represents a major paradigm shift, and teachers will have an ongoing need for guidance and support in curriculum planning, locating, adapting, or developing instructional materials, and teaching academic content to this population. Cameto, Bergland, and colleagues (2010) asked teachers how many hours during the last year they had spent engaged in professional

development in instructional strategies in teaching reading language arts. Thirty-seven percent of teachers said they had received 1 to 5 hours of professional development in instructional strategies in reading/language arts and 16 percent reported they had received no professional development in this area over the past 12 months. Furthermore, 42 percent of teachers reported receiving between 1 and 5 hours in reading/language arts content standards during the last month, and 22 percent reported that they had received no professional development in mathematics content standards during the past 12 months.

The introduction of the Common Core State Standards Initiative (CCSSI) represents a shift from individual state-developed academic content standards to common academic standards. By the time the final common core state standards (CCSS) were released, 41 of the 50 states had committed to adopt them by the end of summer 2010. As of January 2012, 45 states and the District of Columbia have adopted the CCSS in English language arts and mathematics (Center on Education Policy, 2012).

Where students with significant disabilities sit in relation to the common standards would on its face not appear to be in dispute. The regulations implementing IDEA 1997 describe the term "general curriculum" as the same curriculum as that established for students without disabilities (34 C.F.R. § 300.347[a][1][i]). States that adopt the common core standards will have adopted them for all their students with IEPs, including students with SCD.

Work is already underway to explore how students with significant cognitive disabilities can be included in the CCSSI. The National Alternate Assessment Center at the University of Kentucky and the National Center for the Improvement of Educational Assessment has begun to develop learning progressions frameworks (LPFs) in mathematics, language arts, and science (Hess, 2011). During 2010 and 2011, national content experts and researchers in reading and writing met to review and synthesize the research literature about language arts and literacy learning and draft LPFs in this content domain. These individuals then articulated "enduring understandings" for each strand and also reviewed the research literature to describe what the learning targets would look like for students at grade level for these enduring understandings by the end of each grade span (K–4, 5–8, and 9–12). The grade span learning targets become progressively more complex demonstrations of learning across the grade spans for each enduring understanding. The reading and writing committees also identified and described the skills and concepts needed to achieve learning targets and reviewed the literature to develop a general order of how those skills and concepts emerge for most students. Finally, the committee members broke down the learning targets into the K–2, 3–4, 5–6, 7–8 grade spans and high school. To complete the LPF development process, the LPF progress indicators (PIs) were aligned with specific CCSS English language arts content standards.

In the next section, we discuss how the CCSS may have an impact on the development of AA-AAS and in the reading language arts in particular.

Future Directions in the Development of AA-AAS

For the states that adopt the CCSS, the implementation of these new standards will have an impact on their accountability and assessment systems. In May 2010, the U.S. Department of Education announced The Race to the Top Assessment Program, which was a $350 million grant competition to support the work of a consortium of States to develop and implement common, high-quality assessments aligned with common college- and career-ready K–12 standards. In

September 2010, (USDE), the Partnership for Assessment of Readiness for College and Careers (PARCC) and the SMARTER Balanced Assessment Consortium (SBAC) were announced as the winners of competition and were awarded the amounts of approximately $170 and $160 million respectively (Tamayo, 2010). The PARCC is a coalition of 26 states [1] and the SBAC is a coalition of 31 states.[2] The assessments will be ready for use by the 2014–15 school year. Responsibility for developing assessments for students with SCD was given to the Office of Special Education and Rehabilitative Services (OSERS) under a separate grant competition (U.S. Department of Education, 2010a) The Consortium for Citizens with Disabilities (Consortium for Persons with Disabilities, 2010) commented that this exclusion "sends a clear message that a "comprehensive assessment system" for U.S. public school children need not include all students with disabilities." On September 2, 2010, the USDE announced that it had awarded grants to two consortia of states to develop a new generation of alternate assessments for SCD (U.S. Department of Education, 2010b). The new assessments will be designed for a wide range of students with SCD and will be aligned to the CCSS. The tests will assess knowledge of mathematics and English language arts in grades 3–8, and one grade in high school. The alternate assessments are expected to align with the assessment systems being developed by the PARCC and the SBAC consortia. The grants were awarded to the National Center and State Collaborative (NSCC) partnership, a consortium of 18 states, the District of Columbia, and several territories led by the University of Minnesota and the Dynamic Learning Maps Alternate Assessment System Consortium, a consortium of 11 states led by the University of Kansas.[3]

According to the NCSC partners website,[4] the consortium will build a comprehensive assessment system that includes a well-designed summative AA-AAS that will complement the work of the Race to the Top Common State Assessment Program (RTTA) consortia. In addition, the NCSC partners will develop products and processes to support educators to plan for and provide appropriate instruction that addresses the CCSS. These supports will help individualized education program (IEP) teams accurately identify the learner characteristics of students with the most significant cognitive disabilities and make appropriate decisions about how each student participates in the overall system of assessments. The NCSC partners have created four work groups: (1) Assessment Design and Development, (2) Curriculum Design and Instruction, (3) Professional Development, and (4) Validity Evaluation. The NCSC partners are still in the design phase of their AA-AAS and have few documents available that provide insight into the design of the AA-AAS. Martha Thurlow, (2010), Director of the National Center for Educational Outcomes commented, "The ultimate purpose is to provide states with a full set of assessments, curriculum and instructional supports, and professional development tools to implement a research-based, systemic approach to improving academic outcomes for students with the most significant cognitive disabilities."

The DLMAAS group will use learning maps to show the relationships between knowledge and skills in mathematics and reading language arts. Learning maps show multiple learning

[1] AL, AR, AZ, CA, CO, DC, DE, FL, GA, IL, IN, KY, LA, MA, MD, MS, ND, NH, NJ, NY, OH, OK, PA, RI, SC and TN.

[2] AL, CO, CT, DE, GA, HI, IA, ID, KS, KY, ME, MI, MO, MT, NC, ND, NH, NJ, NM, NV, OH, OK, OR, PA, SC, SD, UT, VT, WA, WI, and WV.

[3] For more information, see http://dynamiclearningmaps.com/.

[4] See http://www.ncscpartners.org.

pathways that recognize that there are alternate ways to learn the same skill. According to the project website, learning maps include the following kinds of skills:

- *Tested subject-specific skills.* These skills include things like knowing a vocabulary word or being able to solve a multiplication problem.
- *Related precursor academic skills.* These are the underlying skills necessary to master the tested skill. For example, to solve a multiplication problem, a student first needs to understand what numbers are, be able to order numbers, and so on. For each grade-level skill that is tested, there are numerous precursor skills.
- *Communication skills.* These are skills that allow students to communicate their answers. Communication skills are not limited to speech, but instead include a variety of things like pointing or nodding.
- *Attention skills.* Before a student can show knowledge of a particular subject, the student must first be able to focus on the task or item presented.

The DLMAAS consortia will embed assessment items and tasks in day-to-day instruction so that teaching and testing occurs at the same time. This approach enables teachers to see what students know during the year and modify their instruction accordingly. DLMAAS intends to increase the accessibility of its AA-AAS through the use of assistive technology and tools such as keyboard entered responses, drag-and-drop responses that use the mouse to sort or label, and touch-screen technology. DLMAAS intends to form a coherent system with assessments developed by Race to the Top assessment competition grantees. The proposed system will measure achievement and growth of students with SCD and report on student participation and performance on AA-AAS. Other outcomes will be the creation of guidelines for IEP teams to use in determining which students should be assessed using an AA-AAS and training on those guidelines for IEP teams. Finally, the consortium will create and implement professional development for teachers in instruction and assessment of challenging academic content to show progression of student learning.

Conclusion: The Role of ECD in Future Alternate Assessment Development

The U.S. Department of Education and professional organizations such as the American Psychological Association, the National Council on Measurement in Education, and the American Educational Research Association require that assessments be aligned to grade-level content and be designed and developed following systematic and rigorous processes that are reflective of industry standards. ECD is a framework and set of processes that meet these criteria and lead to the design of families of items that are well aligned to the focal constructs of interest.

ECD uses a replicable assessment design process that can be applied to all content areas and all types of evidence, from performance tasks and portfolio activities to technology-based simulations and animations to traditional multiple-choice item formats (Cameto, Haertel, et al., 2010). The application of ECD to alternate assessment addresses validity issues as described by Shafer (2005) and Tindal et al. (2003) by applying a replicable process that makes explicit the content to be assessed, the evidence to be collected, and the features of tasks to be developed.

The work conducted during this project was innovative in two respects: the application of ECD to assessments for students with significant cognitive disabilities and the integration of UDL into ECD principles in the design of tasks for AA-AAS. A major strength of this unique approach is the support it provides for the development of items and tasks for all students that focus on construct-relevant content, minimize the impact of construct-irrelevant skills, and take into account appropriate accessibility options (Cameto, Bergland, et al., 2010). For example, in the development of reading/ELA design patterns and tasks, reading/ELA content would be targeted and the need for non-construct-relevant skills would be minimized. In addition, designers would consider supports such as use of a large font or alternate response options during item design, which would then be built into the items, thus reducing the need for teacher-selected and administered accommodations or modifications during the assessment process.

The union of ECD and UDL extends current knowledge in the field and contributes much-needed information for improving AA-AAS in mathematics. Moreover, it has demonstrated that ECD and UDL have the potential to improve assessment practices generally across the ability spectrum and specifically for students with significant cognitive disabilities. The use of ECD can enhance the quality of assessments for students with SCD, improve the efficiency with which future assessments are developed, and capture the myriad design decisions required when developing a valid assessment of student learning (Mislevy et al., 2003).

The assessments developed by the four consortia, and offered mostly online, will replace the current state tests given to millions of students each year in reading and math by 2014–15. The new generation of AA-AAs will be the means by which students with SCD can be represented in a broad assessment and accountability system. Developing AA-AAS that have the same technical rigor as traditional large-scale assessments remains a significant undertaking. Four challenges to the development of valid and reliable alternate assessments based on alternate achievement standards were identified: (1) accounting for the characteristics of the students to be tested, (2) overcoming a lack of research, (3) demonstrating alignment between alternate assessments and academic content standards, and (4) aligning classroom instruction with grade-level academic content standards**.** In this technical report, we have discussed these challenges, described the progress made, and considered the impact of the CCSS in reading/English language arts and how ECD may play a role in the development of the next generation of AA-AAS.

References

Achieve. (2002). *Measuring up: A standards and assessment benchmark report for Oklahoma*. Washington, DC: Author. Retrieved from http://www.achieve.org/files/OK_Benchmark.pdf

Agran, M., Alper, S., & Wehmeyer, M. L. (2002). Access to the general curriculum for students with significant disabilities: What it means for teachers. *Education and Training in Mental Retardation and Developmental Disabilities, 37*, 123-133.

Al Otaiba, S. (2004). Providing effective literacy instruction to students with Down syndrome. *TEACHING Exceptional Children, 36*(4), 28-35.

Allor, J. H., Mathes, P. G., Roberts, J. K., Jones, F. G., & Champlin, T. M. (2010). Teaching students with moderate intellectual disabilities to read. An experimental examination of a comprehensive reading intervention. *Education and Training in Autism and Developmental Disabilities, 45*(1), 3-22.

Almond, P., & Bechard, S. (2005). *Alignment of two performance-based alternate assessments with combined content standards from eight states through expanded benchmarks*. Paper presented at the 2005 meeting of the National Council on Measurement in Education, Montreal, Quebec, Canada

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.

Assistive Technology Act of 2004, Pub. L. No. 108-364 (2004).

Beddow, P. A. (2011). Beyond universal design: Accessibility theory to advance testing for all students. In M. Russell & M. Kavanaugh (Eds.), *Assessing students in the margin: Challenges, strategies, and techniques* (pp. 381-406). Charlotte, NC: Information Age Publishing.

Bolt, S. E. (2011). Factors to consider in providing appropriate test accommodations to individual students with disabilities. In M. Russell & M. Kavanaugh (Eds.), *Assessing students in the margin: Challenges, strategies, and techniques* (pp. 3-30). Charlotte, NC: Information Age Press.

Browder, D., Gibbs, S., Ahlgrim-Delzell, L., Courtade, G. R., Mraz, M., & Flowers, C. (2009). Literacy for students with severe developmental disabilities. What should we teach and what should we hope to achieve? *Remedial and Special Education, 30*(5), 269-282. doi: 10.1177/0741932508315054

Browder, D., Spooner, F., Ahlgrim-Delzell, L., Flowers, C., Algozzine, B., & Karvonen, M. (2003). A content analysis of the curricular philosophies reflected in states' alternate assessment performance indicators. *Research and Practice for Persons with Severe Disabilities, 28*(4), 165–181.

Browder, D. M., Spooner, F., Ahlgrim-Delzell, L., Harris, A. A., & Wakeman, S. (2008). A meta-analysis on teaching mathematics to students with significant cognitive disabilities. *Exceptional Children, 74*(4), 407-432.

Browder, D. M., Wakeman, S., & Flowers, C. (2009). Alignment of alternate assessments with state standards. In W. D. Schafer & R. W. Lissitz (Eds.), *Alternate assessment based on alternate achievement standards: Policy, practice, and potential* (pp. 61-92). Baltimore, MD: Paul H. Brookes.

Browder, D. M., Wakeman, S. Y., Spooner, F., Ahlgrim-Delzell, L., & Algozzine, B. (2006). Research on reading instruction for individuals with significant cognitive disabilities. *Exceptional Children, 72*(4), 392-408.

Burdge, M., Clayton, J., Denham, A., & Hess, K. K. (2010). Ensuring access: A four step process for accessing the general curriculum. In H. L. Kleinert & J. Farmer-Kearns (Eds.), *Alternate assessment for students with significant cognitive disabilities: An educator's guide* (pp. 109-148). Baltimore, MD: Paul H. Brookes.

Cameto, R., Bergland, F., Knokey, A.-M., Nagle, K. M., Sanford, C., Kalb, S. C., . . . Ortega, M. (2010). *Teacher perspectives of school level implementation of alternate assessments for students with significant cognitive disabilities. A report from the National Study on Alternate Assessments* (NCSER 2010-3007). Menlo Park, CA: SRI International. Retrieved from http://ies.ed.gov/ncser/pubs/20103007/pdf/20103007.pdf

Cameto, R., Haertel, G., DeBarger, A., & Morrison, K. (2010). *Applying evidence-centered-design to alternate assessments in mathematics for students with significant cognitive disabilities* (Alternate assessment design–mathematics technical report 1: Project overview). Menlo Park, CA: SRI International. Retrieved from http://alternateassessmentdesign.sri.com/documents/AAD_M_TechRpt1.pdf

Cameto, R., Knokey, A.-M., Nagle, K., Stanford, C., Blackorby, J., Sinclair, B., & Riley, D. (2009). *National profile on alternate assessments based on alternate achievement standards. A report from the National Study on Alternate Assessments* (NCSER 2009-3014). Menlo Park, CA: SRI International.

CAST. (2011). *Universal design for learning guidelines version 2.0*. Wakefield, MA: Author.

Center on Education Policy. (2009). *Has progress been made in raising achievement for students with disabilities?* Washington, DC: Author. Retrieved from www.cep-dc.org/index.cfm?fuseaction=Page.viewPage&pageId=496&parentID=481

Center on Education Policy. (2012). *States progress and challenges in implementing common core state standards*. Washington, DC: Author. Retrieved from http://www.cep-dc.org/index.cfm?DocumentTopicID=1

Christensen, L. L., Lazarus, S. S., & Shyyan, V. (2011). Ever-evolving state policies and practices for using technology-based accommodations to improve assessment outcomes for student in the margins. In M. Russell & M. Kavanaugh (Eds.), *Assessing students in the*

margin: Challenges, strategies, and techniques (pp. 127-148). Charlotte, NC: Information Age Press.

Consortium for Persons with Disabilities. (2010). *CCD Education Task Force: Race to the Top (RTTT) notice inviting applications (NIA) for comprehensive assessment systems grants published in the Federal register Vol. 75, No. 68 on April 9, 2010*. Retrieved from http://www.c-c-d.org/task_forces/education/tf-education.htm

Dolan, R. P., & Hall, T. E. (2001). Universal Design for Learning: Implications for large-scale assessment. *IDA Perspectives, 27*(4), 22-25.

Dolan, R. P., & Hall, T. E. (2007). Developing accessible tests with universal design and digital technologies: Ensuring we standardize the right things. In L. L. Cook & C. C. Cahalan (Eds.), *Large-scale assessment and accommodations: What works* (pp. 95-111). Arlington, VA: Council for Exceptional Children.

Downing, J. E. (2005). *Teaching literacy to students with significant disabilities*. Baltimore, MD: Brookes Publishing.

Elliott, S. N., Compton, E., & Roach, A. T. (2007). Building validity evidence for scores on a statewide alternative assessment: A contrasting groups, multi-methods approach. *Educational Measurement: Issues and Practice, 26*(2), 30-43.

Erickson, K., Hanser, G., Hatch, P., & Sanders, E. (2009). *Research based practices for creating access to the general curriculum in reading and literacy for students with significant intellectual disabilities*. Chapel Hill, NC: Center for Literacy and Disabilities Studies.

Erickson, K. A., Koppenhaver, D. A., & Cunningham, J. W. (2006). Balanced reading. Intervention in augmentative communication. In R. McCauley & M. Fey (Eds.), *Treatment of language disorders in children: Conventional and controversial interventions* (pp. 309-345). Baltimore, MD: Paul H. Brookes.

Flowers, C. (2009). Does the emperor have new clothes? In W. D. Schafe & R. W. Lissitz (Eds.), *Alternate assessment based on alternate achievement standards: Policy, practice and potential* (pp. 347-350). Baltimore, MD: Paul H. Brookes.

Flowers, C., Ahlgrim-Delzell, L., Browder, D. M., & Spooner, F. (2005). Teachers' perceptions of alternate assessment. *Research and Practice for Persons with Severe Disabilities, 30*(2), 81-92.

Flowers, C., Browder, D. M., & Ahlgrim-Delzell, L. (2006). An analysis of three states' alignment between language arts and mathematics standards and alternate assessments. *Exceptional Children, 72*(2), 201-215.

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

Flowers, C., Wakeman, S., & Browder, D. M. (2009). Links for Academic Learning (LAL): A conceptual model for investigating alignment of alternate assessments based on alternate achievement standard. *Educational Measurement: Issues and Practice, 28*(1), 25-37.

Government Accountability Office. (2009). *No Child Left Behind Act: Enhancements in the Department of Education's peer review process could improve state academic assessments*. Washington, DC: Author. Retrieved from http://www.gao.gov/new.items/d09911.pdf

Hess, K. K. (2011). *Learning progressions frameworks designed for use with the Common Core State Standards in English Language Arts and literacy K-12*. Dover, NH: National alternate Assessment Center at the University of Kentucky and the National Center for the Improvement of Educational Assessment. Retrieved from http://www.naacpartners.org/

Johnstone, C., Thurlow, M., Moore, M., & Altman, J. (2006). *Using systematic item selection methods to improve universal design of assessments* (Policy Directions 18). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://education.umn.edu/NCEO/OnlinePubs/Policy18/

Karvonen, M., Wakeman, S., & Flowers, C. (2006). *Alignment of standards, large scale assessments and curriculum: A review of the methodological and empirical literature*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. Retrieved from http://www.naacpartners.org/products.aspx

Karvonen, M., Wakeman, S. Y., Browder, D. M., Rogers, M. A. S., & Flowers, C. (2011). *Academic curriculum for students with significant cognitive disabilities: Special education teacher perspectives a decade after IDEA 1997*. (ERIC Document Reproduction Service No. ED521407).

Katims, D. (2000). Literacy instruction for people with mental retardation: Historical highlights and contemporary analysis. *Education and Training in Mental Retardation and Developmental Disabilities, 35*(1), 3-15.

Kearns, J. (2010). Principles and practices for achievement assessments in school accountability systems. In H. L. Kleinert & J. Farmer-Kearns (Eds.), *Alternate assessment for students with significant cognitive disabilities: An educator's guide* (pp. 19-40). Baltimore, MD: Paul H. Brookes Publishing.

Kearns, J., Towles-Reeves, E., Kleinert, H. L., & Kleinert, J. (2009). Who are the children who take alternate achievement standards assessments? In W. D. Schafer & R. W. Lissitz (Eds.), *Alternate assessment based on alternate achievement standards: Policy, practice, and potential* (pp. 3-22). Baltimore, MD: Paul H. Brookes Publishing.

Ketterlin-Geller, L. R., Yovanoff, P., & Tindal, G. (2007). Developing a new paradigm for conducting research on accommodations in mathematics testing. *Exceptional Children, 73*, 331-347.

Kettler, R. J., Elliott, S. N., Beddow, P. A., Compton, E., McGrath, D., Kaase, K. J., . . . Hinton, K. (2010). What do alternate assessments of alternate academic achievement standards measure? A multitrait-multimethod analysis. *Exceptional Children, 76*(4), 457-474.

Kleinert, H. L., Quenemoen, R. F., & Thurlow, M. (2010). An introduction to alternate assessments: Historical foundations, essential parameters, and guiding principles. In H. L. Kleinert & J. Farmer-Kearns (Eds.), *Alternate assessment for students with significant cognitive disabilities: An educator's guide* (pp. 3-18). Baltimore, MD: Paul H. Brookes Publishing.

Kleinert, J., Kearns, J., & Kleinert, H. L. (2010). Students in the AA-AAS and the importance of communicative competence. In H. L. Kleinert & J. Farmer-Kearns (Eds.), *Alternate assessment for students with significant cognitive disabilities: An educator's guide* (pp. 41-74). Baltimore, MD: Paul H. Brookes Publishing.

Kliewer, C. (1998). *Schooling children with Down syndrome*. New York, NY: Teachers College Press.

Kliewer, C., Biklen, D., & Kasa-Hendrickson, C. (2006). Who may be literate? Disability and resistance to the cultural denial of competence. *American Educational Research Journal, 43*, 163-192.

Kohl, F. L., McLaughlin, M. J., & Nagle, K. M. (2006). Alternate achievement standards and assessments: A descriptive investigation of 16 states. *Exceptional Children, 73*(1), 107-123.

Koppenhaver, D. A. (2000). Literacy in AAC: What should be written on the envelope we push? *Augmentative and Alternative Communication, 16*, 267-277.

Land, L., Pugalee, D. K., Denham, A., & Kleinert, H. L. (2010). Math instruction and assessment linked to grade-level standards. In H. L. Kleinert & J. Farmer-Kearns (Eds.), *Alternate assessment for students with significant cognitive disabilities: An educator's guide* (pp. 173-202). Baltimore, MD: Paul H. Brookes Publishing.

Lehr, C., & Thurlow, M. (2003). *Putting it all together: Including students with disabilities in assessment and accountability systems*: National Center on Educational Outcomes. Retrieved from http://www.cehd.umn.edu/NCEO/OnlinePubs/PolicyDirections16.pdf

Mace, R., Hardie, G., & Plaice, K. (1991). Accessible environments: Towards universal design. In W. E. Preiser, J. C. Visher, & E. T. White (Eds.), *Design interventions towards a more humane architecture*. New York, NY: Van Nostrant Reinhold.

Marion S. & Pellegrino, J. (2009, April). *Validity framework for evaluating the technical quality of alternate assessments based on alternate achievement standards.* Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL. (PDF : 102 Kb)

Marion, S. F., & Pellegrino, J. W. (2006). A validity framework for evaluating the technical quality of alternate assessments. *Educational Measurement: Issues and Practice, 25*(4), 47-57.

McDonnell, J. J., Hardman, M. L., McDonnell, A. P., & Kiefer-O'Donnell, R. (1995). *An introduction to persons with severe disabilities: Educational and social issues.* Boston: Allyn & Bacon.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3-67.

Nagle, K. M., Cameto, R., Haertel, G., Haydel DeBager, A., & Morrison, K. (2011). *Alternate assessment design-mathematics technical report 2: Current state of mathematics assessment in alternate assessment*. Menlo Park, CA: SRI International.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Orelove, F. P., & Sobsey, D. (1991). *Educating children with multiple disabilities: A transdisciplinary approach*. Baltimore, MD: Paul H. Brookes Publishing

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14.

Quenemoen, R. F. (2008). *A brief history of alternate assessments based on alternate achievement standards* (Synthesis Report 68). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Quenemoen, R. F. (2009). The long and winding road of alternate assessments: Where we started, where we are now, and the road ahead. In W. D. Schafer & R. W. Lissitz (Eds.), *Alternate assessments based on alternate achievement standards: Policy, practice, and potential* (pp. 127-156). Baltimore, MD: Paul H. Brookes Publishing.

Quenemoen, R. F., Kearns, J., Quenemoen, M., Flowers, C., & Kleinert, H. (2010). *Common misperceptions and research-based recommendations for alternate assessment based on alternate achievement standards* (Syntheses Report 73). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://www.cehd.umn.edu/NCEO/OnlinePubs/Synthesis73/Synthesis73.pdf

Rigney, S. L. (2009). Public policy and the development of alternate assessments for students with cognitive disabilities. In W. D. Schafer & R. W. Lissitz (Eds.), *Alternate assessments based on alternate achievement standards. Policy, practice, and potential* (pp. 41-60). Baltimore, MD: Paul H. Brookes Publishing.

Roach, A. T., Elliott, S., & Webb, N. L. (2005). Alignment of an alternate assessment with state academic standards: Evidence of content validity of the Wisconsin alternate assessment. *The Journal of Special Education, 38*(4), 218-231.

Russell, M. (2011). Reflecting on milestones and mapping a route for the future. In M. Russell & M. Kavanaugh (Eds.), *Assessing students in the margin: Challenges, strategies, and techniques* (pp. 465-468). Charlotte, NC: Information Age Press.

Russell, M., & Kavanaugh, M. (Eds.). (2011). *Assessing students in the margin: Challenges, strategies, and techniques*. Charlotte, NC: Information Age Publishing.

Ryndak, D. L., Morrison, A. P., & Sommerstein, L. (1999). Literacy before and after inclusion in general education settings: A case study. *Journal of the Association for Persons with Severe Handicaps, 24*, 5-22.

Salend, S. (2009). Using technology to create and administer accessible test. *TEACHING Exceptional Children, 41*(3), 40-51.

Schafer, W. D. (2005). Technical documentation for alternate assessment. *Practical Assessment, Research and Evaluation, 10*(10), 1-9.

Shafer, W. D. (2005). Technical documentation for alternate assessment. *Practical Assessment, Research and Evaluation, 10*(10), 1-9.

Tamayo, J. R. (2010). *Assessment 2.0: Next generation comprehensive assessment systems. An analysis of the proposal by the Partnership for the Assessment of Readiness for College and Careers and SMARTER Balanced Assessment Consortium*. Washington, DC: Aspen Institute. Retrieved from http://www.aspeninstitute.org/publications/assessment-20-next-generation-comprehensive-assessment-systems

Thompson, S., & Thurlow, M. (2003). *2003 State special education outcomes: Marching on*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://www.education.umn.edu/NCEO/OnlinePubs/2003StateReport.htm

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: National Center on Educational Outcomes. Retrieved from http://education.umn.edu/NCEO/OnlinePubs/Synthesis44.html

Thurlow, M. (2010). U of M's NCEO receives $45 million to form new national collaborative partnership for students with the most significant cognitive disabilities. Retrieved from http://www1.umn.edu/news/news-releases/2010/UR_CONTENT_257101.html

Thurlow, M. L., Moen, R. E., Liu, K. K., Scullin, S., Hausmann, K. E., & Shyyan, V. (2009). *Disabilities and reading: Understanding the effects of disabilities and their relationship to reading instruction and assessment*. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Tindal, G. (2009). Sampling the domain for evidence of achievement. In W. D. Schafer & R. W. Lissitz (Eds.), *Alternate assessment based on alternate achievement standards: Policy practice, and potential* (pp. 33 - 41). Baltimore, MD: Paul H. Brookes Publishing

Tindal, G., McDonald, M., Tedesco, M., Glasgow, A., Almond, P., Crawford, L., & Hollenbeck, K. (2003). Alternate assessments in reading and math: Development and validation for students with significant disabilities. *Exceptional Children, 69*(4), 481-494.

Towles-Reeves, E., Kearns, J., Kleinert, H., & Kleinert, J. (2009). An analysis of the learning characteristics of students taking alternate assessments based on alternate achievement standards. *The Journal of Special Education, 42*(4), 241-254. doi: 10.1177/0022466907313451

Towles-Reeves, E., Kleinert, H., & Muhomba, M. (2009). Alternate assessment: Have we learned anything new? *Exceptional Children, 75*(2), 233-252.

U.S. Department of Education. (2003). *Title I: Improving the academic achievement of the disadvantaged*. Final rule 68, Fed. Reg. 236 Washington, DC: Retrieved from http://www2.ed.gov/legislation/FedRegister/finrule/2003-4/120903a.html.

U.S. Department of Education. (2005). *Alternate achievement standards for students with the most significant cognitive disabilities. Non-regulatory guidance*. Washington, DC: Author. Retrieved from http://www.ed.gov/policy/elsec/guid/altguidance.doc

U.S. Department of Education. (2007). *Standards and assessment peer review guidance. Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Washington, DC: Author. Retrieved from www2.ed.gov/policy/elsec/guid/saaprguidance.doc

U.S. Department of Education. (2010a). *Race to the Top, Notice Inviting Applications, Federal Register/Vol. 75, No. 68*. Retrieved from http://edocket.access.gpo.gov/2010/pdf/2010-8176.pdf.

U.S. Department of Education. (2010b). *U.S. Secretary of Education Duncan announces winners of competition to improve student assessment* (Press release). Author. Retrieved from http://www.ed.gov/news/press-releases/us-education-department-awards-grants-improv-assessments-students-disabilities

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2009). *State and local implementation of the No Child Left Behind Act, Volume V-Implementation of the 1 percent rule and the 2 percent interim policy options*. Washington, DC: Author. Retrieved from http://www2.ed.gov/rschstat/eval/disadv/nclb-disab/nclb-disab.pdf

Van Garderen, D., & Whittaker, C. (2006). Planning differential multicultural instruction for secondary inclusive classrooms. *TEACHING Exceptional Children, 38*(3), 12-20.

Wakeman, S. Y., Browder, D. M., Flowers, C., & Karvonen, M. (2011). Alternate achievement standards for alternate assessments: Considerations for policy and practice. In M. Russell &

M. Kavanaugh (Eds.), *Assessing students in the margin. Challenges, strategies, and techniques* (pp. 149-170). Charlotte, NC: Information Age Publishing.

Wakeman, S. Y., Browder, D. M., Jimenez, B. A., & Mims, P. J. (2010). Aligning curriculum with grade specific content standards: Using eight criteria to create access. In H. L. Kleinert & J. Farmer Kearns (Eds.), *Alternate assessment for students with significant cognitive disabilities: An educator's guide* (pp. 77-105). Baltimore, MD: Paul H. Brookes Publishing.

Webb, N. L. (1997). *Criteria for alignment of expectation and assessments in mathematics and science education* (Research Monograph No. 6). Washington, DC: Council of Chief State School Officers, National Institute for Science Education, University of Wisconsin-Madison. Retrieved from http://facstaff.wcer.wic.edu/normw/WEBBMonograph6criteria.pdf

Wehman, P. (2001). *Life beyond the classroom: Transition strategies for young people with disabilities*. Baltimore, MD: Paul H. Brookes Publishing.

Wehmeyer, M. L., Agran, M., & Hughes, C. (1998). *Teaching self-determination to students with disabilities: Basic skills for successful transition*. Baltimore, MD: Paul H. Brookes Publishing.

Wiener, D. (2005). *One state's story: Access and alignment to the GRADE-LEVEL content for students with significant cognitive disabilities* (Synthesis Report 57). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.