SRI International

Alternate Assessment Design–English Language Arts/Reading

Technical Report 1: Project Overview

Applying Evidence-Centered Design to Alternate Assessments in English Language Arts/Reading for Students with Significant Cognitive Disabilities

November 28, 2011

SRI International
Renée Cameto
Geneva Haertel
Angela Haydel DeBarger
Kathryn Morrison

SRI Project No. P19474

**Alternate Assessment Design–English Language Arts/Reading Technical Report 1:
Project Overview**

**Applying Evidence-Centered Design to Alternate Assessments in
English Language Arts/Reading for Students with
Significant Cognitive Disabilities**

Students with significant cognitive disabilities challenge conventions with respect to the teaching, learning, and assessing of academic content. Assessment has been instrumental in changing expectations for these students' learning and in turn is beginning to influence classroom instructional practices. Assessment designers are challenged to develop assessments that adequately and reliably show what these students know and can do. The sheer variability in this target population, the assumptions about measuring their achievement, and the variability of design implementation procedures make traditional assessment design approaches inapplicable without some reformulation (Gong & Marion, 2006; Government Accountability Office, GAO, 2009; Ryan, Quenemoen, & Thurlow, 2004). The methods used to date in designing alternate assessments and selecting their content are varied but typically do not match the technical rigor used in designing general education assessments (Bechard, 2005). The Alternate Assessment Design–English Language Arts/Reading (AAD-ELA/R) project is one of the first to address systematically the specification of grade-level academic content for alternate assessments of students with significant cognitive disabilities through the application of evidence-centered design (ECD) and the principles of universal design for learning (UDL).

ECD directly addresses these most pressing issues by using 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. The use of ECD can enhance the quality of assessments and improve the efficiency with which future assessments are developed while documenting the myriad design decisions required when developing a valid assessment of student learning (Mislevy, Steinberg, & Almond, 2003). The AAD-ELA/R project is innovative in two aspects: It is applying ECD to assessments for students with significant cognitive disabilities, and it is integrating ECD and UDL approaches in the design of tasks for alternate assessments based on alternate academic achievement standards (AA-AAS). This work extends current knowledge in the field and provides a prototype for future alternate assessment development.

Utah, Idaho, and Kansas have formed a consortium with SRI International to improve their AA-AAS using ECD to design and develop assessment tasks that are linked to Common Core State Standards in English language arts/reading. In this report, we describe

- The project goals and activities

- The development of assessments for accountability purposes for students with significant cognitive disabilities

- ECD and UDL frameworks and how they are applied through a co-design process

- Our plan to produce a series of technical reports, including procedural guidelines, design documents, and associated sample assessment tasks

- Our dissemination plan including the project website, www.alternateassessmentdesign.sri.com

History

A succession of federal laws, including the Individuals with Disabilities Education Act (IDEA) of 1997 and 2004 and the 2001 reauthorization of the Elementary and Secondary Education Act (ESEA), require that *all* students be assessed in reading/language arts, mathematics, and science and be included in state accountability systems. Most students with disabilities participate in general assessments even with accommodations, but some students, including those with significant cognitive disabilities, may need alternate ways to access assessments. To include these students in educational accountability systems, all states have developed alternate assessments based on alternate achievement standards (Kohl, McLaughlin, & Nagle, 2006; Thompson & Thurlow, 2003). However, states have faced a number of challenges, including (1) clearly documenting the links between their general education content standards and their alternate assessments, (2) developing a clear rationale for their choice of particular content standards in their alternate assessments, and (3) providing strong evidence that the intended assessment content is actually being assessed, as called for by Flowers, Wakeman, Browder, and Karvonen (2007) in *Links for Academic Learning.* Although valuable work has been conducted in the area of technical adequacy of alternate assessments (for example, by the New Hampshire Enhanced Assessment Initiative and the National Alternate Assessment Center), the reliability and validity of alternate assessments remain problematic and complete confidence cannot be placed in the results of such tests (GAO, 2009; Quenemoen, 2008; Quenemoen, Kearns, Quenemoen, Flowers, & Kleinert, 2010). A compelling need exists for well-designed evidence-based AA-AAS to measure and document the performance of students with significant cognitive disabilities.

Federal education laws enacted during the past decade have produced a frenetic pace of change in alternate assessments and generated a marked shift to the full inclusion of students with significant cognitive disabilities in accountability systems across the states, accompanied by a shift in instructional emphasis from functional skills to academic content (Thompson, Johnstone, Thurlow, & Altman, 2005). A review conducted by Quenemoen (2008) indicated that states use several different approaches when gathering information on the performance and progress of these students. These approaches include rating scales, portfolios, performance tasks, multiple choice, or a blend of multiple formats (Cameto et al., 2009). These, in turn, are implemented with varying degrees of local decisionmaking, Individualized Education Plan team involvement, scoring, and criteria for inclusion in calculations for adequate yearly progress (Cameto et al., 2009). Design and implementation of alternate assessments are in considerable flux (for example, assessment consortia are designing and developing assessments that will align with the Common Core State Standards Initiative[1]).

The *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on

---

[1] The Common Core State Standards Initiative is a state-led effort coordinated by the National Governors Association Center for Best Practices (NGA Center) and the Council of Chief State School Officers (CCSSO). Governors and state commissioners of education from 48 states, two territories, and the District of Columbia developed a common core of state standards in English/language arts and mathematics for grades K–12.

Measurement in Education, 1999) continues to be the authoritative source of information on test validity. Several standards are particularly relevant to the design of AA-AAS. Among other dictates, the *Standards* requires that procedures for specifying and generating test content be described, that the relationship of the items to the dimensions of the domain be stated clearly, and that steps be taken to ensure that test score inferences accurately reflect the intended construct rather than any disabilities.

Nonregulatory guidance explained the December 9, 2003, regulation to ensure that students with the most significant cognitive disabilities were fully included in state accountability systems and that students had access to challenging instruction linked to state content standards. The guidance clarified that states are responsible for designing assessment systems that permit all students in the tested grades to be assessed against grade-level content and achievement standards, ensuring that assessments are based on state content standards. States are expected to field-test assessments by sampling the types of students expected to participate in the final assessment administration, define the assessment's measurement constructs precisely, and develop accessible test forms that allow for a wide range of accommodations in test administration. For AA-AAS in grades 3 through 8 and at the high school level, the assessment materials should show a clear link to the content standards for the grade the student is enrolled in, although the grade-level content may be reduced in complexity or modified to reflect prerequisite skills. The AAD-ELA/R holds this guidance as the target for performance task assessment and design and selected evidence-centered design and alternate assessment development methods.

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. Furthermore, this process is generalizable and can be applied to all content areas and types of evidence. The use of this approach in the AAD-ELA/R project will contribute much-needed information for improving AA-AAS and will further inform efforts to improve assessment practices generally across the ability spectrum and specifically for students with significant cognitive disabilities.

A review of relevant literature on alternate assessments and the results of the peer review process applied by the U.S. Department of Education to alternate assessments indicate that the technical quality of alternate assessments continues to be a significant challenge (GAO, 2009; Quenemoen, 2008; Quenemoen, Lehr, Thurlow, & Massanari, 2010). Meeting the standards adopted by the American Psychological Association and the American Educational Research Association requires access to large item pools, large samples of students to establish item and scale functioning and difficulty characteristics, and the use of standardization at every step in the assessment development process. Alternate assessment systems vary greatly in the design of the system, type of evidence collected, and the standardization that is applied in part because of the nature of the evidence collected in the alternate assessment systems. Portfolios are still the most common, and they frequently lack evidence supporting their reliability and validity (Cameto et al., 2009; Quenemoen, Thompson, & Thurlow, 2003; Thompson et al., 2005). Some states have been moving toward use of performance tasks to assess students with significant cognitive disabilities, which have the advantage of producing scores that can be evaluated through modern item response theory methods and can be administered to groups of students. In addition, formative assessments for this population are beginning to be explored.

Historically, large-scale assessments have not focused on how content, design, or task characteristics influence the ability of students to perform, especially those students in the tails of the achievement distribution. Alternate assessment designers in particular have often lacked systematic design processes that (1) define the focal knowledge, skills, and abilities (KSAs) required to demonstrate proficiency in academic content areas; (2) design assessment tasks with features that are well aligned with the focal KSAs; (3) design assessment tasks that minimize nonfocal KSAs and thereby mitigate construct-irrelevant variance; and (4) take into account the many ways that students perceive test content and express their responses. Those using the rigorous multistep design process that is central to ECD carefully consider how the content, task, and learner characteristics interact in the creation of assessment tasks.

Application of ECD to the Assessment Needs of a Challenging Population

Students with significant cognitive disabilities may come from any of the 13 regulatory categories included in the Individuals with Disabilities Improvement Act. In a survey of special education teachers of students with significant cognitive disabilities in several states, however, Cameto and colleagues (2010) found that, when asked to report on a randomly selected "target" student with significant cognitive disabilities, teachers reported these students were primarily clustered into three disability categories: mental retardation, autism, and multiple disabilities. Although these data represent only three states, they are consistent with findings reported by Kearns (2007). An additional finding was that the majority of these teachers reported that the target students had multiple disabling conditions (Cameto et al., 2010). The teachers surveyed by Cameto and colleagues (2010) also provided information on students' communication level and academic ability. Teachers were provided with descriptions of three communication levels developed by Browder, Flowers, and Wakeman (2008[2]) and asked to indicate which one best reflected the highest level at which their target student currently communicated. A majority of teachers (68%) reported that the target students with significant cognitive disabilities communicated with symbols or words and had basic or emerging functional academic skills. A small percentage (12%) indicated that the target students had no reliable communicative response.

In the past, students with significant cognitive disabilities typically lived in residential institutions where they were provided with little in the way of education (McDonnell, Hardman, & McDonnell, 2003). Since the passage of Public Law 94-142, also known as Part B of the Education of the Handicapped Act (1975), renamed in 1990 the Individuals with Disabilities Education Act (IDEA), these students have become integrated into home and community life,

---

[2] Communication levels were described as follows:
- 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 may not use objects to communicate. May or may 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 of objects 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.

living with their families, attending their neighborhood schools, learning to read, traveling independently in their communities, and engaging in productive employment as adults. Browder and Spooner (2003) reviewed the evolution of special education from the developmental mental-age-based perspective of the 1970s through the functional life-skills view of the 1980s, the social inclusion and self-determination view of the 1990s, and the academic standards-based demands since the turn of the century. Each time expectations have been raised, students have exceeded previous expectations, and now most are members of their communities, have friends, and enjoy social memberships like their nondisabled peers (Wagner, Cadwallader, & Marder, 2003).

Although federal requirements hold students with the most significant cognitive disabilities to high academic expectations, the strongest argument for such high expectations for these students is their own performance over the last three decades (Marion & Pellegrino, 2006). Since the advent of IDEA, expectations for students with disabilities have been raised repeatedly, and students have consistently outperformed what had previously been perceived to be their limits. Initial research indicates that including students with disabilities in large-scale accountability testing results in higher expectations, improved instruction, and improved performance for those students (Cortiella, 2007; Kleinert, Kennedy, & Kearns, 1999; Quenemoen et al., 2001; Towles-Reeves, Garrett, Burdette, & Burge, 2006; Ysseldyke, Dennison, & Nelson, 2003). The collaborating states—Utah, Idaho, and Kansas—and the SRI team recognize that efforts to design alternate assessments must proceed within this context of the possibilities signaled by previous advances in special education.

High expectations—a hallmark of good education—now include academic performance for these students. But how can their academic performance be assessed? For general education students, most if not all statewide assessments have been developed following careful plans and blueprints linking content standards to assessment items with known psychometric properties, and processes and links have been well documented. Thus, their validity is well understood. For students with many types of disabilities, such assessments may be accommodated or modified. For students with significant cognitive disabilities, accommodations or modifications to the general education assessments are not sufficient. Although assessments for these students must by law be linked to general education content standards, they may use alternate academic achievement standards to measure KSAs.

Many students with significant cognitive disabilities also have coexisting physical or sensory disabilities that can interfere with their assessment performance. In recent years, augmentative and alternative communication devices and assistive technologies have reshaped the way such students are taught and learn, raising even further our expectations about what they may achieve. For students eligible to take AA-AAS, at least three important factors must be attended to: accommodations and technology, including universal design for learning and assessment; alternative and augmentative communication systems; and systematic prompting with feedback that has been used extensively in research with students with severe disabilities (Browder & Cooper-Duffy, 2003). Each of these considerations will be integrated into the design of assessment tasks based on the ECD process in this project.

The AAD-ELA/R project emulates and extends the ECD approach to the design of alternate assessment tasks in English language arts/reading. ECD is a practical theory-based approach to developing quality assessments that combines developments in cognitive psychology and advances in measurement theory and technology. ECD is a well-understood process that can be used in all stages of assessment design and development, from domain analysis to the

specification of student, evidence, and task models to the creation of items and tasks and finally to the design of an assessment delivery system. Although each of the collaborating states has unique needs, ECD provides a robust and suitable approach that can be customized to meet these needs. The tasks designed in the AAD-ELA/R project can be implemented in portfolio or performance task assessment systems or in formative benchmark applications. The implementation of tasks can be guided according to the assessment specifications of each participating state—portfolio systems in Idaho and Kansas and on-demand summative assessments in Utah.

A synergistic application of ECD and UDL facilitates the development of assessment tasks aligned with academic content standards, increases the accessibility of these tasks, and raises expectations for the performance of students with significant cognitive disabilities.

Project Description and Goals

The AAD–English Language Arts/Reading project combines current knowledge from multiple disciplines to advance the design of alternate assessment performance tasks for students with significant cognitive disabilities. The approach integrates recent work in (1) the pedagogy of special education for students with significant cognitive disabilities (Browder & Spooner, 2003), (2) alternate assessment design (Bechard, 2005), and (3) universal design for learning (CAST, 2008) with (4) evidence-centered assessment design (Mislevy & Haertel, 2006). This work is guided by federal guidelines for alternate assessment design that specify that "all students, including students with disabilities, be held to grade-level achievement standards when taking assessments" (U.S. Department of Education, 2005).

The goals of the project are to

1. Extend the conceptual framework of evidence-centered design to alternate assessment based on alternate achievement standards using the Principled Assessment Design for Inquiry (PADI) assessment design system

2. Integrate the principles of universal design for learning with ECD to guide the development of tasks that are accessible to all learners

3. Select a set of Common Core State Standards that represent critical areas of learning common across all three states in Reading Literature, Reading Informational Text, Foundational Reading, Writing, Language, and Speaking and Listening on which to base the design and development of assessment tasks using ECD

4. Develop AA–AAS assessment Design Patterns, task templates, assessment task specifications, and exemplar tasks that address Common Core State Standards in English language arts for students with significant cognitive disabilities

5. Evaluate the exemplar assessment tasks produced using ECD through task tryouts in all three states

6. Enhance the human capital of state departments of education staff in learning how to use ECD to design and develop assessment tasks for use with students with significant cognitive disabilities, including the provision of procedural guidelines

7. Support state departments of education staff and teachers in the development of additional performance tasks in English language arts to expand the task bank for each state

8. Provide participating states with a library of Design Patterns, task templates, and task specifications that are reusable and extensible to the authoring of additional performance tasks in English language arts.

Evidence-Centered Design

Evidence-centered design is a recommended approach for the development of educational assessments and can be applied to a range of content standards and assessment types. The rigorous multilayer design process central to ECD enables designers to consider systematically the content, task, and learner characteristics that influence student performance. ECD provides a foundation for assessments that states can use to address the validity of their assessment systems.

A strength of ECD 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. For example, in a mathematics examination, math content would be targeted and the need for non-construct-relevant skills such as reading would be minimized; designers would consider supports such as use of a large font or alternate response options during item design instead of modifying items and tasks after they have been written.

The ECD process involves five layers of activity:

1. Domain analysis—determining the specific content to be included in the assessment. Use of the common core standards and existing state standards exemplify starting points for domain analysis.

2. Domain modeling—creating and documenting a high-level description of the proposed assessment. Design Patterns are one example of this type of activity.

3. Conceptual assessment framework design—specifying in detail the knowledge, skills, and abilities to be assessed, the evidence that needs to be collected, and the features of the tasks that will elicit the evidence. Also identified are nontargeted KSAs, which, although required for successful performance on an item, are not the intended target of the assessment. By identifying nontargeted KSAs, designers can minimize construct-irrelevant variance and maximize accessibility. Finally, the psychometric model and evaluative decision rules for task scoring are considered and assessment task features are detailed and carefully aligned with the targeted and nontargeted KSAs.

4. Implementation—creating the assessment items or tasks, along with appropriate accessible alternate representations of item or task content.

5. Delivery—specifying the processes for the assessment administration, scoring, and reporting, including accessibility features that are allowed without violating the targeted KSAs.

Universal Design for Learning

Universal design emphasizes the importance of addressing accessibility for the broadest range of potential users during the initial stages of designing a product and throughout the

development and implementation of the product. The use of universal design principles creates flexible solutions because designers consider from the start the diverse ways individuals will interact with a product and the environment.

The tenets of universal design have been extended to the education arena through universal design for learning (UDL). When sources of construct-irrelevant variance in an assessment are identified by ECD, UDL principles can be used to minimize the variance by incorporating appropriate options for how students interact within the assessment environment. In this way, ECD works synergistically with UDL. By considering multiple means of perception, expression, cognition, language and symbol use, executive functioning, and engagement, the application of UDL in the ECD process accounts for individual differences in how students recognize, strategize, and engage in learning and testing situations. This can include consideration of augmentative and alternative communication systems. This synergistic process minimizes the unintended negative influence that access needs may have on student performance and maximizes the opportunities for students to show what they know and can do.

Selection of Common Core State Standards in English Language Arts/Reading for the Development of Design Patterns and Tasks

In an analysis of test design and development methods, Bechard (2005) reported that the best approach to designing alternate assessments aligns the assessment content with a state's academic content standards, thus both promoting access to the general curriculum and increasing instructional opportunities for students with significant cognitive disabilities. States have developed alternate assessment items, tasks, or types of evidence of student performance in two ways. Most states based the design of their items/tasks for their AA-AAS on extensions of the grade-level content standards referred to as extended standards and adopted by the state boards of education (Cameto et al., 2009). Other states based the design of the items/tasks for their AA-AAS on the grade-level content standards adopted by their boards of education for all students, often referred to as general education grade-level standards. The three AAD-ELA/R collaborating states had adopted extended content standards in English language arts and reading for their existing AA-AAS. With the adoption of the Common Core State Standards (National Governors Association Center for Best Practices and The Council of Chief State School Officers, 2010) by the three states, they elected to develop assessment tasks based on the CCSS in English language arts/reading rather than their extended state standards.

The project developed criteria to guide the selection of common core standards for which tasks would be developed. The criteria allowed for key ideas in each of the CCSS strands to be included and for selection of standards across grade levels 3–8 and high school. This process, domain analysis, resulted in the identification of 30 common core standards for which tasks could be designed. The North Carolina Department of Public Instruction's Extended Common Core State Standards were consulted to guide and delineate the scope of the ELA/R content to be assessed. From this process, a final set of 21 Design Patterns and associated families of tasks were developed.

The ECD Co-Design Process

Co-design is a process of bringing together the expertise of assessment specialists, special educators, and content area specialists to create Design Patterns, task templates, task

specifications, and exemplar tasks. In this project, the co-design team members were specialists in large-scale and formative assessment, special educators of students with significant cognitive disabilities, and English language arts/reading educators with experience in instruction and assessment. This constellation of expertise was contributed by members of each state department of education, by SRI International, and by nationally recognized experts in special education, English language arts/reading, and assessment.

To facilitate the design process, project staff used products that are associated with the PADI online assessment design system. This technology systematically supports the design of evidence-based assessment items and tasks using Design Patterns and task templates

Step 1: Co-design team develops Design Patterns. Design Patterns are guiding structures that are part of the domain modeling layer of ECD. Design Patterns comprise attributes that are necessary for constructing an evidentiary-based assessment. These attributes are based on the work of Messick (1994) and Mislevy and his colleagues (Mislevy, Hamel, et al., 2003; Mislevy, Steinberg, et al., 2003). Each Design Pattern articulates an assessment argument by identifying the focal KSAs that are to be measured, the kinds of observations that can provide evidence of this knowledge or skill, and the features of task situations that allow the students to provide this evidence. Also specified in the Design Patterns are any nonfocal KSAs that may be required for students to respond correctly to the assessment tasks but are not the target of the assessment task (for example, reading comprehension and decoding skills needed to respond to a mathematics word problem). Design Patterns also capture the ways assessment tasks can be varied to increase or decrease demands for knowledge and specify the work products and rubrics that the assessment designer may want to use. In the AAD-ELA/R project, 21 Design Patterns in English language arts/reading were created and apply to the associated English language arts/reading common core standards.

Step 2. Co-design team develops summary task templates. A summary task template was completed for each state. This template provides an overview of the assessment system the state uses including an overview of its student model, which consists of the constructs to be assessed (for example, overall English language arts/reading proficiency; subdomain proficiency as appropriate, such as reading literature and informational text, reading foundations, speaking and listening, language, and writing.); scoring and evaluation rubrics; measurement models; and descriptions of the kinds of stimulus materials and presentation used in each state's tasks and items.

Step 3. Co-design team develops task specifications and authors tasks. Task specifications provide guidelines for the design of individual assessment tasks. Designers specify the particular stimuli and response options that will be presented to students. For example, in a task specification, the designer indicates that four examples of hyberbole (rather than three or five) will be presented to students in a passage. Students will be expected to restate the meaning of the hyperbole in their own words. Completing the task specification, designers also specify how students' responses will be scored, give administration guidelines, and identify the variable features that can be used to increase or decrease the difficulty of the tasks. Prompts, graphics, multimedia presentations, and supporting materials are described in detail for each task.

Twenty-two task specifications were developed, and each is linked to one of the 21 Design Patterns; four items are associated with each task specification. For each task specification, the first item is designed to be the most cognitively complex and to assess one of the following

depth-of-knowledge (DOK) levels: Application, Comprehension, or Performance (Flowers et al., 2007). The second item is designed to be less complex and targets a lower DOK level (either Performance or Recall). The third item is even less complex and targets the Recall DOK level. If students are unable to respond to the third item at the Recall level, they are asked to respond to a item at the Attention DOK level (the fourth item). The first and second items are designed to align with a single focal KSA, which was selected to be the target of the assessment task at the beginning of the co-design process. The third item is designed to align with an additional KSA, which is also selected at the beginning of the co-design process. The additional KSA could be described as a foundational skill in that it is typically a prerequisite for successful performance on the first and second items.

Step 4. State departments of education staff and teachers create additional performance tasks. Using the Design Patterns, summary task templates, task specifications, and the written description of exemplar tasks, state department of education staff members and selected teachers are provided with the materials needed to replicate the process of generating performance tasks for AA-AAS in English language arts/reading with guidance from the SRI team.

Design for Pilot Task Tryouts

The collaborating states will pilot-test the newly developed assessment tasks with teachers administering them to students eligible to take state AA-AAS. Grant funds will be used to reimburse teachers for their time in administering the pilot tests. The pilot-testing is summarized below. Details on the pilot-testing, such as sampling criteria and size, timing and scheduling, recruitment, administration, and data collection activities, will be described in Technical Report 7. To pilot the 88 newly designed items, each state will gather information by administering tasks using common instructions. The focus of the pilot will be to collect information about task variability and the appropriateness of the tasks to measure a range of student performance levels.

Task viability. Teachers will administer pilot task tryouts to students to judge the viability of the tasks and items. Can the four items associated with a Design Pattern be administered as designed? Are the item instructions and materials clear to the teacher? Are they clear to the student? Data will be collected through a teacher questionnaire and observations of item administration. Data will inform improvement of the items.

Appropriateness of items to measure a range of student performance levels. Items will be administered to students with significant cognitive disabilities whose teachers characterize them as demonstrating low, medium, or high symbolic functioning (Browder et al., 2008). Which students successfully perform the first item (the most complex item) and at what level of symbolic functioning? Which students successfully perform the second, third, and fourth items (decreasingly complex items), and at what level of symbolic functioning do students perform correctly? Data will inform modification of items so all or most students can gain access to at least one item associated with each Design Pattern.

Each of the 22 tasks will be administered to nine students across the three communication levels (i.e., three students at each level). These tasks will be piloted with a minimum of 198 students in each of the three states ($N = 594$). (See the appendix for a list of the common core standards linked to the Design Patterns.)

Dissemination Plan

Holding high expectations for the academic achievement of students with significant cognitive disabilities is now a widely accepted practice, in part because it is included in legislation: IDEA in 1997 and 2004, ESEA in 2001. Evidence-centered design holds promise to further this practice by ensuring that challenging academic content will be the focus of AA-AAS test design. By applying ECD principles to the development of alternate assessments for students with significant cognitive disabilities, the AAD-ELA/R project is breaking new ground. It is essential that the outcomes and lessons learned from this research be shared with other states so that the quality of assessment for this population can be improved nationally. To achieve this goal, several dissemination activities are planned, as described below.

Website. SRI has developed a project website (http://alternateassessmentdesign.sri.com) to post study reports and selected products that will be available to the public. The site includes links to relevant websites, such as those for the CCSSO (Council of Chief State School Officers), NAAC (National Alternate Assessment Center), NCEO (National Center on Education Outcomes), and the NCSA (National Conference on Student Assessment), and has links to the websites for each of the collaborating states and related contact information. The URL for the site will be distributed to interested parties at national meetings and presentations and via a listserv provided by SRI.

Webinar. SRI will host a free 2-hour web-based interactive seminar presentation to disseminate project procedures and outcomes and to answer questions. Information about and invitations to attend the webinar will be on the website and distributed via e-mail using the listserv described above.

Technical Report Series including Procedural Guidelines. As each report in the Technical Report series is completed, it will be posted on the project website. Technical Report 9 is a user-friendly document that describes the process and procedures used in the AAD-ELA/R project so that states outside the consortium can learn about the ECD principles used. The planned technical reports are as follows:

1. *Project Overview: Applying Evidence-Centered Design to Alternate Assessments in English Language Arts/Reading for Students with Significant Cognitive Disabilities.*

2. *Current State of English Language Arts/Reading Assessment in Alternate Assessment.* A description of (1) the state of the art in alternate assessment in English language arts/reading and (2) the current state of practice in alternate assessment design in English language arts/reading.

3. *Domain Analysis—Selection of Common Core State Standards in English Language Arts/Reading for the Development of Design Patterns and Tasks.* The results of an analysis of the CCSS in English language arts/reading to identify a set of standards for which evidence-centered Design Patterns, task specifications, and assessment tasks were completed.

4. *Design Patterns—Developing Design Patterns for Students with Significant Cognitive Disabilities in English Language Arts*. A description of the theoretical foundations of evidence-centered design that underlie design patterns, the processes used to create design patterns, and an example design pattern and associated English language arts task.

5. *Synergistic Use of Evidence-Centered Design and Universal Design for Learning for Improved Assessment Design*. A brief description of the integration of the ECD and UDL approaches used to develop the assessment design tools in this project.

6. *Design and Development of Assessment Tasks.* A library of English language arts/reading items and tasks available from this project.

7. *Pilot Task Tryouts Design*. The design of the tryouts, including a description of the sample of students to be tested, the logistics of the data collection, and the qualitative and quantitative analyses to be conducted.

8. *Finding, Conclusions, and Recommendations for AAD-ELA/R Tasks*. The qualitative and quantitative findings of the assessment task tryouts, conclusions about the design and development process, and recommendations for further research and development.

9. *Procedural Guidelines for Design Patterns and Assessment Tasks*. The steps to follow in applying the co-design process to the creation of design patterns using ECD and UDL.

References

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

Bechard, S. (2005). *Developing alternate assessments using expanded benchmarks from a nine-state consensus framework in reading, writing, mathematics, and science.* Paper presented at the 2005 National Council on Measurement in Education (NCME) Annual Conference, Montreal, Canada.

Browder, D. M., & Cooper-Duffy, K. (2003). Evidence-based practices for students with severe disabilities and the requirement for accountability in "No Child Left Behind." *The Journal of Special Education, 37*(3), 157–163.

Browder, D. M., & Spooner, F. (2003). Understanding the purpose and process of alternate assessment. In D. Ryndak & S. Alper (Eds), *Curriculum and instruction for students with significant disabilities in inclusive settings* (2nd ed., pp. 51-72). Boston, MA: Allyn & Bacon.

Browder, D. M., Flowers, C., & Wakeman, S. Y. (2008). Facilitating participation in assessments and the general curriculum: Level of symbolic communication classification for students with significant cognitive disabilities. *Assessment in Education: Principles, Policy, and Practice*, *15*(2), 137–151.

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

Cameto, R., Bergland, F., Knokey, A.-M., Nagle, K., Sanford, C., Kalb, S. C., Blackorby, J., Sinclair, B., Riley, D., & Ortega, M. (2010, April). *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.

CAST (2008). *Universal design for learning guidelines version 1.0.* Wakefield, MA: Author.

Cortiella, C. (2007). *Rewards and roadblocks: How special education students are faring under No Child Left Behind*. Retrieved from http://www.ncld.org/images/ stories/oncapitolhill/ policyrelatedpublications/rewardsandroadblocks/rewardsandroadblocks.pdf

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

Gong, B., & Marion, S. (2006). *Dealing with flexibility in assessments for students with significant cognitive disabilities.* Minneapolis: University of Minnesota, National Center On Educational Outcomes. Available at http://cehd.umn.edu/NCEO/onlinepubs/ synthesis60. html

Government Accountability Office. (2009). *No Child Left Behind: Enhancements in the Department Of Education's review process could improve state academic assessments* (GAO 09-911). Retrieved from www. gao. gov/new. items/D09911. pdf

Kearns, J. (2007). *Defining the population to be served.* Paper presented at the 8th annual Maryland Conference on Alternate Assessments, University of Maryland, College Park.

Kleinert, H. L., Kennedy, S., & Kearns, J. F. (1999). The impact of alternate assessments: A statewide teacher survey. *The Journal of Special Education, 33*(2), 93–102.

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.

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. (2003). A*n introduction to persons with moderate and severe disabilities: Educational and social issues* (2nd Ed.). Boston, MA: Allyn and Bacon.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13–23.

Mislevy, R., & Haertel, G. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 24*(4), 6–20.

Mislevy, R., Hamel, L., Fried, R., Gaffney, T., Haertel, G., Hafter, A., Wenk, A. (2003). *Design patterns for assessing science inquiry* (PADI Technical Report 1). Menlo Park, CA: SRI International.

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

National Council of Teachers of Mathematics. (2005). *Principles and standards for school mathematics.* Reston, VA: Author.

National Governors Association Center for Best Practices and The Council of Chief State School Officers. (2010). *Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects*. Common Core State Standards Initiative: Preparing America's Students for College and Career. Washington, DC: author. Retrieved from http://www.corestandards.org/the-standards/english-language-arts-standards

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

Quenemoen, R., Kearns, J., Quenemoen, M., Flowers, C., & Kleinert, H. (2010). *Common misperceptions and research-based recommendations for alternate assessment based on alternate achievement standards* (Synthesis Report 73). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved from http:// www. cehd. umn. edu/NCEO/onlinepubs/synthesis73/synthesis73. pdf

Quenemoen, R. F., Lehr, C. A., Thurlow, M. L., & Massanari, C. B. (2001). *Students with disabilities in standards-based assessment and accountability systems: Emerging issues, strategies, and recommendations* (Synthesis Report 37). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Quenemoen, R., Thompson, S., & Thurlow, M. (2003). *Measuring academic achievement of students with significant cognitive disabilities: Building understanding of alternate assessment scoring criteria* (Synthesis Report 50). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Ryan, J. M., Quenemoen, R. F., & Thurlow, M. L. (2004). *I say potato, you say potato: The assessment-speak gap between general and alternate assessment experts.* Minneapolis: University of Minnesota, National Center on Educational Outcomes.

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

Thompson, S. J., Johnstone, C. J., Thurlow, M. L., & Altman, J. R. (2005). *2005 State special education outcomes: Steps forward in a decade of change*. Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Thompson, S., & Thurlow, M. (2003). *2003 State special education outcomes: Marching on*. Minneapolis: University of Minnesota, National Center on Educational Outcomes. Available at http://www.education.umn.edu/NCEO/onlinepubs/2003statereport.html

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., Garrett, B., Burdette, P., & Burdge, M. (2006). Validation of large-scale alternate assessment systems and their influence on instruction—What are the consequences? *Assessment for Effective Intervention*, *31*(3), 45–57.

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

Wagner, M., Cadwallader, T. M., & Marder, C. (with Cameto, R., Cardoso, D., Garza, N., Levine, P., & Newman, L.). (2003). *Life outside the classroom for youth with disabilities. A report from the National Longitudinal Transition Study-2 (NLTS2).* Menlo Park, CA: SRI International. Available at www.nlts2.org/reports/2003_04-2/nlts2_report_2003_04-2_complete.pdf

Ysseldyke, J., Dennison, A., & Nelson, R. (2003). *Large-scale assessment and accountability systems: Positive consequences for students with disabilities* (Synthesis Report 51). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

**Appendix: Common Core State Standards Selected for the
Development of Tasks for Pilot-Testing**

Reading Informational Text (RI)

1. Grade 4 (RI.4.7): Interpret information presented visually, orally, or quantitatively (e.g., in charts, graphs, diagrams, timelines, animations, and interactive elements) and explain how the information contributes to understanding of the text in which it appears.
2. Grade 7 (RI.7.5): Analyze the structure an author uses to organize a text, including how the major sections contribute to the whole and to the development of the ideas.
3. Grade 8 (RI.8.7): Evaluate the advantages and disadvantages of using different mediums (e.g., print or digital text, video, multimedia) to present a particular topic or idea.
4. Grade 11/12 (RI.11-12.6): Determine an author's point of view or purpose in a text in which the rhetoric is particularly effective, analyzing how style and content contribute to the power, persuasiveness, or beauty of the text.

Reading Literary Text (RL)

5. Grade 3 (RL.3.1): Ask and answer questions to demonstrate understanding of a text, explicitly using the text as the basis for the answers.
6. Grade 3 (RL.3.3): Describe characters in a story (e.g., their traits, motivations, or feelings) and explain how they contribute to the sequence of events.
7. Grade 7 (RL.7.3): Analyze how particular elements of a story or drama interact (e.g., how setting shapes the characters or plot).
8. Grade 9/10 (RL.9-10.3): Analyze how complex characters (e.g., those with multiple or conflicting motivations) develop over the course of a text, interact with other characters, and advance the plot or develop the theme.

Reading Foundations (RF)

9. Grade 5 (RF.5.4): Read with sufficient accuracy and fluency to support comprehension.
   a. Read on-level text with purpose and understanding.
   b. Read on-level prose and poetry orally with accuracy, appropriate rate, and expression on successive readings.
   c. Use context to confirm or self-correct word recognition and understanding, rereading as necessary.

Language (L)

10. Grade 4 (L.4.2): Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing.

      a. Use correct capitalization.

      b. Use commas and quotation marks to mark direct speech and quotations from a text.

      c. Use a comma before a coordinating conjunction in a compound sentence.

      d. Spell grade-appropriate words correctly, consulting references as needed.

11. Grade 5 (L.5.5): Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.

      a. Interpret figurative language, including similes and metaphors, in context.

      b. Recognize and explain the meaning of common idioms, adages, and proverbs.

      c. Use the relationship between particular words (e.g., synonyms, antonyms, homographs) to better understand each of the words.

12. Grade 8 (L.8.2): Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing.

      a. Use punctuation (comma, ellipsis [...], dash) to indicate a pause or break.

      b. Use an ellipsis to indicate an omission.

      c. Spell correctly.

13. Grade 8 (L.8.5): Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.

      a. Interpret figures of speech (e.g. verbal irony, puns) in context.

      b. Use the relationship between particular words to better understand each of the words.

      c. Distinguish among the connotations (associations) of words with similar denotations (definitions) (e.g., bullheaded, willful, firm, persistent, resolute).

14. Grade 11/12 (L.11-12.5): Demonstrate understanding of figurative language, word relationships, and nuances in word meanings.

      a. Interpret figures of speech (e.g., hyperbole, paradox) in context and analyze their role in the text.

      b. Analyze nuances in the meaning of words with similar denotations.

Speaking and Listening (SL)

15. Grade 3 (SL.3.2): Determine the main ideas and supporting details of a text read aloud or information presented in diverse media and formats, including visually, quantitatively, and orally.

16. Grade 7 (SL.7.2): Analyze the main ideas and supporting details presented in diverse media and formats (e.g., visually, quantitatively, orally) and explain how the ideas clarify a topic, text, or issue under study.

17. Grade 9/10 (SL.9-10.5): Make strategic use of digital media (e.g., textual, graphical, audio, visual, and interactive elements) in presentations to enhance understanding of findings, reasoning, and evidence and to add interest.

Writing (W)

18. Grade 3 (W.3.8): Recall information from experiences or gather information from print and digital sources; take brief notes on sources and sort evidence into provided categories.
19. Grade 5 (W.5.1): Write opinion pieces on topics or texts, supporting a point of view with reasons and information.
    a. Introduce a topic or text clearly, state an opinion, and create an organizational structure in which ideas are logically grouped to support the writer's purpose.
    b. Provide logically ordered reasons that are supported by facts and details.
    c. Link opinion and reasons using words, phrases, and clauses (e.g., consequently, specifically).
    d. Provide a concluding statement or section related to the opinion presented.
20. Grade 5 (W.5.2): Write informative/explanatory texts to examine a topic and convey ideas and information clearly.
    a. Introduce a topic clearly, provide a general observation and focus, and group related information logically; include formatting (e.g., headings), illustrations, and multimedia when useful to aiding comprehension.
    b. Develop the topic with facts, definitions, concrete details, quotations, or other information and examples related to the topic.
    c. Link ideas within and across categories of information using words, phrases, and clauses (e.g., in contrast, especially).
    d. Use precise language and domain-specific vocabulary to inform about or explain the topic.
    e. Provide a concluding statement or section related to the information or explanation presented.
21. Grade 11/12 (W.11-12.2): Write informative/explanatory texts to examine and convey complex ideas, concepts, and information clearly and accurately through the effective selection, organization, and analysis of content.
    a. Introduce a topic; organize complex ideas, concepts, and information so that each new element builds on that which precedes it to create a unified whole; include formatting (e.g., headings), graphics (e.g., figures, tables), and multimedia when useful to aiding comprehension.
    b. Develop the topic thoroughly by selecting the most significant and relevant facts, extended definitions, concrete details, quotations, or other information and examples appropriate to the audience's knowledge of the topic.
    c. Use appropriate and varied transitions and syntax to link the major sections of the text, create cohesion, and clarify the relationships among complex ideas and concepts.
    d. Use precise language, domain-specific vocabulary, and techniques such as metaphor, simile, and analogy to manage the complexity of the topic.
    e. Establish and maintain a formal style and objective tone while attending to the norms

and conventions of the discipline in which they are writing.
f. Provide a concluding statement or section that follows from and supports the information or explanation presented (e.g., articulating implications or the significance of the topic).