SRI International

Alternate Assessment Design–Mathematics

Technical Report 8:

An Expert Panel Review of Alternate Assessment Design–Mathematics' Application of Evidence-Centered Design and Universal Design for Learning to the Development of Alternate Assessment Tasks

March 3, 2011

SRI International
Renée Cameto
Kathryn Morrison
Geneva Haertel
Angela DeBarger

SRI Project No.: P18879

This report was supported in part by the Elementary and Secondary Education Act, Enhanced Assessment Grant (ESEA grant No. # S368A08008) through a partnership with the Utah (Lead), Idaho, and Florida State Departments of Education and SRI International as the research entity. Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the view of ESEA or the U. S. Department of Education.

333 Ravenswood Avenue • Menlo Park, California 94025-3493 • 650.859.2000 • www.sri.com

Contents

Tables

An Expert Panel Review of Alternate Assessment Design–Mathematics' Application of Evidence-Centered Design and Universal Design for Learning to the Development of Alternate Assessment Tasks

Students with significant cognitive disabilities challenge conventions with respect to the teaching, learning, and assessing of academic content. Assessment has been instrumental in changing the learning expectations of these students which 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; Ryan, Quenemoen, & Thurlow, 2004; U.S. Government Accountability Office, 2009). The methods used to date in designing alternate assessments and selecting their content are varied but typically do not match the technical rigor used for designing general education assessments (Bechard, 2005). The Alternate Assessment Design–Mathematics (AAD-M) project is 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-M project is innovative in two aspects: It is applying ECD for the first time 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 achievement standards (AA-AAS). This work extends current knowledge in the field and provides a prototype for future alternate assessment development.

Utah, Idaho, and Florida have formed a consortium with SRI International to improve their AA-AAS using ECD to design and develop assessment tasks that are linked to state extended content standards in mathematics. In this report, we describe

- Project goals and activities
- The development of assessment tasks for accountability purposes for students with significant cognitive disabilities
- ECD and UDL frameworks and describe how they are applied through a co-design process
- The findings from an expert review of a sample of the project products

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 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 is under way 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 results of such tests (Quenemoen, 2008; Quenemoen, Kearns, Quenemoen, Flowers, & Kleinert, 2010; U.S. Government Accountability Office, 2009). 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., 2009a). 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., 2009a). Design and implementation of alternate assessments are in considerable flux (for example, existing assessments are likely to be revised to align with the 2010 Common Core State Standards Initiative[1]).

The *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on 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* require that procedures for specifying and generating test content be described, that the relation of

[1] The Common Core State Standards Initiative (CCSSI) 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, 2 territories, and the District of Columbia developed a draft common core of state standards in English/language arts and mathematics for grades K–12. When the draft is formalized, the participating states will adopt the Common Core Standards. The CCSSI plans to develop a common core of standards in science.

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-M 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-M 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 the technical quality of alternate assessments continues to be a significant challenge (Quenemoen, 2008; Quenemoen et al., 2010; U.S. Government Accountability Office, 2009). 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., 2009b; 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" SWSCD, 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 these teachers reported that the target SWSCD 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 SWSCD 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 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, 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

[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 reads 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.

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, Lehr, Thurlow, & Massanari, 2001; Towles-Reeves, Garrett, Burdette, & Burdge, 2006; Ysseldyke, Dennison, & Nelson, 2003). The collaborating states—Utah, Idaho, and Florida—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-M project emulates and extends the ECD approach to the design of alternate assessment tasks in mathematics. 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 each state's needs. The tasks designed in the AAD-M 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, on-demand summative assessments in Utah, and diagnostic formative assessments in Florida.

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-M 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 that are addressed in this paper 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. Use the National Council of Teachers of Mathematics (NCTM) Principles and Standards for School Mathematics to identify common expectations that represent critical areas of learning common across all three states in number and operations, algebra, geometry, measurement, and data analysis and probability
4. Develop AA-AAS assessment Design Patterns, Summary Task Templates, and Development Specifications and Exemplar Task Templates that address priority state academic standards in mathematics for students with significant cognitive disabilities
5. Evaluate the exemplar assessment tasks produced using ECD through pilot-testing in all three states and through review by nationally recognized experts in mathematics, UDL, and the education and assessment of students with significant cognitive disabilities

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. The layers focus in turn on the identification of the content to be assessed; the creation of a model of the assessment; the design of assessment elements such as potential observations, work products, rubrics, and psychometric models; the creation of these elements including the assessment tasks; and the design of the assessment delivery, scoring, and reporting. Each layer is described below.

1. Domain analysis involves 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 entails creation and documentation of a high-level description of the assessment. Design patterns are one example of this type of activity.
3. Conceptual assessment framework specifies 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 is the creation of the assessment items or tasks, along with appropriate accessible alternate representations of item or task content.
5. Delivery involves specification of 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 in which individuals will interact with a product and the environment.

The tenets of universal design have been extended to the education arena; this extension is referred to as UDL. When sources of construct-irrelevant variance in an assessment are identified by ECD, the application of UDL principles can be used to minimize construct-irrelevant 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 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.

UDL principles are incorporated into this ECD process during assessment design and item authoring by considering multiple means of perception, expression, cognition, language and symbol use, executive functioning, and engagement. This can include consideration of augmentative and alternative communication systems.

Domain Analysis: Crosswalk of States' Extended Standards in Mathematics and the National Council of Teachers of Mathematics (NCTM) Expectations

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., 2009b). 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-M collaborating states had adopted extended content standards in mathematics for their AA-AAS. The state extended standards were the foundation for the domain analysis reported, the first layer of the ECD process.

This study team analyzed the three states' content standard extensions and aligned them with the mathematics expectations put forth in the *Principles and Standards for School Mathematics* (National Council of Teachers of Mathematics, NCTM, 2000. The NCTM expectations are intended to help in focusing curricula within a grade band and in developing curricula progressively across grade bands. They are used to design instructional programs and curricular frameworks for K–12 mathematics. The expectations represent a consensus among educators about the content and processes that are essential to understanding the knowledge, skills, and abilities required for successful performance in mathematics. NCTM identifies expectations in five domains: numbers and operations, algebra, geometry, measurement, and data analysis and probability. To focus the states' efforts on a common set of expectations, the project team conducted a crosswalk between the NCTM expectations and the three states' extended mathematics standards. This crosswalk identified areas of overlap and uniqueness in the NCTM expectations that each state reflected through its extended standards. The crosswalk resulted in identification of a total of 30 NCTM expectations that were common across all three states in grades 3 through 8 and high school. Ten of these expectations were in the domain of numbers and operations and between four to six expectations were in each of the remaining four domains. These expectations were the foundation on which the design patterns and associated 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, Summary Task Templates, and Development Specifications and Exemplar Task Templates and to use the results from expert panel review and pilot testing to refine tasks. In this project, the co-design team members were specialists in large-scale and formative assessment, special educators of SWSCD, and mathematics 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, mathematics, 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. The ECD components are described briefly below; the steps for developing the two major products, Design Patterns and Development Specifications and Exemplar Task Templates, are provided in the appendices.

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. The list of attributes in the Design Pattern and their definitions are presented in Table 1. 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-M project, 30 Design Patterns in mathematics were created and apply to the mathematics content covered in each state.[3] A detailed description of the process used to create Design Patterns is provided in Appendix A.

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 used by the state including an overview of its student model, which consists of the constructs to be assessed (for example, overall mathematics proficiency; subdomain proficiency as appropriate, such as numbers and operation, geometry, etc.); scoring and evaluation rubrics; measurement models; and descriptions of the kinds of stimulus materials and presentation used in each state's tasks and items.

---

[3] For more information about design patterns, see PADI Technical Report 1, *Design Patterns for Assessing Science Inquiry* (Mislevy, Hamel, et al., 2003); PADI Technical Report 5, *The Case for an Integrated Design Framework for Assessing Science Inquiry* (Baxter & Mislevy, 2005); and Technical Report 8, *An Example-Based Exploration of Design Patterns in Measurement* (DeBarger & Riconscente, 2005). Technical reports are available at padi.sri.com.

Table 1. *Design Pattern* Attributes and Definitions

***Design Pattern* Attribute**	**Attribute Definition**
Title	Short name for the Design Pattern (DP)
Summary	Brief description of the family of tasks implied by the DP
Rationale	Nature of the KSAs of interest and why they are important
Focal Knowledge, Skills & Abilities (KSAs)	The primary KSAs targeted by this DP
Additional KSAs	Other KSAs that may be required by tasks from this DP, some of which can be supported by Universal Design for Learning (UDL) and accommodations
Potential Observations	Observed behaviors of students that can provide evidence of Focal KSAs
Potential Work Products	What students say, do, or make that provides evidence about the Focal KSAs
Potential Rubrics	Some evaluation techniques that may apply
Characteristic Task Features	Aspects of assessment situations likely to evoke the desired evidence
Variable Task Features	Aspects of assessment situations that can be varied in order to control difficulty or target emphasis on various KSAs
Educational Standards	State extended standards (if appropriate)

The co-design team develops Development Specifications and Exemplar Task Templates.

Development Specifications and Exemplar Task Templates provide guidelines for the design of individual assessment tasks. Appendix B includes an example *Development Specifications and Exemplar Task Template* for Number and Operations A3 (grades 3–5). Designers specify the particular stimuli and response options that will be presented to students. For example, the designer may indicate that four data points (rather than three or five) will be presented to students who are asked to create a line graph. In completing the Development Specifications and Exemplar Task Templates, 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, diagrams, and supporting materials are described in detail for each task.

A Development Specifications and Exemplar Task Template is linked to each of the 30 Design Patterns; four items are associated with each exemplar task. For each exemplar task, the first task is designed to be the most cognitively complex and to assess one of the following depth-of-knowledge (DOK) levels: Application, Comprehension, or Performance.[4] The second task is designed to be less complex and targets a lower DOK level (either Performance or Recall). The third task is even less complex and targets the Recall DOK level. If students are unable to

[4] 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, NC: National Alternate Assessment Center, University Of North Carolina at Charlotte.

respond to the third task at the Recall level, they are asked to respond to a task at the Attention DOK level (the fourth task). The first and second tasks are designed to align with a single Focal KSA which was selected to be the target of the assessment tasks at the beginning of the co-design process. The third and fourth tasks are 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 tasks. A detailed description of the process used to create Development Specifications and Exemplar Task Templates is provided in Appendix C.

Design for Pilot Task Tryouts

The collaborating states pilot-tested the newly developed assessment tasks with teachers administering them to students eligible to take state AA-AAS. The details of the pilot-testing, such as sampling criteria and size, timing and scheduling, recruitment, administration, and data collection activities, are summarized below and are fully described in a technical report (Technical Report 7: Pilot Task Tryout Design[5]). To pilot the 120 newly designed items (based on 30 design patterns with a suite of 4 items for each design pattern), each state gathered information by administering the items using common instructions and a common data collection system. In addition to item scores, survey data was collected on each item's characteristics as they were administered, whether students had an opportunity to learn the knowledge and skills assessed, and the characteristics of the students and teachers involved in the piloting. The pilot survey analysis focuses on items variability and the appropriateness of the items to measure a range of student performance levels.

Task viability. Teachers who administered the pilot task tryouts to students judged the viability of the each of the 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 inform future improvement of the tasks.

Appropriateness of tasks to measure a range of student performance levels. Tasks were 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 most complex item and at what level of symbolic functioning? Which students successfully perform the decreasingly complex items and at what level of symbolic functioning? Data will inform future modification of tasks so all or most students can gain access to at least one task associated with each design pattern.

(See Appendix D for a list of NCTM expectations linked to the design patterns.)

Design for Expert Panel Review

In the final months of the project, the products produced were submitted to a panel of three experts for their review. This section will describe the panelists, the instructions they received, and the materials distributed to them.

Expert Panelists

Three panelists with expertise in ECD, UDL, mathematics content, and/or special education were recruited to participate in the review process. The three panelists are described below:

5 Technical Report Series available at http://alternateassessmentdesign.sri.com

Panelist 1 is a Distinguished Professor of Special Education at a major U.S. university and has more than two decades of experience conducting research and writing about teaching students with moderate and severe developmental disabilities. The panelist was recognized for these achievements as a Distinguished Researcher by the American Education Research Association Special Education SIG and a First Citizen's Scholar's Award. The panelist's work also focuses on providing resources for practitioners that utilize and extend this research. The panelist has co-authored several curricula and also is co-author of a comprehensive textbook.

Panelist 2 is engaged in research and scholarship focused on supporting all students in mathematics education through application of instructional leadership principles and practices. The panelist has served as Principal Investigator for federally and locally funded research grants emphasizing the development of formative assessment procedures in mathematics and valid decision-making systems for students with diverse needs in the general education curriculum. Specifically, the panelist investigates:

- Formative assessment design frameworks using modern test theory, including Item Response Theory;
- Empirical impact of accommodations and other test changes on the validity of test-score interpretations and uses;
- Implications of using technology to implement universal design of assessment principles to support accessibility; and
- Mathematics teachers' decision-making with a focus on integrating research-based instructional design and delivery principles with teacher content knowledge.

The panelist publishes and presents to audiences focused on research as well as practice in the areas of mathematics education, measurement and assessment, and special education. The panelist's research is informed by experience in K-12 education: the panelist taught high-school science and is trained as a K-12 administrator.

Panelist 3 applies technical skills, mathematics background, special education experience, and clinical specialties to the development of universally-designed software. The panelist is Co-Principal Investigator for a major program funded by the U.S. Department of Education and has also served as Co-Principal Investigator for a project, funded by the U.S. Department of Education's Office of Special Education Programs, to develop and test a prototype of a digital high school biology text containing embedded supports for student learning.

Previously the panelist was an assistant professor in the Department of Occupational Therapy and Assistant Director of Occupational Therapy for major U.S. universities. During the panelist's tenure as an occupational therapist, the panelist presented extensively on the effects of deficits in visual perception on academic learning in children. The panelist's doctoral research analyzed the relationship between visual-spatial abilities and mathematics achievement in boys with and without learning disabilities. The panelist received training in mathematics assessment at the Learning Disorders Unit of Children's Hospital in Boston and conducted evaluations of visual-spatial and mathematics abilities of children with a variety of disabilities for 15 years. The panelist has also worked as an elementary and special education teacher.

Materials and Instructions Sent to Expert Panelists

On January 11, 2001, project materials were sent to the expert panelists. All materials were sent electronically. The materials included:

1. An overview of the AAD-Math project
2. Instructions for Reviewing Design Patterns and Task Templates for AAD-Mathematics EAG
3. 30 Design Patterns
4. 30 “horizontal views” of the Design Patterns which align the critical components of the Focal KSAs, the Potential Observations, and the Potential Work Products
5. 30 “horizontal views” of the Design Patterns which align the critical components of the Additional KSAs and the Variable Features
6. The 30 Development Specifications and Exemplar Task Templates

The Instructions for Reviewing Design Patterns and Task Templates for AAD-Mathematics EAG provided guidelines for reviewing the project documents. The full instructions document is in Appendix E. The instructions included the following guiding statements for reviewing the Design Patterns:

For each Design Pattern:

1. Determine whether you agree with the identification of the Focal KSAs associated with the NCTM Expectations.
2. Judge whether the Potential Observations are appropriate examples of the kinds of evidence that you would expect of students who had mastered the Focal KSAs.
3. Judge whether the Work Products are appropriate for collecting the kinds of evidence identified in the Potential Observation.
4. Determine whether you agree with the Cognitive Background Knowledge Additional KSAs that have been identified as prerequisite or precursor knowledge to the Focal KSAs.
5. Judge the importance of [the remaining 6 categories of Additional KSAs] for the performance of students with significant cognitive disabilities on tasks associated with the NCTM Expectation.
6. Judge whether the Variable Features are adequate supports for the Cognitive Background Knowledge Additional KSAs.

The lists of Additional KSAs and Variable Features are the same in every Design Pattern. Panelists were asked to review these lists just one time using the following guiding statement/question:

7. Judge whether the Variable Features are adequate supports for the Additional KSAs in each of the 6 UDL categories. If the Variable Features were built into the assessment tasks, would they mitigate the barriers to performance created by the Additional KSAs?

The instructions included the following guiding statements and questions for reviewing the tasks:

8. Consider whether [the Item Directives, the Correct Answer, the Description of Stimulus Items, Materials for the Examiner, and Variable Features for Administration to Individual Students] adequately address the Focal KSA (for items 1 and 2).

9. Consider whether [the Item Directives, the Correct Answer, the Description of Stimulus Items, Materials for the Examiner, and Variable Features for Administration to Individual Students] adequately address the Cognitive Background Additional KSA (items 3a and 3b).
10. Are the items specified in the Task Template appropriate for students with significant cognitive disabilities at the grade levels indicated?
11. Are the items well aligned to the specifications in the Design Pattern?

The Design Patterns and task sets developed were divided into two groups: core (9 that all states included in their pilot tests) and additional (21 - 7 piloted in each of the three states). Reviewers were asked to focus initially on the core Design Patterns and tasks and then move to the additional if time permitted. The nine core Design Patterns and tasks reviewed by all three expert panelists include the following expectations/standards[6]:

1. Number and Operations [A3], grades 3-5: Develop understanding of fractions as parts of unit wholes, as parts of a collection, as locations on number lines, and as divisions of whole numbers
2. Number and Operations [A1], grades 6-8: Work flexibly with fractions, decimals, and percents to solve problems
3. Algebra [B2], grades 3-5: Represent the idea of a variable as an unknown quantity using a letter or a symbol
4. Algebra [B1], grades 6-8: Develop an initial conceptual understanding of different uses of variables
5. Algebra [B3], grades 9-12: Use symbolic algebra to represent and explain mathematical relationships
6. Geometry [A1], grades 3-5: Identify, compare and analyze attributes of two- and three-dimensional shapes and develop vocabulary to describe the attributes
7. Geometry [A1], grades 6-8: Precisely describe, classify, and understand relationships among types of two- and three-dimensional objects using their defining properties
8. Geometry [A1], grades 9-12: Analyze properties and determine attributes of two- and three-dimensional objects
9. Measurement [B2], grades 3-5: Select and apply appropriate standard units and tools to measure length, area, volume, weight, time, temperature, and the size of angles.

Results

Each panelist submitted their responses within three weeks of receipt of the materials. Two of the panelists responded to each guiding statement/question for each of the Design Patterns and Tasks. One panelist submitted specific comments about the Design Patterns and tasks but did not respond to each guiding statement/question. The comments of each of the reviewers focused on their areas of expertise. Panelist 1's comments focused on the appropriateness of the components of the Design Patterns and Tasks for students with significant cognitive disabilities. Panelist 2's comments focused primarily on the accuracy of the Design Patterns and Tasks in assessing the

[6] Notations such as A1, B2, etc. following the content subdomain were created by the project to catalogue the expectations and are not part of the NCTM expectations.

mathematical constructs. Panelist 3's comments focused on the appropriateness of the application of UDL principles to various components of the Design Patterns and Tasks.

The individual panelist responses were compiled into tables, one table for each Design Pattern and its associated task. The responses for each of the questions for each Design Pattern and Task pair were reviewed and categorized. Finally, responses were reviewed across all Design Patterns and tasks in order to detect crosscutting themes, The crosscutting themes identified are briefly described in Table 2.

Table 2. Panelist Comments: Crosscutting Themes

Guiding Statements/Questions	Panelist Comments
1. Determine whether you agree with the identification of the Focal KSAs associated with the NCTM Expectations.	One panelist noted that overall the KSAs did a very good job of capturing the key or essential ideas of the content with a strong rationale for each. This panelist commented that the identification of the KSAs would be very useful to the states for planning AA-AAS and for teachers for planning instruction. Another panelist thought that the Focal KSAs were reasonable and represented the range of skills and knowledge students should have to demonstrate proficiency across the expectations. This panelist provided a few specific suggestions to clarify Focal KSAs.
2. Judge whether the Potential Observations are appropriate examples of the kinds of evidence that you would expect of students who had mastered the Focal KSAs.	One panelist indicated that the Potential Observations reflected a range of depth of knowledge. This panelist commented that some of the Potential Observations could benefit from more "hands on" examples that do not complicate assessing mathematical learning with communication ability. Another panelist stated that all the Potential Observations were appropriate examples of evidence that would be expected of students proficient in the construct identified in the Focal KSA. Suggestions for improving several Potential Observations were provided.
3. Judge whether the Work Products are appropriate for collecting the kinds of evidence identified in the Potential Observation.	The two panelists who provided comments on this guiding statement thought that the Potential Work Products were appropriate for the targeted expectation and reflected a range of depth of knowledge.
4. Determine whether you agree with the Cognitive Background Knowledge Additional KSAs that have been identified as prerequisite or precursor knowledge to the Focal KSAs.	The two panelists who provided comments on this guiding statement agreed that in most cases the Cognitive Background Knowledge specified was appropriate for the Focal KSAs and drew on previous grades nicely. One panelist noted, however, that most background knowledge will need to be taught concurrently with the new grade band content to students with significant cognitive disabilities because they will progress across grades based on their age even if they do not master the content. Another panelist suggested that the cognitive background knowledge may too close to the Focal KSA in a few instances and suggested a few additional cognitive background knowledge KSAs to be considered.

Table 2: Panelist Comments: Crosscutting Themes (concluded)

Guiding Statements/Questions	Panelist Comments
5. Judge the importance of the remaining 6 categories of Additional KSAs for the performance of students with significant cognitive disabilities on tasks associated with the NCTM Expectation.	One panelist commented that because it is anticipated that student with students with significant cognitive disabilities will have deficits in precursor cognitive background knowledge, the identification of these KSAs are particularly important to plan for their accommodation. The additional KSAs were found to be useful to help identify areas to consider in planning UDLs. Another panelist also found these KSAs reasonable and important for success on the targeted construct. Another panelist commented that it was difficult to determine which Additional KSAs would be important for a specific NCTM expectation. This panelist felt that all of the perceptual, skill and fluency, language, and affective Additional KSAs were generic enough to be relevant for all tasks but that the other two categories were more dependent on the specific task.
6. Judge whether the Variable Features are adequate supports for the Cognitive Background Knowledge Additional KSAs.	Two panelists noted that the Variable Features for the Cognitive Background Knowledge Additional KSAs were overall very helpful and typically targeted appropriate adaptations and accommodations.
Judge whether the Variable Features are adequate supports for the Additional KSAs in each of the 6 UDL categories. If the Variable Features were built into the assessment tasks, would they mitigate the barriers to performance created by the Additional KSAs?	Two panelists noted that the Variable Features for the UDL categories were overall very helpful and typically targeted appropriate adaptations and accommodations. One panelist cautioned against the use of “providing definitions” as not likely to be helpful to many students with SCD. This panelist also suggested giving a demonstration as a more viable Variable Feature for this population, provided some examples, and commented that a single demonstration would help the student know what the examiner expected and would be unlikely to affect whether the student understood the construct being addressed by the Additional KSA.
8-11. Task Templates- are the items appropriate for students with SCDs? Are they well aligned to the Design Pattern?	One panelist overall thought the tasks were good choices. This panelist wanted to see all three items in a suite of items associated with the same Focal KSA or content and thought that the extensions across DOK could be improved possibly by using objects from everyday activities. Another panelist thought that the tasks appeared appropriate for the target population and appropriately measured the Focal and Additional KSAs. Another panelist suggested that while there were several potential variable UDL features listed, the process would benefit from greater documentation of how they were incorporated into the actual tasks. This panelist also suggested that the variable features for administration to individual students include options for individualizing the materials used.

Summary

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 has been 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. The ECD process provides an approach to large-scale assessment that focuses on how content, design, and task characteristics can influence the ability of students with significant cognitive disabilities to perform. 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. The ECD process carefully considers how the content, task, and learner characteristics interact in the creation of assessment tasks.

Many students with significant cognitive disabilities have coexisting physical or sensory disabilities that can interfere with their assessment performance. Augmentative and alternative communication devices and assistive technologies have reshaped the way such students are taught and learn. Three important factors must be attended to particularly for students eligible to take AA-AAS: 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 was integrated into the design of assessment tasks based on the ECD process in this project. The tasks designed in the AAD-M project can be implemented in portfolio or performance task assessment systems or in formative benchmark applications.

The synergistic application of ECD and UDL resulted in the development of assessment tasks aligned with academic content standards, increased the accessibility of these tasks, and raised expectations for the performance of students with significant cognitive disabilities in the states participating in the project.

This paper described a number of the goals for the project, and focused on The goals of the project that are addressed in this paper are to the development of AA-AAS assessment Design Patterns and Development Specifications for Task that address priority state academic standards in mathematics for students with significant cognitive disabilities. As part of the validation process typically undertaken in assessment task development, task refinement can benefit from and evaluation of the tasks through pilot-testing and through review by recognized experts in mathematics, UDL, and the education and assessment of students with significant cognitive disabilities. The review of a core set of tasks was conducted near the completion of the project to identify areas when specific adjustments could improve the task and to comments of the adequacy of the approach. General cross-cutting themes expressed by the expert panelists were presented.

The work of the AAD-Math project was summarized by one panelist, "it is clear that considerable work has been invested in creating assessment tasks and that these tasks will accurately measure the knowledge, skills, and abilities of students with significant cognitive disabilities." Another panelist noted that the AAD-Math project provided an excellent set of KSAs for the NCTM standards. “The KSAs captured the “biggest” ideas within the content with appropriate rationale. In applying them for students with SCD, consideration has been given to providing appropriate accommodations and supports so that the tasks can be accessible. The specific tasks selected link to the content and are overall teachable for students with SCD given intensive, systematic instruction. Some more development of the tasks by DOK is recommended.” The recommendations of the reviewers for improving specific task components were considered and modifications made as deemed appropriate.

References

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

Baxter, G., & Mislevy, R. (2005). *The case for an integrated design framework for assessing science inquiry* (PADI Technical Report 5). Menlo Park, CA: SRI International. Retrieved from http://padi.sri.com/downloads/TR5_IDFramework.pdf

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, 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., Flowers, C., & Wakeman, S. Y. (2008). Facillitating 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.

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: Allyn & Bacon.

Cameto, R., Bergland, F., Knokey, A.-M., Nagle, K. M., Sanford, C., Kalb, S. C., et al. (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.

Cameto, R., Knokey, A.-M., Nagle, K., Sanford, C., Blackorby, J., Sinclair, B., & Riley, D. (2009a). *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.

Cameto, R., Knokey, A.-M., Nagle, K., Sanford, C., Blackorby, J., Sinclair, B., & Riley, D. (2009b). *State profiles on alternate assessments based on alternate achievement standards. A report from the National Study on Alternate Assessments* (NCSER 2009-3013). 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*. New York, NY: National Center for Learning Disabilities. Retrieved from http://www.cehd.umn.edu/NCEO/OnlinePubs/NCLD/NCLDRewardsandRoadblocks.pdf

DeBarger, A. H., & Riconscente, M. (2005). *An example-based exploration of design patterns in measurement* (PADI Technical Report 8). Menlo Park, CA: SRI International. Retrieved from http://padi.sri.com/downloads/TR8_VersionForDesign.pdf

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.

Gong, B., & Marion, S. (2006). *Dealing with flexibility in assessments for students with significant cognitive disabilities*. Minneapolis: MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://cehd.umn.edu/NCEO/ OnlinePubs/Synthesis60.html

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. doi: 10.1177/002246699903300203

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). *An introduction to persons with moderate and severe disabilities: Educational and social issues* (2nd ed.). Boston, MA: Allyn and Bacon.

Messick, S. (1994). *Alternative Modes of Assessment, Uniform Standards of Validity* (No. ETS RR-94-60). Princeton, NJ: Educational Testing Service.

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

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6-20.

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

Quenemoen, M., 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, MN: University of Minnesota, National Center on Educational Outcomes.

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., 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

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, MN: University of Minnesota, National Center on Educational Outcomes.

Ryan, J. M., Quenemoen, R. F., & Thurlow, M. (2004). *I say potato, you say potato: The assessment-speak gao between general and alternative assessment experts*. Minneapolis, MN: National Center on Educational Outcomes.

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

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., & Altman, J. R. (2005). *2005 State special education outcomes: Steps forward in a decade of change*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://www.cehd.umn.edu/nceo/onlinePubs/2005StateReport.htm

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*(1), 45-57.

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

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

Wagner, M., Cadwallader, T. W., & Marder, C. (2003). *Life outside the classroom for youth with disabilities. A report from the National Longitudinal Transition Study-2 (NLTS2)*. Menlo Park, CA: SRI International.

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

Appendix A: Design Pattern Development Guidelines

The nine steps in the following pages describe the process to complete a *Design Pattern*. However, it is possible for the process to be more iterative than implied by these steps; that is, prior steps may be revisited and the *Design Pattern* refined accordingly to further specify attributes or make the assessment argument more explicit. The example described in the steps that follow was developed by the AAD-M project for the Number and Operations expectation: **"Develop understanding of fractions as parts of unit wholes, as parts of a collection, as locations on number lines, and as divisions of whole numbers."**

Step 1. Create a Title and Summary

The title is a name for the *Design Pattern* that briefly describes the content or skills addressed in it. It is important to adopt a naming convention and to consistently use it. For the AAD-M project, the title was comprised of three elements: the content area subdomain, the label or code of the standard or "expectation" addressed, and the grade-level range. For example, the title "Number and Operations A3 (grades3–5)" was created from the expectation of the NCTM *Principles and Standards for School Mathematics* mentioned above. The content area subdomain was Number and Operations. The next element in the title was the code for the NCTM expectation being addressed, A3, in which "A" referred to the second standard in the Number and Operations subdomain, and the "3" referred to the third expectation within that first standard.[7] The final element of the title was the grade-level range, grades 3–5.

The summary provides more detail about the scope or breadth of knowledge and skills to be addressed in the *Design Pattern*. To operationalize this attribute, the AAD-M project used the verbatim wording of the expectation from the NCTM *Principles and Standards for School Mathematics*. For instance, the summary for Number and Operations A3 (grades3–5) was "**Develop understanding of fractions as parts of unit wholes, as parts of a collection, as locations on number lines, and as divisions of whole numbers,**" which is the exact wording of the NCTM expectation.

Note that although the AAD-M project chose to use the NCTM standards and expectations to guide the work, other standards, including the state's extended standards or the Common Core State Standards, could serve as the base for the ECD approach.

Step 2. Add Relevant Educational Standards

In the AAD-M project, prior to developing the *Design Pattern*, the co-design team engaged an expert in mathematics to create a crosswalk linking NCTM expectations to state extended standards. Extended standards from Utah, Idaho, and Florida related to the NCTM expectation were included in the *Design Pattern* as a reference to show what mathematics content and skills each state identified as essential for students with significant cognitive disabilities.

Step 3. Develop Rationale Statement

The Rationale identifies why the construct(s) identified in the summary are important to assess. Creating a rationale statement requires input from a math content expert, who can situate the mathematics constructs targeted by the *Design Pattern* within the broader domain of

[7] The NCTM does not label specific standards and expectations within or across domains using the naming conventions described above (i.e., A3). This naming convention was a creation of the AAD-M project to distinguish among various standards and expectations within an NCTM subdomain (e.g., Number and Operations).

mathematics. For instance, the rationale statement for Number and Operations A3 (grades3–5) was "**Fractions represent a significant extension of children's knowledge about numbers. When children possess a sound understanding of fractions, they can use this knowledge to describe real world phenomena and apply it to problems involving measurement, probability, and statistics.**"

Step 4. Identify Focal Knowledge, Skills, and Abilities (Focal KSAs)

Standards are often written at a grain-size that is too large for assessment purposes. Focal KSAs reflect the standard when it is further unpacked into its essential, assessable elements. The content expert(s) on the co-design team draft Focal KSAs by reviewing the standard. In co-design meetings, Focal KSAs are discussed and further refined.

The focus and grain-size of the Focal KSAs need to be agreed upon by the co-design team. It is possible to generate multiple standards-based Focal KSAs, each of which only addresses one facet of a standard. For example, Number and Operations A3 (grades3–5), included six Focal KSAs each of which addresses a component of the expectation:

1. Ability to recognize a whole and divide it into or recognize equal parts (e.g., halves, thirds, or quarters)
2. Ability to identify fractions by the number of parts in the whole and in the fractional amount
3. Ability to identify a collection as a whole, and consider groups of objects in the collection as parts of the whole
4. Knowledge that fractions are numbers and identify points on the number line corresponding to particular fractions, between 0 and 1, and greater than 1
5. Knowledge that a division operation may not have a whole number result
6. Ability to solve problems involving fractions

It is critical to take the time with the co-design team to consider how the content or skills in the standard should be parsed because Focal KSAs will influence other attributes of the *Design Pattern*. Focal KSAs may be refined or deleted after their influence with respect to other attributes of the *Design Pattern* becomes more apparent. For example, in Step 4 Focal KSAs are "operationalized" when observations of student behaviors that are likely to provide evidence of each Focal KSA are specified. At this point the co-design team may realize that a Focal KSA is too vague or too complex to create these observations. If this occurs, the co-design team should revisit Step 3 and refine the Focal KSA. The co-design may also go back and add new Focal KSAs as they progress through the design steps.

Step 5. Develop Potential Observations and Potential Work Products

Potential Observations help to make each Focal KSA more concrete by describing the evidence (in the form of a specific student behavior) that indicates that a student has acquired the KSA. Potential Observations are phrased to describe the highest quality of student performance that demonstrates evidence of the Focal KSA. Qualifiers such as "accurate" and "correct" are used in all Potential Observation statements. Co-design teams also may find it helpful to generate specific examples for each Potential Observation (i.e., given a particular mathematics problem or context, describe the observed behavior). In constructing *Design Patterns* for the AAD-M project, the extended standards from each participating state also were considered when determining the range and qualities of behaviors that would likely be observed for students with significant cognitive disabilities.

Potential Work Products are descriptions of the form of the information that can be gathered from students (e.g., written explanation or selection of a response). When possible, work products should be stated such that they do not reflect bias in how students express their response. Often, Potential Observations can be expressed in multiple ways (e.g., in speech or in writing). Thus, the Potential Work Product "Expression of a mathematical pattern" is preferable to "Student writes the mathematical pattern," since not all students can write. However, in some cases, a Potential Work Product must be specific to a particular mode of expression for a Potential Observation. In these cases, additional Potential Observations and associated Potential Work Products should be specified that reflect alternate modes of expression.

A "horizontal view" of the *Design Pattern* is used during co-design meetings to make the connections among each Focal KSA and its associated Potential Observations and Potential Work Products explicit. Table 1 shows an excerpt of the "horizontal view" for Number and Operations A3 (grades3–5).

Table A-1. "Horizontal View" of Excerpt from Number and Operations A3 (grades 3–5) Design Pattern Focal KSAs, Potential Observations and Potential Work Products

Focal KSAs	Potential Observations	Potential Work Products
Ability to recognize a whole and divide it into or recognize equal parts (e.g., halves, thirds, or quarters)	Student correctly divides an object into a specified number of equal parts Student correctly identifies a pictorial representation of a fraction Student correctly distinguishes a whole from fractions of a whole	Student worksheet with multiple pictorial representations of fractions A whole object divided into fractional pieces Selection of a whole orange and fractional pieces of it

Step 6. Develop Characteristic Features of Tasks

In reviewing the Focal KSAs, Potential Observations and Potential Work Products, the co-design team identifies the key features of tasks that will be developed using a particular *Design Pattern*. These Characteristic Features must apply to all tasks created from a *Design Pattern*. For example, one Characteristic Feature developed for the Number and Operations A3 (grades3–5) *Design Pattern* is "**All problems will involve the use of fractions.**" In addition, Characteristic Features can define ways to constrain tasks in relation to the content (e.g., limitations on which numbers should be used). Characteristic Features also can pertain to more general task features desired in tasks associated with a *Design Pattern*. These may include task features such as prompting for individual student responses (not group responses), allowing accommodations, and involving a test administrator who knows the student's comprehensive/response abilities.

Step 7. Identify Cognitive Background Knowledge Additional KSAs

Steps 2–6 make explicit relationships among the standard (or, in the case of the AAD-M project, the NCTM expectation), the Focal KSAs, student behaviors and work products that provide evidence of the Focal KSAs, and characteristic features of tasks to elicit the desired student behaviors. In Step 7 the co-design team describes the Additional KSAs that are not construct relevant but may be required for successful performance on tasks associated with a particular *Design Pattern*.

To determine the Cognitive Background Knowledge Additional KSAs, the co-design team must consider the prerequisite knowledge and skills that may be needed for each Focal KSA. For example, the Number and Operations A3 (grades3–5) *Design Pattern* includes the Focal KSA, **Ability to recognize a whole and divide it into or recognize equal parts (e.g., halves, thirds, or quarters)**. In order for students to be able to demonstrate this ability, the co-design team determined that students may need additional background KSAs, such as:

- Ability to count using whole numbers
- Ability to use a number line to model whole numbers and operations on them
- Knowledge that there is "space" on the number line between each whole number
- Ability to perform division operations (e.g., grouping)

Step 8. Create Cognitive Background Knowledge Variable Features of Tasks

In order to prevent Cognitive Background Knowledge Additional KSAs from impinging on a student's ability to demonstrate what they know about the Focal KSAs, the co-design team considers how these Additional KSAs may be supported. These supports are Cognitive Background Knowledge Variable Features. For example, for Number and Operations A3 (grades3–5), the following Cognitive Background Knowledge Variable Features were identified:

- Provision of a table, chart, or tactile reminder of the numbers to support understanding that numbers occur in a specified sequence
- Supports for use of a number line (supported, unsupported, degree of support) (e.g., provide a number line, model use of a number line, re-teaching use of a number line just prior to assessment)
- Supports for division skills (supported, unsupported, degree of support) (e.g., provide a calculator, use number line to provide visual representation of division, counters, counter mats)

Step 9. Review and Select UDL Additional KSAs and Variable Features

In the AAD-M project six categories of UDL were used: (1) Perceptual (Receptive), (2) Skill and Fluency (Expressive), (3) Language and Symbols, (4) Cognitive, (5) Executive, and (6) Affective. UDL Additional KSAs are nonconstruct relevant knowledge, skills, and abilities in these categories that may be required for successful performance on tasks associated with a *Design Pattern*. UDL Variable Features are used to support student abilities associated with perceiving task stimuli, expressing responses to tasks, comprehending linguistic components of tasks, information processing, executive functioning, and engagement. Unlike the Cognitive Background Knowledge Additional KSAs and Variable Features, which are created afresh for each Design Pattern, the UDL Additional KSAs and associated Variable Features have been standardized and are prepopulated in each Design Pattern (for a full list of these associations, see *Implementing Evidence-Centered Design to Develop Assessments for Students with Significant Cognitive Disabilities: Guidelines for Creating Design Patterns and Development Specifications and Exemplar Task Templates for Mathematics*). The co-design team is responsible for reviewing this standardized list and selecting those Additional KSAs and associated Variable Features that are most relevant for the task.

Appendix B: Example Development Specifications and Exemplar Task Template

Task/Item Development

Number & Operations A3 (grades 3-5)

Attributes	General Information
Title	Number & Operations A3 (grades 3-5)
Summary	Develop understanding of fractions as parts of unit wholes, as parts of a collection, as locations on number lines, and as divisions of whole numbers
Rationale	Fractions represent a significant extension of children's knowledge about numbers. When children possess a sound understanding of fractions, they can use this knowledge to describe real world phenomena and apply it to problems involving measurement, probability, and statistics.
Grade level standards (from NCTM)	Understand the structure of numbers and the relationships among numbers. Explore a variety of models of fractions focused on familiar fractions: halves, thirds, fourths, fifths, sixths, eighths, and tenths. Develop strategies for ordering and comparing fractions using benchmark fractions such as ½ and 1. Use parallel number lines show a unit fraction and its multiples.

Attributes	Definition	**Item 1** Application/Comprehension/ Performance	**Item 2** Performance/Recall	**Item 3a/3b** Recall/Attention
Focal KSA	Focal KSA from DP for Items 1 & 2; Add'l KSA from DP for Item 3	• Ability to identify fractions by the number of parts in the whole and in the fractional amount. (This focal KSA deals with both the numerator and the denominator)		• Knowledge that a whole number can be divided into fractions of the whole.
Potential Observations from DP	Observed behaviors of students that can provide evidence of the Focal KSA	• Student correctly matches fraction to pictorial or object representation of that amount • Student correctly names the fraction represented by a picture or object		*Not addressed in DP*
Potential Work Products	What students say, do, or make that provides evidence about the Focal KSA	• Student worksheet that depicts one half, one quarter, and three quarters of a pizza and student indicates which picture represents ¾ • Student worksheet with one picture of a portion of a pizza in slices and student indicates fraction represented		*Not addressed in DP*
Characteristic Features	Aspects of assessment situations likely to evoke the desired evidence	• All problems will involve the use of fractions		

Attributes	Definition	Item 1 Application/Comprehension/ Performance	Item 2 Performance/Recall	Item 3a/3b Recall/Attention
Potential Variable Features/ Scaffolding	Features that could be changed to impact item difficulty and scope	• The number of representations presented to the student. • Models (fraction circles, card board representations, other manipulatives) pictures. • Types of representation: Presents fractions or wholes. • Presentation of fraction: Verbal, symbolic • Size of the denominator (2, 3, or 4) • Depth of knowledge of the content (e.g., fractions used (halves, thirds, quarters, etc.)	• The number of representations presented to the student. • Models (fraction circles, card board representations, other manipulatives) pictures. • Types of representation: Presents fractions or wholes. • Presentation of fraction: Verbal, symbolic • Size of the denominator (2, 3, or 4) • Depth of knowledge of the content (e.g., fractions used (halves, thirds, quarters, etc.)	• The number of representations presented to the student. • Models (fraction circles, card board representations, other manipulatives) pictures • Types of representation: Presents fractions or wholes. • Presentation of fraction: Verbal, symbolic • Size of the denominator (2, 3, or 4) • Depth of knowledge of the content (e.g., fractions used (halves, thirds, quarters, etc.)
Selected Variable Features/ Scaffolding for the Item	From Item 1 to Item 3: • Reduce DOK • Reduce scope • Increase scaffolding	• 3 representations presented • Model: Drawing of fraction circles • Presents 3 models each of which is a fraction • Presents verbal and symbolic representations of numeric fractions • DOK math content: ¼, 2/4, ¾ • DOK level: Comprehension (Translate) • Scaffolding: - Use of a diagram - Multiple representations of fractions	• 2 representations presented • Model: Photograph of pizza/pie • Presents 2 models one of which one is a fraction and one is a whole • Presents verbal and symbolic representations of numeric fractions • DOK math content: ½ and 1 whole • DOK level: Performance (Locate) • Scaffolding: - Use of photograph of familiar stimuli - Multiple representations of fractions	• 2 representations presented • Model: Photograph of pizza/pie • Does not present verbal and symbolic representations of numeric fractions • Presents verbal and symbolic representation of part versus whole • DOK math content: 1 whole versus part of a whole (1/4, ½, ¾) • DOK level: Recall (Identify "not a whole") • Scaffolding: - Use of photograph of familiar stimuli - Multiple representation of part vs. whole

Attributes	Definition	Item 1 Application/Comprehension/ Performance	Item 2 Performance/Recall	Item 3a/3b Recall/Attention
Item Directive	The stem or question (includes description and number of distractors if applicable)	Examiner presents student with three drawings of pizzas/pies and says, "**Here are three drawings of parts of a pizza/pie**." Examiner then presents a card with the numeric fraction "¾" down on the table in front of the student and says, "**Which drawing shows three fourths of a pizza/pie?**"	Examiner presents students with two photos, one with a whole pizza/pie and the other with a half of a pizza/pie. Examiner says, **"Here are two photos of pizzas/pies."** Examiner then presents a card with the numeric fraction "½"down on the table in front of the student and says, **"Show me which photo shows half of a pizza/pie."**	3a) Examiner presents two unlabeled pizza/pies in a row to student and says, **"Here are two photos of pizza/pie (a whole pizza/pie and part of a pizza/pie). Show me the photo that is a part of a pizza/pie."** 3b) If student cannot respond, remove all stimuli but the photo of part of a pizza/pie. Teacher says, **"Look at/touch the photo of part of the pizza/pie."**
Correct Answers	Correct answer for the item	Student indicates ¾ of a pizza/pie picture on the worksheet	Student indicates ½ pizza/pie picture	3a) Student indicates the picture that is not a whole pizza/pie. 3b) Student looks/touch the picture of a half of pizza/pie
Description of Stimulus Items	Description of the *graphics or objects* used in administration of the task	Three **unlabeled** drawings of three pizzas/pies presented in a row (these are bird's eye view of the pizza/pie, not a side view with perspective). Each pizza/pie is divided into four slices. One has 2 of the four remaining, one has one of the four remaining, and the third has three of the four remaining. Every pizza/pie shows the quarters outlined with a dotted line. Every pizza/pie has four sections outlines even if they are missing. The teacher has a card with ¾ on it.	Two **unlabeled** photos of two pizzas/pies presented in a row (these are bird's eye view of the pizza/pie, not a side view with perspective). Teacher has card with ½ on it. One photo is of a whole pizza/pie and the other photo is of a half of pizza/pie.	Two **unlabeled** photos of two pizza/pies presented in a row (these are bird's eye view of the pizza/pie, not a side view with perspective). One photo is of a whole pizza/pie. The second photo is of half of pizza/pie.

Attributes	Definition	Item 1 Application/Comprehension/ Performance	Item 2 Performance/Recall	Item 3a/3b Recall/Attention
Materials for Examiner	Materials required to administer, document, and score the task (e.g., worksheet, camera to take picture of product, manipulatives)	3 pictures of pizzas/pies 1 card with numeric ¾ Recording sheet for teacher to complete	2 photographs of pizzas/pies 1 card with numeric ½ Recording sheet for teacher to complete	2 photographs of pizzas/pies Recording sheet for teacher to complete
Variable Features for Administra-tion to Individual Students	Features that could be changed to impact item accessibility for individual student needs (e.g., as specified in the student's IEP)	• Question presentation individualized (e.g., related in sign language) • Response format individualized based on student communication system • Remind student of prior experiences • Verbal/gestural prompts individualized • Use of tactile graphics	• Question presentation individualized (e.g., related in sign language) • Response format individualized based on student communication system • Remind student of prior experiences • Verbal/gestural prompts individualized • Use of tactile graphics	• Question presentation individualized (e.g., related in sign language) • Response format individualized based on student communication system • Remind student of prior experiences • Verbal/gestural prompts individualized • Use of tactile graphics

Updated Flowers/Browder Math DOK[8]:

1. **Attention:** touch, look, listen, repeat what the teacher said, vocalize, respond, attend, recognize
2. **Memorize/recall:** list, describe (facts), state math facts, identify, state, define, match, recognize, label, follow a pattern
3. **Performance:** answer, follow 1 step directions, find answer, present, read, separate, spell, tell time, map, model demonstration, perform, demonstrate, follow, choose, count, locate, group by given attributes, solve simple (one computation skill) problems, measure
4. **Comprehension:** understand, extend a pattern, sketch, ask and answer questions, categorize/group by unknown attributes, explain, conclude, group, restate, review, translate, classify/sort with understanding, simplify (equivalent forms)
5. **Application:** compute, organize, collect (such as data), apply, revise, construct, solve complex (multiple computation skills) problems, use given formulas in novel situations (formula may or may not be identified), explain a process, conduct research
6. **Analysis, Synthesis, Evaluation:** create a complex pattern, analyze, compare, contrast, compose, predict, plan, judge, evaluate, interpret data, generalize findings, create hypotheses

---

8 Bechard, S., Almond, P., Karvonen, M., Wakeman, S., Turner, C., Bowen, T., & Turner, L. (2009). *Hitting a moving target: A discussion of ten alignment studies for AA-AAS.* Paper presented at the National Conference on Student Assessment. Los Angeles, CA June 23, 2009.

Appendix C: Development Specifications and Exemplar Task Template Development Guidelines

Once a *Design Pattern* has been reviewed and finalized, the creation of the *Development Specifications and Exemplar Task* Template can commence. In this section the methodology involved in this enterprise is described. Specifically, the following section provides guidelines and suggestions for the development of tasks.[9]

Step 1. Pre-populate Section A of the Development Specifications and Exemplar Task Template

Step 1 involves pre-populating some of the attributes within the *Development Specifications and Exemplar Task Template* with information taken directly from the *Design Pattern*. The first three attributes (Title, Summary, and Rationale) come directly from the associated *Design Pattern*.

Grade-Level Standards are included to provide background information about what general education children should be able to do in light of specific standards or expectations and at certain grade levels. For example, the following Grade-Level Standards information was included for Number and Operations A3 (grades 3–5), which focuses on understanding fractions:

- Understand the structure of numbers and the relationships among numbers
- Explore a variety of models of fractions focused on familiar fractions: halves, thirds, fourths, fifths, sixths, eighths, and tenths
- Develop strategies for ordering and comparing fractions using benchmark fractions such as ½ and 1
- Use parallel number lines to show a unit fraction and its multiples

Whether this information is extracted from a standards document or generated for the project, the informed perspective of a math education or content expert is required. The co-design teams refer to this information as they build items that are accessible and appropriate for the 1% population, while considering alignment to grade-level expectations for the general education population.

Step 2. Pre-populate Section B of the Development Specifications and Exemplar Task Template

Step 2 involves pre-populating the attributes within Section B of the *Development Specifications and Exemplar Task Template* with information taken directly from the *Design Pattern*. In this step all Focal KSAs, Cognitive Background Knowledge Additional KSAs, Potential Observations, Potential Work Products, and Characteristic Features are copied from the associate *Design Pattern* into the *Development Specifications and Exemplar Task Template*.

Finally, the UDL Variable Features selected in the *Design Pattern* as most relevant for the task are copied into the Potential Variable Features section of the *Development Specifications and Exemplar Task Template.* These Variable Features will be “set” as part of the item development process to document precisely how task features are manipulated to influence item difficulty. For instance, in Number and Operations A3, grades 3–5, the Potential Variable Features included:

- Number of representations presented to the student

[9] The methodology we describe here is a result of our experiences in developing the AAD-M project’s task design and development specifications template. Although there may be minor variations among different co-design teams in their implementation, these are the general guidelines that were followed.

- Models (fraction circles, card board representations, other manipulatives) and pictures
- Types of representation (fractions or wholes)
- Presentation of fraction (verbal, symbolic)
- Size of the denominator (2, 3, or 4)
- DOK of the content (e.g., fractions used [halves, thirds, quarters, etc.])

Step 3. Review and/or Revise the Pre-populated Attributes in Section B

It is an important and necessary step to review the pre-populated components of the template, as well as reflect again on the extended standards aligned to the standard/expectation being addressed. This reflection provides a sense of how and if the participating states are currently assessing content related to the expectation and also helps to shed light on how the emphases placed on a particular expectation may vary by state.

From among the Focal KSAs, the co-design team will select the Focal KSA that will serve as the foundation for Items 1 and 2. The choice of the Focal KSA can depend on several factors:

- Alignment of the particular focal KSA to the intended emphases of the state's extended standards.
- Complexity of the KSA (e.g., number of steps involved, level of cognitive skill required, and whether this level is appropriate for the target population). During the ECD design pattern process, the expectation is deconstructed into a set of distinct focal KSAs. Some co-design teams may prefer to select more fine-grained or more comprehensive Focal KSAs.
- Clarity or relative simplicity of the intended KSA to be assessed.
- Feasibility for developing tasks that can be "worked down"[10] to encourage content accessibility for a wider spectrum of the target population.

Once the Focal KSA is selected, the next step is to determine the Potential Observations and Potential Work Products that will be targeted for Items 1 and 2. Within the design pattern each Focal KSA is associated with one or more Potential Observations (i.e., which represent different ways of gathering evidence of the focal KSA) and one or more corresponding Potential Work Products. A decision must be made about which Potential Observation and Potential Work Product will be used to provide evidence about the chosen Focal KSA. Although it is usually the case that the Potential Observation for Items 1 and 2 is selected from the list of Potential Observations detailed in the *Design Pattern* for the chosen Focal KSA, the co-design team may identify others at this point. If there is not a Potential Observation and/or Potential Work Product within the list from the *Design Pattern*, then a more appropriate Potential Observation and/or Potential Work Product that embodies the Focal KSA can be suggested, selected, and subsequently added to the *Development Specifications and Exemplar Task Template*. The selection of the Potential Observation and Potential Work Product may depend on several factors including:

- Cognitive complexity of the observed behavior for the target population (e.g., number of steps or skills involved in providing an answer)
- How characteristics of students from this population might limit their ability to demonstrate evidence about their knowledge in a specific way

[10] Browder uses the phrase "work it down" to describe how to develop alternate assessments (AA) for students with significant cognitive disabilities that are linked to grade-level academic content standards. She suggests starting with content standards at grade level then considering how items can be translated so that students at different levels of functioning or communication would be able to access it.

Once the Focal KSA, Potential Observation, and Potential Work Product are decided upon, Characteristic Features are reviewed to remind the co-design team about the critical task features that must be present. Potential Variable Features are also reviewed so that the co-design team can consider possible ways to vary the four items. It is possible that the co-design team will propose additional Characteristic Features and Potential Variable Features. If it is determined that a proposed Characteristic Feature (not already within the *Design Pattern*) applies to all tasks created from a *Design Pattern*, it should be added.

For consistency the co-design team should update the *Design Pattern* by adding any new Potential Observations, Potential Work Products, Characteristic Features, and Variable Features that are generated during the task development process. Consistency of content between the *Design Pattern* and *Development Specifications and Exemplar Task Template* is critical. Note that this reconsideration or revision to the *Design Pattern* illustrates the iterative nature of the ECD process for developing both *Design Patterns* and *Development Specifications and Exemplar Task Templates*.

Step 4. Determine the Task Requirements for the Item

As items are created it is important to keep the following considerations in mind:

Presence of context—A decision must be made about whether to include a context or to present the task in a decontextualized fashion. For example, if the Focal KSA aims to assess the students' ability to calculate summary statistics, a contextualized item can be developed, "In a recycling contest, students collected aluminum cans. This data table shows how many aluminum cans each student collected. What is the mean number of cans collected?" Alternatively, a decontextualized item can be developed, "Using these 10 data points, calculate the mean." Including context can make an item more interesting and engaging to students, but it can also increase the cognitive demand in a nonconstruct relevant way. If the decision is to have context present, here are further considerations in choosing one that is appropriate:

- **Choose a context that is grade-level appropriate and respectful. For instance, when targeting** students in the grade 9–12 range, a recycling contest was the chosen context for students to demonstrate their ability to answer a question about data by identifying, creating, and using a graphical display, and calculating and using a summary statistic. Although the use of a marbles contest could allow the assessment of the same mathematical skill, it would not have been grade-level appropriate.
- Establish a context that is realistic where possible. For example, if inches of rainfall during the year is the chosen context, the data points included should reflect what is typical and realistic.
- Ensure concrete examples are used in the context where possible. For example, discuss mathematical relations in the context of everyday situations.
- Use a context that is generalized where possible. For example, when discussing rainfall, instead of referring specifically to rainfall within a particular state (e.g., Hawaii or Florida), it is important to discuss rainfall in general so that the technical accuracy of the information (i.e., knowing the amount of rainfall that occurs in a particular state) is not the subject of the question.
- Choose a context that is clear and unambiguous.

Student response mode—A decision must be made about whether students will be asked to select the correct response from a set of response options or whether the student will be asked to construct the correct response on their own. If the student is asked to construct the correct response, another decision must be made about whether students would be asked to construct a verbal response, a graphical representation, a computer generated response, a concrete representation of their response, or a written response. The assessment designer must consider the relationship between the response mode required and the specific cognitive limitations of the students. It is possible that although an item may be designed with a particular response mode in mind, it may need to be modified by the test administrator at the time of administration given a particular student's capabilities.

Presence of data—A decision must be made about whether data will accompany the text, and if so, the following questions should be considered:

- Should the data be presented within a table, graphically, or in a list?
- If data is presented graphically, what type of display should be used (e.g., line graph, pie chart, bar graph)?
- Should the data be rich enough to allow the assessment designers to ask a range of nontrivial or interesting questions or should the data be limited to a specific question without extraneous information, relationships, or variables illustrated?
- How many data points should be presented?
- Should single and/or double-digit numbers be included (e.g., 9, 14)?
- Should categorical and/or numerical data be presented? ;
- How complex should the highlighted relationship be in the data distribution?

Number of questions within an item—A decision must be made about whether one question or multiple questions should be asked of the student. This may depend on the complexity of the Focal KSA and on the approaches states are using and whether item interdependency can be addressed in their measurement model.

- If multiple questions are asked, should they be asked in the same context and/or data set or should multiple contexts and/or data sets be progressively built into the items?
- Should an overall framing or thematic question be included when multiple questions are asked?

Number of steps to the solution—A decision must be made about how many steps should be involved in getting to the final solution.

Step 5. Develop the Item Directive

In Section C of the template specific task information is generated and recorded. This information will be recorded for each of the 4 items within the task. It is suggested that co-design teams work through steps 5–9 for Item 1, then go back and repeat these steps for Item 2, and finally go through them again to create Items 3a and 3b.

The Item Directive segment of the template includes the item prompt or question, the item description and distracters when applicable, and specific instructions that will be presented to students for each item. For the AAD-M project the convention was adopted that text in bold was to be read aloud by the examiner. The Item Directive does not detail specific individual adjustments that can be made (and that are acceptable) in the task administration. This information concerning individual adjustments is presented in the Variable Features for Administration to Individual Students section of the template (described in detail in step 9).

The mathematics expert within the co-design team typically suggests an idea for the Item Directive, taking into consideration the Focal KSA, the decisions made about the task requirements, their experience in the classroom, and the best way to assess the mathematics concepts targeted.

After drafting an initial representation of the idea for the Item Directive, the team discusses and modifies the Item Directive based on insights from differing perspectives, such as the principles of ECD, mathematics education, and classroom experience with special education students. The concerns addressed in these discussions should include:

- Capabilities of students in the target population
- Construct relevant and irrelevant details elicited by the proposed Item Directive:
  - Whether the proposed Item Directive adheres to the Focal KSA
  - What Additional KSAs might be required by the task
  - How to minimize or support the Additional KSAs within the design of the Item Directive
- Evaluation of the content of the Item Directive:
  - Context (see criteria in step 4)
  - Data presentation (see criteria in step 4)

To illustrate this process, the following is an example of the Item Directive for Item 1 of Number and Operations A3 (grades 3–5), which will be further elaborated in steps 6–7. The examiner presents students with three drawings of pizzas/pies and says, "**Here are three drawings of parts of a pizza**." The examiner then presents a card with the numeric fraction "¾", places it on the table in front of the student, and then asks, "**Which drawing shows three-fourths of a pizza?**" (Note: Some items may include multiple options for context information. These options are placed within square brackets [ ]. They are provided within the Item Directive to allow for maximum flexibility and appropriateness according to specific characteristics of the population. For instance, a board game [as opposed to a video game] may be a more appropriate example of a prize for populations from lower socioeconomic backgrounds; hence, it is provided as a possible replacement option).

Step 6. Document the Correct Answer

After the co-design team has reached consensus on the Item Directive, they next document the Correct Answer. The answer can be a number, graph, or description. The team should also specify whether alternative versions of the stated correct answer are also acceptable. For example, for the item created for Number and Operations A3, grades 3–5, the Correct Answer to the Item Directive (e.g., Which drawing shows three-fourths of a pizza?) is "**Student indicates the picture of ¾ of a pizza**."

Step 7. Describe the Stimulus Items and Materials for the Examiner

The Description of the Stimulus Items is a depiction or detailed description of the graphics, objects, or tools to be used in task administration. This might include a table of data presented to the student with which they must create graphics or interpret, synthesize, and/or calculate statistics. If there are multiple questions within an item, there will be a description of the stimulus materials for each question. The Stimulus Materials for Item 1 of Number and Operations A3, grades 3–5 include:

- Illustration of three pizzas divided into quarters; one has two quarters remaining, one has one quarter remaining, and the third has three quarters remaining
- Note card with the numeric fraction ¾

The Materials for the Examiner is a description of the materials examiners will need to administer, document, and score an item (e.g., worksheet, camera with which to take a picture of product, or a manipulative). It includes the task worksheet that describes the item and delivery instructions and task data sheet or other method to record the student's response.

Step 8. Update Selected Variable Features

The co-design team must return to the Selected Variable Features to update the information based on the selections made for the finalized Item Directive. The team first decides on the DOK level for the item. Using the 6-point DOK scale (Flowers et al., 2007), the team decides which level best exemplifies the DOK required by the Item Directive created for the item. This decision is based on a number of factors including:

- **Understanding of the structure** of the DOK levels and the verbs used to exemplify each level, including how each level and verb can be operationalized generally in the context of mathematics and more specifically in the context of the item. For instance, an item that asks students to explain and/or make a conclusion is considered to be at the comprehension level.
- **Determining the mathematical sophistication** of what is elicited by the item based on the abstract nature of the mathematics concept being probed based on (1) the amount of prior mathematics knowledge that has to be drawn upon, (2) the number of mathematical principles required for the solution, and (3) whether the question can be answered with a procedure or routine.
- **Determining the complexity** of what is elicited by the item based on (1) whether the student has to extend or produce novel findings, (2) whether the item has multiple questions or requires multiple or integrated skills, and (3) whether the answer is a constructed response or selected response. In addition, the distracters in a selected response item can be written to impact the item's complexity.

If the DOK assigned to Item 1 is lower than desired, the team may decide to use Item 1 as an Item 2 or may revise the Item Directive to increase the DOK level of the item.

The co-design team should explicitly detail the decisions made for each Variable Feature selected to create the Item Directive. For instance, if the co-design team chooses to ask students to create a histogram (rather than a scatter plot or box plot), then they must document this decision.

Step 9. Document Variable Features for Administration to Individual Students

Variable Features for Administration to Individual Students specify task features that could be changed to impact item accessibility according to individual student needs (e.g., large print, Braille for those with visual impairments). Although the Item Directive will not be modified, it is possible that certain students will require specific accommodations in addition to the accessibility and scaffolding features built in to the design of the item. The boundaries of this category will be determined in part by accommodation policies in individual states. However, it is certain that these Variable Features should not compromise the construct (Focal KSA) targeted. Currently, two types of Variable Features for Administration to Individual Students have been consistently noted in the *Development Specifications and Exemplar Task Template*: (1) the freedom to vary the format of the question presentation (e.g., presented in sign language with Braille, auditory, or with or without gestural prompts) and (2) the students' response format individualized based on their communication system. States need to specify which accommodations or formats are and are not allowed

Step 10. Repeat Steps 5–9 to Develop Item 2

The co-design team should repeat Steps 5–9 to develop Item 2. Item 2 must assess the same Focal KSA as Item 1, but it involves skills that are considered to be at a lower DOK level. In addition, Item 2 is typically less complex, more narrow in scope, and more heavily scaffolded or supported. In creating Item 2, the modifications below should be kept in mind. These modifications help to ensure that the DOK and scope have been appropriately decreased and that supports or scaffolding have been appropriately increased relative to Item 1 while still preserving the Focal KSA.

- **Reduce DOK Levels:**
  - If Item 1 required students to *construct* a response (a higher DOK level), in Item 2 students can be asked to *select* the appropriate answer from a set of response options (a lower DOK level).
- **Reduce Complexity**:
  - If Item 1 asked students to create a scatter plot or box plot, Item 2 can ask for the creation of a histogram, which is technically less sophisticated. A histogram is focused on the frequency of one variable, while a scatter plot is about the relationship between two variables.
  - If Item 1 presents 20 data points to be mathematically represented, Item 2 could present only 10 data points.
  - If Item 1 contains 4 subquestions (i.e., a, b, c, and d), Item 2 could contain only *2 questions* (i.e., a, b).
- **Narrow the Scope of Content to Be Assessed:** If Item 1 assessed a composite set of skills (e.g., students determine the appropriate representation to be used to answer a research question, create that representation, and then use the representation to answer the research question), then Item 2 should assess fewer components of those skills (e.g., perhaps students just create and use the representation).
- **Increase Scaffolding or Support:** If the Focal KSA is about creating mathematical representations, Item 1 might ask students to create the representation with little support. Item 2 will increase the amounts and kinds of scaffolding within the design of the item. For instance, students could be provided graph paper to support the creation of a graphical representation (e.g., histogram) and/or students could be provided with key elements of the graph already completed (e.g., axes, labeled axes, and bins).

Step 11. Repeat Steps 5–9 to Develop Items 3a and 3b

Steps 5–9 also should be followed to complete Items 3a and 3b to ensure systematic development and documentation of design decisions for these items. However, recall that for these items an Additional KSA (not the Focal KSA) is targeted.

Some important considerations developing Item 3a are as follows:

- For consistency, select an Additional KSA that is aligned to the selected Focal KSA.
- The choice and use of an Additional KSA (or prerequisite skill) that is narrowly focused increases the likelihood that the item is less sophisticated than Items 1 and 2.
- Ensure that students at the lower functioning end of the spectrum of students with significant disabilities are taken into account in the design of this item.

Item 3b targets the attention DOK level. This usually involves removing all distracters from Item 3a and leaving only the correct answer for the student. The student is asked to point to or otherwise indicate the remaining stimulus item. This item is included in an effort to ensure that all students, including those with the most severe cognitive disabilities, will be able to participate in the testing experience and encounter some success.

Appendix D: NCTM Expectations Associated with Core and Additional Tasks for Pilot Testing

CORE TASKS

Number and Operations
1. A3 grades 3–5: Develop understanding of fractions as parts of unit wholes, as parts of a collection, as locations on number lines, and as divisions of whole numbers
2. A1 grades 6–8: Work flexibly with fractions, decimals, and percents to solve problems

Algebra
3. B2 grades 3–5: Represent the idea of a variable as an unknown quantity using a letter or a symbol
4. B1 grades 6–8: Develop an initial conceptual understanding of different uses of variables
5. B3 grades 9–12: Use symbolic algebra to represent and explain mathematical relationships

Geometry
6. A1 grades 3–5: Identify, compare, and analyze attributes of two- and three-dimensional shapes and develop vocabulary to describe the attributes
7. A1 grades 6–8: Precisely describe, classify, and understand relationships among types of two- and three-dimensional objects using their defining properties
8. A1 grades 9–12: Analyze properties and determine attributes of two- and three- dimensional objects

Measurement
9. B2 grades 3–5: Select and apply appropriate standard units and tools to measure length, area, volume, weight, time, temperature, and the size of angles

ADDITIONAL TASKS

Utah:
Number and Operations
1. A1 grades 3–5: Understand the place-value structure of the base-ten number system and be able to represent and compare whole numbers and decimals
2. A2 grades 3–5: Recognize equivalent representations for the same number and generate them by decomposing and composing numbers
3. A7 grades 6–8: Develop meaning for integers and represent and compare quantities with them

Algebra
4. C1 grades 3–5: Model problem situations with objects and use representations such as graphs, tables, and equations to draw conclusions

Data Analysis and Probability
5. A2 grades 6–8: Select, create, and use appropriate graphical representations of data, including histograms, box plots, and scatter plots

Measurement
6. B3 grades 3–5: Select and use benchmarks to estimate measurements 7. A1 grades 9–12: Make decisions about units and scales that are appropriate for problem situations involving measurement

Florida:

Number and Operations
1. A4 grades 3–5: Use models, benchmarks, and equivalent forms to judge the size of fractions
2. B1 grades 3–5: Understand various meanings of multiplication and division
3. C4 grades 6–8: Develop, analyze, and explain methods for solving problems involving proportions, such as scaling and finding equivalent ratios
4. A1 grades 9–12: Develop a deeper understanding of very large and very small numbers and of various representations of them

Data Analysis and Probability
5. A3 grades 3–5: Represent data using tables and graphs such as line plots, bar graphs, and line graphs

Geometry
6. A4 grades 3–5: Explore congruence and similarity

Measurement
7. A2 grades 6–8: Understand relationships among units and convert from one unit to another within the same system

Idaho:

Number and Operations
1. B3 grades 3–5: Identify and use relationships between operations, such as division as the inverse of multiplication, to solve problems
2. C2 grades 3–5: Develop fluency in adding, subtracting, multiplying, and dividing whole numbers
3. C1 grades 9–12: Develop fluency in operations with real numbers, vectors, and matrices, using mental computation or paper-and-pencil calculations for simple cases and technology for more-complicated cases

Algebra
4. A1 grades 3–5: Describe, extend, and make generalizations about geometric and numeric pattern

Data Analysis and Probability
5. B1 grades 3–5: Describe the shape and important features of a set of data and compare related data sets, with an emphasis on how the data are distributed
6. B1 grades 9–12: For univariate measurement data, be able to display the distribution, describe its shape, and select and calculate summary statistics

Measurement
7. B2 grades 6–8: Select and apply techniques and tools to accurately find length, area, volume, and angle measures to appropriate levels of precision

Appendix E: Instructions for Reviewing Design Patterns and Task Templates for AAD-Mathematics EAG

1. To review the Design Pattern and its associated tasks created for each of 30 NCTM Expectations that were selected to be part of the EAG project, you will need to read a set of documents for each expectation. You have received one or more zip files that contain 30 folders. Each folder, one for each NCTM Expectation, contains a Design Pattern, Horizontal View 1 (Focal KSA [knowledge, skills, and abilities], Potential Observations and Potential Work Products), Horizontal View 2 (Additional KSAs and Variable Features), and the Task Template of the associated items (e.g., AlgB1 appears at the top of each document and is contained in the file name).

Review of Design Patterns and Horizontal Views

2. Familiarize yourself with the design pattern attributes by reviewing their definitions in the second column of the Design Pattern (the definition in the Summary row is an error; the Summary is actually the verbatim NCTM Expectation).
3. Examine the Design Pattern and *determine whether you agree with the identification of the Focal KSAs associated with the NCTM Expectations* in the Summary row.
4. Using Horizontal View 1 review the appropriateness of the Potential Observations associated with each Focal KSA. *Judge whether the Potential Observations are appropriate examples of the kinds of evidence that you would expect of students who had mastered the Focal KSAs*.
5. Using Horizontal View 1 review the appropriateness of the Work Products associated with each Focal KSA and Potential Observation. *Judge whether the Work Products are appropriate for collecting the kinds of evidence identified in the Potential Observations*.
6. Using the Design Pattern, determine *whether you agree with the Cognitive Background Knowledge Additional KSAs that have been identified as prerequisite or precursor knowledge to the Focal KSAs*. Remember that Additional KSAs are required for successful performance on assessment tasks, but are not the target of the assessment.
7. Using the Design Pattern, review the remaining 6 categories of Additional KSAs. These 6 categories represent aspects of Universal Design for Learning (UDL), including Perceptual/Receptive, Skill and Fluency/Expressive, Language and Symbols, Cognitive Processing, Executive Processing and Affective Skills. *Judge the importance of these Additional KSAs for the performance of students with significant cognitive disabilities on tasks associated with the NCTM Expectation*.
8. Using Horizontal View 2 (ignore the highlighting in these views), review the appropriateness of the Variable Features that will be used to support students' performance on the Additional KSAs. The relationship between the Variable Features and Additional KSAs are unique for the Cognitive Background Knowledge section of each Design Pattern (review these for each Design Pattern). *Judge whether the Variable Features are adequate supports for the Cognitive Background Knowledge Additional KSAs.* The relationship between Variable Features and Additional KSAs for the 6 UDL Additional KSAs are standardized across all Design Patterns (review this relationship only once). *Judge whether the Variable Features are adequate supports for the Additional KSAs in each of the 6 UDL categories. If the Variable Features were built into the assessment tasks, would they mitigate the barriers to performance created by the Additional KSAs?*

9. Using the Design Pattern, note the Characteristic Features that would apply to all tasks created from a Design Pattern. These features help ensure that the constructs being assessed by the Focal KSAs are likely to be elicited by the tasks.
10. FYI, a list of each state's extended educational standards that were determined to be common across the states participating in the EAG project and linked to the NCTM Expectation is presented at the end of the each Design Pattern. These extended standards were judged to be related to the NCTM Expectation based on a crosswalk conducted at the beginning of the project. No review is necessary.

Review of Task Templates

1. Familiarize yourself with the Task Template attributes by reviewing their definitions in the first column. Remember the Summary row is the NCTM Expectation
2. The general information at the top of the Task Template (in green) provides the title of the Design Pattern associated with this suite of tasks, the grade levels for which the tasks were designed, the NCTM Expectation, the Rationale for the Expectation and a list of grade level standards from NCTM associated with this Expectation.
3. The white band in the Task Template is a header that describes the items that are created as exemplar tasks. The Flowers, et al. depth of knowledge (DOK) levels in mathematics for students with significant cognitive disabilities are presented for each of items 1, 2, 3a and 3b. The complete set of DOK levels are specified at the bottom of the Task (Webb's DOK levels for general education students are also specified). Item 3b is an activity at the "attention" level for students who are unable to perform the task at the "recall" level of item 3a.
4. Items 1 and 2 are designed to assess one Focal KSA selected from the associated Design Pattern. Items 3a and 3b are designed to assess a Cognitive Background Additional KSA from the same Design Pattern.
5. Look at the pink section of the Task Template; this section contains the Focal and Additional KSAs, Potential Observation, and Potential Work Product that were selected from the Design Pattern for use in the task design. All of the Characteristic Features from the Design Pattern are carried over to this section. The Potential Variable Features in the pink section include task features that could be varied to influence the difficulty level of each item. These Variable Features are sometimes noted during the Design Pattern development and sometimes during initial Task Development. The information in this section guides subsequent task development.
6. Look at the blue section; this section contains the Selected Variable Features for each of the items 1, 2, 3a and 3b, the Item Directives, the Correct Answer, the Description of Stimulus Items, Materials for the Examiner, and Variable Features for Administration to Individual Students. These are the components of a task. In reviewing the Task Template, *consider whether these components adequately address the Focal KSA (for items 1 and 2) and the Cognitive Background Additional KSA (items 3a and 3b). Are the items specified in the Task Template appropriate for students with significant cognitive disabilities at the grade levels indicated? Are the items well aligned to the specifications in the Design Pattern?*