National Center and State Collaborative General Supervision Enhancement Grant (NCSC GSEG)

A proposal submitted in response to the Request for Proposals under the General Supervision Enhancement Grants Program, CFDA 84.373x

Technical Narrative

Submitted by the National Center on Educational Outcomes

July 22, 2010

DISCLAIMER: AS OF JANUARY 1, 2011, THE NCSC PROJECT WORKPLAN WILL CONTINUE TO EVOLVE AS THE PROJECT PARTNERS TEST OUR ASSUMPTIONS AND RESHAPE TO MEET NEEDS. THIS IS A REFLECTION OF THINKING AT THE TIME THE PROPOSAL WAS SUBMITTED, AND IS SUBJECT TO CHANGE. WORK DONE AFTER OCTOBER 1, 2010 SUBJECT TO DISCLAIMER BELOW:

The project is funded under a grant from the Department of Education (PR/Award #: H373X100002, Project Officer, Susan.Weigert@Ed.gov). However, the contents of the proposal draft do not necessarily represent the policy of the Department of Education and no assumption of endorsement by the Federal government should be made.

National Center and State Collaborative General Supervision Enhancement Grant (NCSC GSEG) Technical Narrative Table of Contents

Chapter 1 – Significance

During the past decade, we have learned much about alternate assessments based on alternate achievement standards (AA-AAS) and the students with the most significant cognitive disabilities (SSCD) who participate in them. We have faced many challenges in developing an AA-AAS that is both appropriate for these students and technically defensible. Variations still exist in the meaning of assessment results from state to state and in the quality of curriculum and instruction. There is limited consensus on academic expectations and outcomes for SSCD.

The National Center and State Collaborative General Supervision Enhancement Grant Project (NCSC) will apply the lessons learned from the past decade of research on AA-AAS to address these challenges by developing a multi-state comprehensive assessment system for SSCD. The project will draw on this strong research base to provide an AA-AAS that is built from the ground up on powerful validity arguments linked to clear learning outcomes and defensible assessment results, to complement the work of the Race to the Top Common State Assessment Program (RTTA) consortia. In addition, we will develop supportive interim and formative assessment tools and management systems to ease the burdens of administration and documentation. The new AA-AAS system will be designed to measure individual student growth and produce aggregate scores to fully address other demands being placed on these assessments through Race to the Top innovations and ESEA reauthorization.

We know that a strong assessment design alone is insufficient to ensure either implementation fidelity or the provision of appropriate curriculum and instructional support systems. We propose a comprehensive system that will coherently address curriculum, instruction, and assessment needs in states by 1) producing technically defensible formative, interim, and summative assessments; 2) incorporating evidence-based instruction and curriculum

models; and 3) developing comprehensive approaches to professional development (PD). By drawing on a strong research base to produce curriculum and instructional materials and PD supports, NCSC will support educators as they plan for and provide appropriate instruction that addresses the Common Core State Standards (CCSS). These supports will help educators and Individualized Education Program (IEP) teams accurately identify the learner characteristics of SSCD and make appropriate decisions about how each student participates in the overall system of assessments. These supports will improve the alignment of the entire system and improve the validity of inferences of the system of assessments.

The NCSC organizational partners understand the complexities of state policies and contexts around AA-AAS and are committed to working collaboratively from the best available research and practice while acknowledging a continuing need for innovation and flexibility. The project's organizational structures, development processes, and deliverables reflect this commitment. The organizational partners include the National Center on Educational Outcomes (NCEO) as the host and fiscal agent, along with the National Center for the Improvement of Educational Assessment (NCIEA), staff from the National Alternate Assessment Center (NAAC) at the Universities of Kentucky (UKY) and North Carolina, Charlotte (UNCC), edCount, LLC, and 19 state partners: Alaska, Arizona, Connecticut, District of Columbia, Florida, Georgia, Indiana, Louisiana, Massachusetts, Nevada, New York, North Dakota, Pacific Assessment Consortium (PAC-6)[1], Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, and Wyoming. These states vary in size, demographics, and past AA-AAS practices. All partners share a commitment to the research-to-practice focus of the project and the development of a comprehensive model of curriculum, instruction, assessment, and supportive PD (See Letters of

[1] The 6 entities (AS, CNMI, FSM, GU, Palau, RMI) partner as 1 state, led by UG CEDDERS.

Support Section, citing agreement with NCSC project Theory of Action (TOA); see Appendix A for key descriptive information about the current status of AA-AAS in these states, including current approach, scoring methodology, and Annual Performance Report data (Altman, Thurlow, & Vang, 2010; Quenemoen, Perie, & Kearns, in press; Quenemoen, Quenemoen, Kearns, & Kennedy, in press). Thoughtful, intensive state transition planning will guide overall project development and implementation as well as state-by-state shifts to the new system of standards and assessments, based on ongoing needs assessments and changing economic and political circumstances in each state.

A system built on research-based understanding. The collected research of the NCSC partners and the larger field, documented in the comprehensive project bibliography in Appendix B, will frame our collaborative work. The bibliography is not summarized here, but the body of work is represented on our partner Web sites, including www.nceo.info, www.nciea.org, www.naacpartners.org, and education.uncc.edu/access/. The project's research-to-practice "learning community" management model taps the synergy of the organizational partners' research-based expertise, the policy and practice wisdom of the state partners, and external critical advice and oversight by national experts in educational measurement, inclusive assessments, special education, curriculum, PD, and universal design.

Relevance to the Assistant Secretary's Priorities and the Race to the Top Common State Assessment Program (RTTA). This project addresses all of the Assistant Secretary's priorities by 1) developing alternate academic achievement standards aligned with the CCSS; 2) developing a high-quality alternate assessment system using design-based principles; 3) building a reporting mechanism that supports appropriate interpretation and use of data from the system; and 4) creating clear and appropriate guidelines for IEP teams to use in participation decision

making, as well as PD to support full implementation of those guidelines. Like the RTTA consortia, NCSC will develop a system of assessments that are valid, reliable, and fair and measure student knowledge and skills against a common set of college- and career-ready standards in English/language arts (ELA) and mathematics in grades 3 through 8 and at least once in high school. We will work closely with the funded RTTA consortia to develop alternate assessment systems that fit coherently in states' comprehensive assessment systems. We also intend to work with other funded GSEG projects, supporting our belief that the knowledge base on AA-AAS is still in development, and that having several well designed initiatives with different designs and strategies is in the interest of the students who can benefit from participating in well-designed AA-AAS systems of assessment. See Appendix C for a statement of work that we have proposed to other GSEG applicants, in collaboration with the funder's cooperative agreements and leadership.

Issues and opportunities for collaborative partners. Articulating intended outcomes of the public school system for all students, including SSCD, is a necessary step in building appropriate system accountability. However, articulating desired outcomes and building assessments based on these outcomes is not sufficient for realizing these outcomes for all students. Given the relatively young fields of academic curriculum, instruction, and assessment for SSCD, we will grapple with, study, and build collective understanding of each of the following key issues:

1. College- and career-readiness. The Race to the Top initiatives and Common Core State Standards have focused on the goal of college- and career-readiness (CCR) for all students by the end of high school. We need ambitious yet attainable goals to ensure that all students have opportunities to continue to learn, grow, and contribute to society throughout their lifetimes. Our project will contribute to definitions of what CCR means for SSCD, building understanding of

what expectations for these students can be (Bowman & Weisenkauf, 2004; Conley, 2007; Kearns, in press; Wolfensberger, 1998).

2. Learning progressions. Cognitive psychologists have been working with content experts and educators to map out learning progressions for the general population. There is evidence that some cognitive processes differ for low-achieving students with disabilities, e.g., deficits in working memory (Pellegrino, in press). However, work on learning progressions is very new for SSCD, and current practices tend to focus on a few, repetitive skills across years. New research is needed to "map" how SSCD build competence in an academic domain. Researchers now are finding strong evidence of academic skill and knowledge development among students who participate in AA-AAS, including abstract concepts and transfer of learning (Jameson, McDonnell, Johnson, Riesen, & Polychronis, 2007; Jimenez, Browder, & Courtade, 2008). Research remains to be conducted on teaching academics to students who participate in AA-AAS, but anecdotal and case study evidence indicates that we must retain a "principle of uncertainty" about what students may be able to do (Kliewer & Biklen, 2001). As long as we are uncertain what many SSCD are capable of, we must operate on the principle of the "least dangerous assumption" (Donnellan, 1984; Jorgensen, 2005). Our project will develop and validate learning progressions with SSCD, linked to the CCSS, including skill sequences that can be used in instruction and assessment.

3. Formative and interim assessment tools. In addition to developing a summative system of alternate assessments that can be coordinated with the summative system for the general population, it will be important to parallel the work of the RTTA in developing formative and interim tools for teachers to use throughout the year. This work will be dependent on setting clear academic goals that can be guided by learning progressions. Our project will develop

formative assessment and progress monitoring tools essential to ensure that SSCD benefit from all components of the overall assessment system.

4. Instruction and curriculum tools. Classroom teachers are the key factor in student achievement (Ball & Cohen; 1999, Borko, 2004; Cohen & Hill, 2000) and are an important part of the assessment system. However, many teachers of SSCD need guidance to know how to translate academic content standards into instruction and assessments. Additionally, educators may not understand what concept or skill to teach next once a student masters an objective. Finally, teachers are challenged by the fact that SSCD are a heterogeneous group with varying entry skills into the content standards and may not know how to differentiate instruction for this population (Karvonen, Wakeman, Flowers, & Browder, 2007). In addition to the assessment tools, our project will develop core content modules and entry point and skill sequences that are aligned to the developed learning progression frameworks, to support teacher understanding of what to teach and how to teach it. These practices will be based on current research-based methods for teaching literacy and mathematics to students with intellectual disabilities (e.g., Browder, Trela, & Jimenez, 2007; Jimenez et al., 2008).

5. Differences within the "1%" population. A NAAC study of learner characteristics of students participating in AA-AAS (Kearns, Towles-Reeves, Kleinert, Kleinert, & Thomas, in press) found the population to be highly varied in expressive and receptive language, with 70% communicating symbolically using oral speech or symbol-based augmentative communication, 15-20% emerging in level of symbolic language and exhibiting consistent receptive responses, 8-12% communicating expressively at a presymbolic level and alert to sensory information, and 5-8% having inconsistent receptive responses. Approximately 70% of the population read basic sight words or read with basic understanding of text. About the same percentage of students

counted with correspondence or used a calculator to solve basic math problems. We will analyze how these data translate to instructional and assessment strategies and how to ensure that more students develop symbolic language as well as move beyond calculator skills in math and sight word recognition in reading. Careful consideration of how and whether goals, learning progressions, and growth targets differ for different groups of students within this population will be part of this discussion. Our project will contribute to better understanding of the learning characteristics of all students, and how to translate that information into an assessment system that minimizes barriers to participation.

6. Balance of flexibility and standardization in test administration. Because of the varying needs of SSCD, current AA-AAS typically require teacher expertise to either design or adapt the tasks for each student's academic level and mode of communication. This flexibility, while intended as an advantage to allow teachers to tailor the assessment to an individual student, may result in poorly designed tasks that do not accurately measure student achievement. Conversely, attempts to highly standardize test administration have often resulted in the inability to determine what some students truly know and can do (see Gong & Marion, 2006). Our project will identify promising new testing formats that find the right balance between flexibility and standardization, working closely with teachers, content, and assessment design experts.

7. Growth. Once we have clearly articulated the goals for SSCD and have a better understanding of an appropriate scope and learning continuum toward meeting those goals, we will need to measure student progress toward reaching those goals. Given the flexibility of many of the AA-AAS systems, along with the small sample sizes and high variability in the students, it has been difficult to scale results in a meaningful way to allow for traditional measures of growth. We will need to first describe growth qualitatively, and then our project will develop methods to

accurately measure growth quantitatively for these students. We will collaborate with the proposed IES Growth Center on this goal.

8. Use of technology. Technology can be used to deliver the "material" of both instruction and assessment, possibly reducing some costs and providing greater and easier accessibility. Our approach to PD will rely heavily on technology, on line and on demand, accessible to all at low cost. The use of technology has many potential benefits, such as providing multiple avenues for students to respond (including eye gaze) that will need to be weighed against research and practice on the challenges of removing construct-irrelevant variance in technology-based instruction and assessments. We want the assessments to reinforce powerful learning and teaching practices, and the use of assistive technology and communications tools are essential to the needs of SSCD. The project will develop more valid and more feasible assessments along with coordinated instructional and PD materials and processes that build on the potential of technology, compatible with RTTA and other initiatives.

9. Comparability. Policymakers have been calling for greater comparability of student-level results across states to provide richer information to help improve teaching and learning. Our project will leverage the opportunity of states working together to develop common learning targets, common achievement standards, and a common AA-AAS to get a bigger and better picture of the relative achievement and growth of SSCD across the nation.

10. Costs of AA-AAS. The cost per pupil of administering an AA-AAS has been reported at $250, $375, and up to $575 per student (Willhoft, Kraft, & Taylor, 2008). An informal poll by some of our state partners showed an even greater range of costs, up to $2500 per student in one small state. We believe our research-to-practice approach and development of a sustainable infrastructure for delivering the operational test can substantially contain costs currently incurred

by any one state once the system is developed, especially in states with very few students being tested. Our project will conduct cost and resource analyses on design choices, and work toward a fully sustainable system that ultimately will result in a cost savings for each partner state.

Importance to the field and to the students. AA-AAS, as well as focused academic instruction for SSCD, reflects a shift in teaching and learning practices for these students and their teachers. As we build understanding in the field of how to design meaningful instructional opportunities in the grade-level academic curriculum, we will find new ways of capturing what these students know and can do on formative and summative assessments for multiple purposes. We anticipate that they will continue to surprise us with their capacity to learn content that, in the past, they were never taught. As we carefully consider the consequences of the shift to increased access to challenging academic curricula, in partnership with SEA and LEA practitioners, we will discern the challenges of a paradigm shift for the professionals and IEP team members, including parents and guardians, and address these challenges to develop capacity of the state, district, school, teacher, and parents to ensure positive consequences and improved outcomes for all students.

Importance to the states and nation. This project is designed to benefit participating states by yielding the first fully coordinated system of formative and summative assessments with curriculum, instruction, and PD supports for improving achievement and outcomes for SSCD. The partner states will guide and shape the work of the research/expert partners in building this system, contributing to the effective transfer of research to practice, resulting in a model demonstration of innovative best practices. When our work is complete, the system will be available to all states, regardless of their participation in the initial development funded by this grant. The ultimate purpose is to provide states with a full set of assessments, curriculum and instructional supports, and PD tools to implement a research-based, systemic approach to

improving academic outcomes for SSCD. Thus, we are not only building a high quality, innovative system of assessments for the 19 states participating in our project; we are building comprehensive model demonstration of research-based best practice in curriculum, instruction, PD, and assessment for national benefit. We are doing so with a budget that is about one million dollars less per participating state than the cost guidelines per state defined in the priority.

Chapter 2 – Quality of Project Design

The issues and opportunities defined by our collaborative partners and described in the significance section inform our project design. We have learned that the development of technically defensible assessments is only one piece of the educational puzzle that results in improved outcomes for SSCD. First and foremost, we need clear content learning targets and curriculum designed around those targets, and the same clear content learning targets for the system of assessments. Then, we need to support teachers with model curriculum and instructional materials, and intensive PD. Finally, throughout the process, ongoing validity evaluations are needed to determine that the curriculum, instruction, and assessment strategies are effective, implemented as intended, and producing desired results.

Theory of Action. To guide the project design, and to identify short and long term outcomes, we have developed an initial Theory of Action (ToA, see Appendix D for graphic). The ToA displays the claims we will make about our system and is divided into components that are the focus of our work—curriculum, instruction, and assessment—that lead us to short-term assessment outcomes and long-term student outcomes. We focus first on our long term goals: to ensure that SSCD achieve increasingly higher academic outcomes and that they leave high school ready for postsecondary options. To that end, the first task we will undertake is to engage the workgroups in defining college- and career-readiness for SSCD. That information, along with

the learning progressions framework being developed by Workgroup (WG) 2, will serve as the foundation for a strong curricular framework to support a research-based assessment design. Our ToA starts with the assumption that there are pre-requisites to delivering quality instruction to SSCD, including high quality instructional materials, resources, and communication devices available to teachers and students (see first column, Appendix D). Next, in order for the assessment system to produce valid and useful results, students must be identified appropriately to participate in the AA-AAS, their instruction must be aligned with content assessed, and the AA-AAS must be administered as intended (second column). Several assumptions underlie each of these claims, closely related to the curriculum and instruction goals of WG 2 and the teacher/IEP team training and PD goals of WG 3.

Having made these claims, we must attend to the design of the AA-AAS (third column). The assessment design must match the learning characteristics and communication needs of the students who participate, while maintaining appropriately rigorous academic content. The claims about the AA-AAS include the technical defensibility of content and construct validity, and results that reflect student knowledge and skill. Student scores should provide useful information for teachers to inform instruction, and for parents to track their progress (short-term outcomes column). The assessment system is anchored in the long-term goals that SSCD should achieve increasingly higher academic outcomes, and leave high school ready for postsecondary options (long-term outcomes column). We will revise the ToA with our state and expert partners in the first few months of the project. It will serve as the foundation for the validity evaluation work (see WG 4).

Project Workgroup/Task Areas and Goals. Our project design is organized by four major workgroup (WG)/task areas, with corresponding goals, objectives, and timelines for each: WG 1.

Assessment design and development; WG 2. Curriculum and instruction supports; WG 3. PD and training; and WG 4. Validity evaluation. These WGs interact throughout the project; WG goals, objectives, and activities are iteratively linked. In addition, a set of four management tasks encompass all areas (see management section). See Appendix E for a summary of WG/task areas, organized by goals, objectives/activities, and timelines.

WG design and process. Each WG consists of research staff and state practitioners with high interest, skills, and experience in the WG areas. Our workgroups are modeled on Community of Practice literature (e.g., Wenger, McDermott, & Snyder, 2002), with staff facilitation and direction and shared leadership among the state members, and coordination by the management team. This model leverages the knowledge and skills of the partners by structuring common learning to generate powerful innovations in current practices. They will use a common process:

1) First, each group will map out existing knowledge that defines the current challenges. This may include convening of ad hoc committees of experts and practitioners for specific work.

2) Next, each WG will investigate alternative options to improve practice, and develop methods to explore and test feasibility. Our project staff will carry out this work with guidance from state partners on each workgroup, tapping into our expert advisors. In some cases, they will refer issues to the Technical Advisory Committee (TAC) for review and advice.

3) Finally, each WG will select alternatives through a process of collegial weighing of evidence, and, if necessary, negotiations on key differences, facilitated by the management team.

The WGs will be coordinated with one another through two mechanisms. First, the monthly meetings of the management team will have staff and state representatives from each WG to identify overlaps or gaps, discuss issues with timeline or products, and decide where new or different activities are required. Second, the project evaluation team will monitor the WGs to

identify opportunities and barriers, help resolve them, or bring them to the management team. The preliminary objectives and activities of each WG are described below. See Appendix E for more detailed information and timelines. Given the iterative, developmental nature of our project, these are subject to revision as the work unfolds. These reflect our current project design to ensure we attain all project goals and desired outcomes.

Workgroup 1: Assessment Design and Development. Our summative assessment framework builds on the Assessment Triangle, first presented in Pellegrino, Chudowsky, and Glaser (2001), then modified under the New Hampshire Enhanced Assessment Initiative (NHEAI), and ultimately documented in Marion and Pellegrino (2006, see Appendix F). This triangle explicates the key interactions between models of student cognition, observations of student work, and the inferences we draw from these observations about what students know and can do. Pellegrino et al. (2001) suggest that every assessment must rest upon three foundations: "a model of how students represent knowledge and develop competence in the subject domain, tasks or situations that allow one to observe students' performance, and an interpretation method for drawing inferences from the performance evidence thus obtained" (p. 2). The conceptual elements of each vertex must be thoroughly defined, each vertex must complement the other, and an argument for how these vertices fit together into a coherent whole must be made for the inferences of the assessment to truly have meaning. All of our assessment design and development goals are supported by the latest, most rigorous research-based theories pertaining to what SSCD know and can learn, how best to measure their knowledge and skills, and how to interpret their assessment outcomes.

Goal 1.1: Define college- and career-readiness (CCR) (Year 1): Our assessment system is anchored by the goal that SSCD will achieve increasingly higher academic outcomes and leave

high school ready for post-secondary life. Defining CCR for these students requires exploring what successful, meaningful, varied, and sustainable outcomes may look like for SSCD (Kearns, in press). During Year 1, we will work with special educators, school transition experts, and content experts to create a vision for CCR that will serve as a basis for our assessment system.

Goal 1.2: Establish construct definitions (Year 1): Our assessment design will be based on research about how SSCD actually progress in academic content, rather than selecting arbitrary sequences of skills aligned to the CCSS. The work done at UNCC by WG 2 during the first year of this project will provide solid evidence about what a diverse group of SSCD are capable of achieving in ELA and mathematics and will yield validated learning pathways that lead to student academic proficiencies across the grades. The learning progressions frameworks (LPFs) and grade-level targets being developed by WG 2 will inform the cognition vertex, along with our understanding of SSCD from the research including the work done on the Learner Characteristics Inventory (LCI) described above. Experts in student cognition will advise us throughout the development process. This work will ensure that challenging academic achievement standards are adopted for SSCD.

Goal 1.3: Design assessment frameworks (Years 1–2): We envision an assessment system comprised of a bank of items and a minimum of two blueprints for each subject and grade to allow states flexibility in selecting an approach that best matches their philosophy about assessing SSCD, or that best matches their full assessment system. The item bank will be linked to the CCSS, coded with standards assessed and level of complexity. We will develop multiple items for each standard at different levels of complexity, ranging from entry points into the standard through a full compression of the standard.

We will use evidence-centered assessment design (ECD) to determine the appropriate item type for each standard and to develop sets of items for each content target. ECD is a process of assessment design that involves gathering, organizing, and transforming information in a variety of representational forms, within the framework of a clearly articulated assessment argument. Messick (1994, p. 16) concisely lays out the key aspects of an assessment argument by asking "what complex of knowledge, skills, or other attributes should be assessed? Next, what behaviors or performances should reveal those constructs, and what tasks or situations should elicit those behaviors?" All of the terms, concepts, representations, and structures in ECD are aimed at constructing a coherent assessment argument and building machinery to implement it. "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" (Cameto, Haertel, Morrison, & Russell, 2010, p. 1; Cameto and Haertel are expert advisors on this work, see Personnel Section).

We will continue to use LCI data from participating states to guide universal design considerations, test design choices, participation and accommodations policies and guidelines, and monitoring strategies on implementation of guidelines, focusing particularly on characteristics involving communication, sensory, and motor needs. Our design will include multiple ways for students to access the material and to demonstrate their knowledge along an articulated learning continuum. Our PD will support implementation of best practices for curriculum and instruction, and focus on strengthening students' communication skills. As students' receptive and expressive language skills increase, we hypothesize that they will have greater access to the curriculum and greater achievement. Our understanding about student characteristics and universal design principles will help us design each item, and any necessary

additional adaptations and accommodations that do not interfere with the construct being measured. Our goal for the AA-AAS is to balance the need for standardization with the need for flexibility (Gong & Marion, 2006). We also anticipate changes to the proposed design shared here based on our WG process and findings.

We will collaborate with funded RTTA consortia to ensure complementary components of the assessment system. These decisions will be made during the first year of the grant, working with all internal and external partners and stakeholders. Bid processes will be designed from these decisions, and vendors will carry out components of the design.

Goal 1.4: Develop/field-test assessment items and draft performance level descriptors (PLDs) (Years 2–4): Beginning with a framework of item type by content standard, we will build our item bank for students to be assessed in grades 3 through 8 and once in high school in ELA and mathematics. Items that require a more complex scoring approach than "right/wrong" will also include a scoring rubric meant to fully decode student understanding. We will design an accommodations manual and an administration manual to accompany this work.

Each blueprint (to be developed in coordination with WG 2 under Goal 2.6) will include multiple item types, such as multiple choice, short constructed response, and performance tasks. The blueprints will be specific in their description of how many items teachers must select from each standard, but flexible in which items they select. The goal is to reduce the burden on teachers for developing fully-aligned items or tasks at the appropriate level of complexity, but to provide enough flexibility for teachers in selecting and implementing the task that is appropriate for each student. This approach raises a concern that teachers may select the easiest item to ensure a high performance score. We will address this concern in our workgroups, but early suggestions raised include requiring teachers to test third-graders in the fall to establish a

baseline measure of what constitutes appropriate challenge for each student; providing a screener to help teachers select the most appropriate items; or weighting complexity over performance in the final score to incentivize teachers to select more challenging items.

Items will be subject to multiple analyses. First, the entire item bank will undergo an alignment study, and each item will undergo field testing. Then, items will be analyzed for difficulty and discrimination as well as statistical bias. One benefit of the large number of states participating in this project is that we will have a larger sample of students on which to conduct the field test using multiple blueprints. This larger sample will allow us to run statistics we normally cannot, such as differential item functioning. We can analyze responses both by demographic information and by profiles of communication, auditory, vision, and motor skills.

During the field-test stage, each response will also be coded by independent scorers, and a study will be conducted to determine the reliability and accuracy of teacher scores. Depending on the results, either the central scoring will continue for each student or will act as a read-behind for a randomly selected group of students as a reliability check. Final scores will consist of the performance score as determined by the teachers and independent scorers and the complexity scores pre-coded for each item.

Early in development, we will develop draft performance level descriptors (PLDs) that will be used for all blueprints to ensure that items are developed to clearly distinguish among the performance levels. We will start with the concept of CCR for the high school PLDs, and ensure alignment with those concepts in the earlier grades. After the field test, we will set cut scores separately for each blueprint. The percent in each performance category can be aggregated across students and integrated into a state accountability system to be used as desired by the state. See Goal 1.6 for reporting system.

Goal 1.5: Establish technology platform (Years 2–4): Technology can contribute to improved assessment system design in many ways. We will carefully explore ways to implement technology into every aspect of the test design, while avoiding the introduction of construct-irrelevant variables. We envision that teachers will use technology to design and administer the assessment by selecting items according to the blueprint at an appropriate level for each student, and entering or uploading the student's score. We will conduct a usability study of the administration system to be sure that teachers find the assessment system, including assistive and test administration technologies, easy to understand and administer. Once the WG determines key features for technology implementation, from student and teacher interface to administration and reporting infrastructures, a bid process will be designed and vendor(s) secured.

Goal 1.6: Develop reporting system (Years 2–4): We will develop a reporting system that allows scores and interpretive information to be disseminated electronically. Teachers will be able to view both individual score reports and classroom reports. Principals will see school reports, and districts will have district reports that can be used for program evaluation. Finally, each state will receive a state report that allows them to compare the percentage of students at each performance level to those in other states that use the same blueprint. Our goal will be to develop reports that are user-friendly and actionable. The reporting system will be analyzed by focus groups to determine its usability. Finally, we will produce both teacher and parent guides to help them interpret reports and determine next steps, as well as develop training on appropriate use of score reports delivered to educators**.**

Goal 1.7: Establish operational assessment system (Year 4): In order to complete the development process to match the RTTA timeline for full operational testing, we will conduct a census field test in the final year of the project. We will finalize all items, blueprints, and

reporting system procedures and products; produce final administration, participation/accommodations, and technical manuals; develop final PLDs; and set final cut scores. The C & I and PD processes developed by WGs 2 and 3 will be fully implemented to carry out support within each state by the end of the project, including the teacher networks (COPs, see WG 3), ensuring capacity for high quality implementation after the project's end.

Workgroup 2: Curriculum and Instruction Supports. The most recent and innovative approach that has emerged for developing assessment targets for AA-AAS is to create content-specific learning progressions. Learning progressions (LPs) are descriptive continuums of how students develop and demonstrate more sophisticated understanding over time (Hess, 2010). Learning progressions, also referred to as learning continua or progress maps, have become an increasingly important tool in today's general education classrooms for tracking progress and using formative assessment data to inform instruction (Hess, Burdge, & Clayton, in press). A few states have begun to use a deliberate development process to create content-specific learning progressions for their AA-AAS, while other states discovered after developing curriculum frameworks with varying levels of complexity that they had developed a type of learning continuum (Wiener, 2005). NCSC is using deliberate and research-based processes to develop and validate a learning progressions framework (LPF) as the foundation for developing the AA-AAS. With support from NAAC, NCIEA designed and facilitated a process during spring 2010 for developing LPFs in reading, writing, and mathematics aligned to the CCSS. National content experts and special education experts were convened to: (1) identify the big ideas/enduring understandings (Wiggins & McTighe, 2001) of each content area; (2) describe broad learning targets for elementary, middle, and high school that would illustrate how a student would show an understanding of the bigger idea/enduring understanding along a continuum; and (3) break

down learning targets into an instructional focus at smaller grade spans. These descriptions represent the "chunks" of learning that happen in sequence by grade or grade-span, which begin to form the LPFs across grade levels for each big idea (see example, Appendix G). In July 2010, content experts, classroom teachers, and special education teachers will use the LPFs to create curricular modules to support achievement of the learning targets at each smaller grade span. We will build on these modules.

Specifying the type of instruction is a necessary, essential feature for interpreting and using the LPs. Given the lack of research on academic instructional methods for SSCD, we will develop and implement instructional methods and skill sequences that support the LPFs. Results gathered from the classrooms will provide validity evidence about the pathways described. The project team will contribute their extensive expertise in previous development of LPs as well as experimental research on teaching academic content to SSCD and alignment of AA-AAS to address the challenge of translating the theoretical constructs (i.e., LPs), general education unit plans, and related research into target skills that SSCD can learn. These skills will inform the development of specific AA-AAS items.

Goal 2.1: Validating the LPFs and entry points in ELA and math (Years 1-2): Several strategies will be used to validate our LPFs: cognitive labs, analysis of student work samples, and teacher interviews and surveys collected during year 1. Additionally, we plan to conduct a second review of the draft LPFs using another group of national content experts as part of the validation process. The validation process will serve several key purposes, including identifying and addressing potential gaps in the LPFs; documenting effective instructional approaches to teaching academic content to SSCD; validating progressions of learning when best practices in

instruction are employed; and identifying potential item types for assessments and progress monitoring. These activities will make a major contribution to the field.

Following the validation of the LPFs, WG 2 will examine prioritization of the content within the LPs using models from literature on educating SSCD (e.g., Browder et al., 2009) and recommendations from the expert advisors and state partners. Then, the team will develop entry points into each of the identified progressions to allow maximum accessibility for all students. Once entry points are developed, we will validate them with up to 20 teachers and 4 content experts, including the degree to which the entry points represent the content and how useful the entry points are for instruction. For this review, the team will use a talk aloud procedure with individuals and/or focus groups to gather data on accessibility of the LPs for students with different needs, such as receptive and expressive communication levels. After entry points have been established, the draft LPs will be submitted to at least 4 content experts to rate the extent to which each progression aligns to the CCSS using rating scales within the Links for Academic Learning (Flowers, Wakeman, Browder, & Karvonen, 2009). The LPs will be revised once more and again submitted to the WGs, experts, and TAC (including all state partners) for review and consensus building. This validated content will then be available for the assessment design team for use in the AA-AAS.

Goal 2.2: Develop skill sequences within each learning progression (Years 2–3): WG 2 will create universally designed units of instruction, informed by research from Spooner, Baker, Harris, Alhgrim-Delzell, and Browder (2007) and Browder, Spooner, Wakeman, Trela, and Baker (2006). Based on ideas on increasing participation and learning opportunities through the use of representation, expression, and engagement by SSCD, WG 2 will create skill sequences linked to specific standards and units of study within selected LPs. For validation, up to 80

teachers will be trained on site to teach these sequences using evidence-based practices for instruction of SSCD including systematic instruction (Collins, 2007), self-determined learning (Wehmeyer, Field, Doran, Jones, & Mason, 2004) and embedded instruction (Wakeman, Browder, Jimenez, & Mims, in press). Teachers will submit student data through a dedicated website using DVDs and permanent products. Teachers will be given a survey about how the students learned the skills, followed by a conference call for teacher think-aloud interviews. These data will be used to refine the skill sequences. Final skill sequence and data will be shared with the larger team for use in the AA-AAS. The sequences are refinements of LPs developed in Year 1 based on an iterative process using field testing with students.

Goal 2.3: Develop generalizable skill sequences for each content area (Year 4): WG 2 will use advances from the research literature to develop sequences for generalized skills that cross strands of the LPs and may be pivotal to learning across a larger scope of general education content. These sequences will have more of a process focus than those sequences developed earlier. For example, a sequence may focus on the process for deciding what tools are needed to solve a math problem. Research from the self-determined model of learning will help to guide this development (Wehmeyer et al., 2004). The sequences will be taught to up to 20 teachers in partner states using on-site training. Teachers will submit data and DVDs across the school year to show evidence of student learning via the dedicated website. These data will also provide a model for progress monitoring.

Goal 2.4: Piloting and validating formative and interim uses of assessments that support progress monitoring (Year 2): WG 2 will gather data (including student work analyses and teacher focus groups) to design formative assessments and progress monitoring tools. These activities will be supported by our PD activities to ensure high quality implementation and

fidelity to the content. Formative assessments and progress monitoring tools will then be revised as needed and re-piloted by teachers in year 3.

Goal 2.5: Develop content support for special education teachers (Year 2): The team will collaborate with ELA and math content experts, as well as with WG 3, to develop on-line modules that support special education teacher understanding of the core content reflected in the LPs. The module content will be aligned to the prioritized content within the LPs. The modules will be reviewed by up to 2 content experts for their accuracy and effectiveness. The team will pilot the modules with up to 80 teachers to field test their usefulness in supporting teacher understanding of the content. State partners will be invited to participate in on-line trainings. Teachers who participate will be asked to provide evidence of their content understanding by drafting alternate assessment or lesson plan/formative assessment entries/tasks/items. These will be rated by the research team for alignment with the designated content standards and LP.

Goal 2.6: Develop test blueprint aligned to the CCSS (Year 2): WG 2 will work collaboratively with WG 1 (under Goal 1.4) to develop a thoughtful test blueprint that reflects work being done on LPFs and skill sequences aligned to the CCSS. Based on Year 1 data from participating teachers, a deliberate process will be used to establish a distribution of emphasis, an appropriate range of cognitive complexity, and categorical concurrence (major strands assessed at a given grade level for each content area) for the assessment. Content experts will work with WG 1 to develop and refine PLDs for the AA-AAS using teacher feedback and student work samples collected during years 1–3. These data will also be used to refine curricular modules to reflect areas assessed with the summative assessment, create consistent scoring criteria, and design more effective uses of the ongoing progress monitoring tools (see Goal 1.4, above).

Workgroup 3: PD and Training. SSCD must have access to high quality curriculum and instruction as an essential pre-requisite for assessment validity. WG 3 will collaborate on the development and deployment of PD materials to support cross-WG products and processes. We anticipate that the training materials will include the following: 1) assessment administration for both the interim and summative assessments; 2) interpretation of assessment results; 3) selection and use of assessment accommodations and other administration choices; 4) communication triage for students not using symbolic language in grades K-8; and 5) curriculum based on the CCSS and our LPFs. In addition, we will develop protocols for use in evaluation of teacher and principal effectiveness, determining PD needs, and improving practice.

To scale-up the use of academic curriculum based on the CCSS and assist teachers with developing communicative competence, we will develop teacher-based communities of practice (COPs) (Barber & Fullan, 2005) in each state, led by project staff. The PD component and the teacher evaluation designs will be guided by the findings of the National Commission on Teaching and America's Future, on learning-team principles that are the most effective for improving teaching and student achievement (Carroll, Fulton, & Doerr, 2010). The COPs will support teachers by: 1) providing a network of support in each state to build the curriculum, instruction, and assessment infrastructure through training and technical assistance; 2) providing model curriculum resources for teachers serving SSCD and supports for developing communicative competence; 3) assembling stakeholders across the participating states willing to participate in project activities like cognitive labs and small-scale tryouts, item reviews, field tests, standard setting, and validity evaluation activities.

Goal 3.1: Establish communities of practice and introduce CCSS (Years 1-2): In year 1, we will invite 10–40 teachers in each state (depending on size) to participate in COPs that will

help implement model curricula. COPs will meet for Summer Institutes to receive training on the CCSS, the relationship between content and achievement standards, curriculum, assessment, and access to the general curriculum. The Institutes will occur within large states; smaller states with close geographic proximity may collaborate in a central location. Following the Summer Institutes, we anticipate meeting with the COPs for 1.5–2 hour meetings, 5-6 times per year in the first 2 years, for coaching and follow-up. These may be face-to-face or distance meetings, depending on state preferences. We anticipate meetings will occur more frequently in early years, as we fully integrate the new system into state practices using technology-based training.

We will provide external staff to serve as "critical friends" to each state in a ratio of 1 staff to approximately 3 states (based on size) to establish and mentor the COPs. It may also be necessary to assist states to fund staff through budget allocation, determined annually with the Management Team, based on ongoing needs assessment, state transition plans, and changing economic and political realities in each state. We will provide stipends for after school meetings, and reimburse travel expenses and the cost of substitutes. In addition, we will support the delivery of on-line, web-based training and all materials.

Goal 3.2: Implement CCSS (Years 3–4): We will work closely with WG 2 to support the development of curricula based on the CCSS in ELA and math. The COPs will meet at the Summer Institutes to learn about new curriculum applications and research findings, and for project staff to learn from the COP members what is working and where more assistance is needed within states. We will continue to develop curriculum training materials, including lesson plans with appropriate supports. We will host Curriculum Summits with the COPs from all the member states during Years 2 and 4 of the project to introduce the evidence base produced by WG 2 and assist teachers in designing new curriculum materials based on the CCSS.

Goal 3.3: Communication triage K- 8 (Years 1-4): As discussed in the significance section, approximately 30% of the students participating in AA-AAS do not use symbolic communication, presenting a significant challenge for collecting reliable, valid data on their achievement. However, researchers suggest that symbolic communication can be developed for nearly all students with the right instruction and supports (National Joint Committee for the Communicative Needs of Persons with Severe Disabilities, 1992). We will develop materials and provide training to teachers to properly identify students in grades K-8 who would benefit from language development instruction and appropriate augmentative communication strategies. The goal of these materials and training is to ensure that each student has an expressive and receptive response repertoire, prior to the first administration of the new assessment, so that their assessment results will be valid indications of their knowledge and skills. During years 2–4, WG 3 will review the communication items from the LCI for grades K–8 to monitor changes in students' communication competence and adjust the training modules as needed. We will train the teacher COPs in communication triage and will focus at least three coaching calls on strategies for building communicative competence. We will host a Communication Summit in Years 2 and 4 to establish triage teams of special education teachers and speech language therapists to coach teachers in developing communicative competence. We have identified existing resource experts in our participating states to engage in this work.

Goal 3.4: Develop assessment administration training (Years 2–3): WG 3 lead staff, from UKY, has extensive experience in designing training and qualifying examinations for the administration of AA-AAS. We propose using electronic training for this purpose to control the consistency of the training and provide extended opportunities for teachers to engage in the training. In addition, we can ensure that the assessment administrators meet a minimum

competency by providing a qualifying examination. We will work closely with WG 1 to develop the most succinct and direct training necessary to achieve reliable procedural fidelity for administering the assessment. In addition**,** we will work with WG 1 to identify appropriate accommodations for students who may have hearing, vision, motor, or other barriers that may interfere with assessment responses. Our accommodations training module will assist teams in making sound participation decisions.

Goal 3.5: Assessment results interpretation training (Years 3-4): WG 3 will develop training modules for teachers and principals for appropriately interpreting both interim and summative alternate assessment results in order to make defensible instructional decisions at the school level. Teacher survey data suggest that AA-AAS results are rarely used to inform program improvement (Flowers, Ahlgrim-Delzell, Browder, & Spooner, 2005). While a number of reasons could account for this, specific training opportunities would improve the use of these data for instructional decision making. We propose to work closely with WG 1 and each state team in the design and development of these trainings.

Goal 3.6: Develop teacher/principal evaluation tools (Years 2-4): Little is known about evaluating teacher and principal effectiveness related to this population of students. Principals regularly report a "lack of understanding" of what teachers serving this population should know and do (Towles-Reeves, Kleinert, & Anderman, 2008). Similarly, teachers report that principals have low expectations of achievement from these students and a lack of resources for effectively providing access to the general curriculum (Flowers et al., 2005). We will develop, field test, and validate a multiple-measures approach to evaluating teacher and principal effectiveness. These may include but are not limited to: 1) a teacher self-assessment; 2) school/program assessment; 3) principal walk-through tools; and 4) coaching processes. During years 2 and 3, we will

convene an expert panel group to assist us in the design and development of these instruments. We will field test with COP members, and prepare supporting PD materials.

Workgroup 4: Validity Evaluation. As noted above, this project will employ an argument-based approach to validity evaluation (Kane, 2006). The evaluation process will begin immediately upon the award of the grant with the refinement of the draft ToA (see Appendix D) and development of the Interpretive Argument (IA) for the AA-AAS and continue through the design, implementation, and synthesis of studies related to specific claims and assumptions defined in this Interpretive Argument.

Goal 4.1: Establish the Theory of Action (ToA) and Interpretive Argument (IA) (Year 1): The first step in the validity evaluation process is to articulate the general ToA, or the logic that underlies the design of the assessment and the use of its scores, and to create the IA that lays out the claims and assumptions inherent to the ToA. The process for developing the ToA and IA will be an iterative series of discussions with state representatives and other project participants from each WG, including the TAC. At the first all-project meeting, WG 4 leaders will facilitate a discussion of the major goals and purposes for the AA-AAS and refine the ToA. A draft IA will be developed by WG leaders, and then be vetted by project participants and refined.

Goal 4.2: Prioritize issues and develop study designs (Year 1): Although some priorities for studies have been determined through the proposal process, states' priorities for validity studies will begin to emerge during the IA vetting process. Validity evaluation leaders will specifically vet these priorities and elicit others among state and WG participants once the IA is completed. They will identify and confirm priorities and designs for studies that address each of the five major aspects of validity:

- Test content - an alignment study will be conducted as part of the validity evaluation;

- Response processes – for example, cognitive laboratory studies to identify how students engage with and solve problems posed in assessment items;
- Reliability and accuracy of teacher scoring – for example, collecting and analyzing think-aloud data from teacher and expert scorers;
- Relationships with other variables – for example, comparisons of student achievement scores or ratings based on other data sources;
- Consequences of assessment processes and scores – for example, evaluations of classroom instruction to determine the degree to which academic content is being addressed.

Goal 4.3: Implement validity studies (Years 2–4): Over the course of the years 2-4 of the project, WG 4 will implement research studies designed to address the priorities identified in Year 1. As necessary, the evaluation plan may be modified during this period to ensure that it addresses emerging issues and remains aligned with expectations and innovations that may arise during assessment development.

Goal 4.4: Synthesize results and produce reports (Years 2–4): Throughout the evaluation process, information gained via specific studies will be integrated into an evolving validity argument (VA), the evidence-laden version of the IA. Two types of reports will be produced: project-focused reports, which are meant to inform project participants, the U.S. Department of Education, and others, and external reports, which are meant to inform others about the project and its outcomes. Project-focused reports will be produced for the project as a whole and for individual or groups of states as requested by states and project leadership. External reports include journal articles, book chapters, and articles for presentations at professional conferences. All reports will be written to support full access by a range of target audiences and will be available at no charge via the project website which will include relevant information and

documents in a format that meets a government or industry-recognized standard for accessibility. All products will be vetted by the validity evaluation team.

Chapter 3 –Project Personnel

Our project team includes accomplished scholars, researchers, writers and presenters; experienced and respected state leaders in assessment and special education; and expert partners with specific expertise related to the key issues identified in the significance section. The management team includes seven leading experts in AA-AAS, representing five nationally-respected organizations. **Dr. Martha Thurlow** will serve as overall Principal Investigator and convene the Technical Advisory Committee. **Ms. Rachel Quenemoen** will be the Project Director, convene the management team and full project meetings, and work with each state to develop and implement an individualized standards and assessment transition plan. The other lead project staff members of the management team are listed in Table 1 by workgroup/task leadership roles, including management tasks. **State leaders** representing each of the workgroups join the staff on the management team. Staff bio statements and FTEs follow the table.

Table 1. Leaders and Key Contributors for Each Workgroup/Task Area

Workgroup/Task	Leader	Key Contributors
Task 1 – Assmt Workgroup	Perie	Hess, Flowers, Kearns, NCIEA
Task 2 – C/I Supports Workgroup	Wakeman	Hess, Browder, Flowers, Kearns
Task 3 – PD/Training Workgroup	Kearns/UKY	Hess, Browder, Flowers, Wakeman

Task 4 – Evaluation Workgroup	Forte	Towles-Reeves
Task 5 – Convene TAC	Thurlow/NCIEA	Quenemoen
Task 6 –State Transition Planning	Quenemoen	Towles-Reeves, NCEO
Task 7 – Public Dissemination	Quenemoen	Towles-Reeves, NCEO
Task 8 – Project Management	Quenemoen	Towles-Reeves

Summary bios, roles, FTEs for project lead staff (also see CV Appendix). Dr. Martha Thurlow (.10 FTE) has been the principal investigator of more than 15 projects funded by the Office of Special Education Programs, the Institute of Education Sciences, and by various states, served on numerous national, state, and project panels, and published extensively on inclusive assessment and accountability systems. Her work focuses on how to obtain valid, reliable and comparable measures of student performance that assess the knowledge and skills of students with disabilities rather than their disabilities. **Rachel Quenemoen** (.65 FTE), has worked for 25 years as an educational sociologist focused on research to practice efforts and has been a multi-district cooperative administrator in both general and special education. For the past 15 years, she has worked on educational change processes related to standards-based reform and students with disabilities, building consensus and capacity among national, state, and local practitioners and policymakers. She has facilitated several multi-state projects, served on multiple project state, and federal panels, is co-author of a book on alternate assessments and has written and presented extensively on AA-AAS, standards-based IEPs, students with disabilities in standards-based reforms. *NCEO is the fiscal host for NCSC and leads the management team, TAC, and*

individual state transition planning. **Dr. Karin Hess** (.26 FTE), Senior Associate with NCIEA, has worked with more than a quarter of states in the nation developing grade level expectations, revising state content standards, and creating detailed assessment specifications aligned to educationally and technically-sound content standards. Dr. Hess is a recognized national leader in developing practical approaches for addressing cognitive rigor, as well as designing and using learning progressions as a foundation for formative, interim, and large-scale alternate assessments supporting improved achievement. **Dr. Marianne Perie** (.50 FTE), Senior Associate with NCIEA, has worked with multiple states on assessment and accountability systems. She has focused on alternate assessments, standard setting, validity studies, and technical documentation, serving on three state TACs. She has worked on the National Assessment of Educational Progress (NAEP) and international assessments. She recently wrote a book on setting cut scores published by Educational Testing Services. *NCIEA leads WG 1 and will provide content and assessment design expertise across WGs 2 and 3.* **Dr. Diane Browder** (.125 FTE) is the Lake and Edward P. Snyder Distinguished Professor of Special Education at UNCC. Researching effective instruction and assessment for SSCD for more than two decades, she focuses on AA-AAS and linking both assessment and instruction to the general curriculum. She has two IES-funded research projects, early literacy and math/science for SSCD. **Dr. Claudia Flowers** (.125 FTE) is a UNCC professor of educational research, measurement and statistics. She has several peer-reviewed publications focusing on research in academic achievement and assessment of students with developmental disabilities. She serves on numerous committees including six state TACs and additional expert panels advising on assessment of students with disabilities. **Dr. Shawnee Wakeman** (.25 FTE), a Clinical Assistant Professor in UNCC's Department of Special Education and Child Development, has contributed to more than

15 alignment studies of state AA-AAS. She is researching the principal's relationship to the education of students with disabilities, curriculum access, alignment of educational systems as well as policy issues related to alignment issues and AA-AAS. Her most recent work focuses on learning progressions in ELA and the alignment of content within units of study. *UNCC leads WG 2 while providing expertise to WGs 1 and 3.* **Dr. Jacqui Kearns** (.50 FTE) directs the NAAC and a GSEG project studying the validity of AA-AAS. She has worked on AA-AAS since the development of the first alternate assessments in the early 1990s, in Kentucky and Maryland, and has guided provision of professional development through teacher networks since that time. She has co-authored a book on AA-AAS as well as numerous chapters and peer-reviewed journal articles. *UKY leads WG 3 and provides expertise in severe disabilities, communication strategies, and learner characteristics to the other WGs.* **Dr. Ellen Forte** (.30 FTE) is the President and founder of edCount, LLC. She is nationally known for assisting state and local education agencies in their successful interpretation and implementation of education policies related to standards, assessments, and accountability. She sits on numerous technical advisory committees, and is conducting validity evaluations for four state assessment systems. She has worked as a large urban district assessment director, and with a state assessment office. **Dr. Elizabeth Towles-Reeves** (.75 FTE), served as project coordinator for the NAAC for 5 years, including coordinating a five-state NAAC GSEG consortium to test and then strengthen the validity of the participating states' AA-AAS systems. At edCount, LLC, she continues work studying the validity of AA-AAS, serves on one technical advisory committee, and has authored and co-authored more than a dozen research reports, most in peer-reviewed academic journals. *edCount, LLC leads WG 4 and provides formative and summative validity evaluation findings and feedback to the management team and each workgroup. Additionally, edCount will provide*

direct oversight to the external project evaluator to ensure the GPRA goals of quality, relevance, and usefulness are a foundation of the external evaluation and in required project reporting. edCount will host the vendor contracts for components of assessment implementation.

The NCSC leadership is augmented by other staff with expertise in inclusive assessment and accountability systems. An NCIEA team will be part of the assessment and validity workgroups and co-facilitate TAC/expert work with NCEO: C. Domaleski, B. Gong, S. Marion, P. Paek. An NCEO team will support extant data reports, document/disseminate workgroup findings, conduct cognitive labs, and lead universal design and accommodations work: J. Altman, L. Christensen, C. Johnstone, S. Lazarus, M. Moore, C. Rogers, Y. Wu. A UKY team will work directly with states on all PD products and processes: A. Cole, S. Kennedy, J. Kleinert, L. Land, J. Norman, M. Zeller. See the CV Section and budget justifications for roles.

NCSC Technical Advisors. The project Technical Advisory Committee (TAC) members have national stature in cutting edge theory and practice, and were chosen to challenge and refine our thinking (see CV Appendix). Thurlow, Quenemoen, and the NCIEA team will facilitate their work. **Dr. Derek Briggs** is Associate Professor of Quantitative Methods and Policy Analysis at the University of Colorado – Boulder and chairs the Research and Evaluation Methodology Program. His research on building sound methodological approaches for measurement and evaluation of growth in student understanding focuses on applying psychometric models to test hypotheses about learning progressions and exploring conceptual issues in how developmental test score scales are established and interpreted. **Dr. George Engelhard** is an Emory University Professor of Educational Measurement and Policy in the Division of Educational Studies. Co-editor of four books and author or co-author of over 100 journal articles, book chapters, and monographs, Professor Engelhard also serves on several technical advisory committees on

educational measurement policy in 8 states. He is an AERA fellow. **Dr. Michael Kolen** is Professor of Educational Measurement and Statistics at the University of Iowa. He is Past-President of the National Council on Measurement in Education (NCME), received an NCME Career Award in 2008, is a member of the Joint Committee for the Revision of the Standards for Educational and Psychological Testing, is a Fellow of AERA, a Fellow of APA, and serves on multiple state and organizational technical advisory committees. Dr. Kolen has numerous publications in journals and books on the subject of test scaling, linking, and equating. **Dr. Suzanne Lane** is a University of Pittsburgh professor in the Research Methodology program. Her research and professional interests are in educational measurement, testing, and policy, particularly technical, validity, and policy issues related to the use of large scale assessments. She has been President of NCME (2003 – 2004), Vice President of AERA Division D (2000 – 2002), a member of the AERA, APA and NCME Joint Committee for Revision of *Joint Standards* (1993 – 1999), a member of the Management Committee for the next revision of the *Standards* (2006 – 2009) and a member of the National Research Council Committee on Test Design for K-12 Science Achievement (2003 – 2006). **Dr. Jim Pellegrino** is Distinguished Professor of Education and Liberal Arts and Sciences Distinguished Professor at the University of Illinois at Chicago where he is also Co-Director of the Interdisciplinary Sciences Research Institute. He has chaired several National Academy of Science study committees, is a past member of the Board on Testing and Assessment of the National Research Council and is a lifetime member of the National Academy of Education. His research interests focus on thinking and learning as well as implications of cognitive research and theory for assessment and instructional practice, including specifically partnering with NAAC on how it applies to SSCD.

Expert Panel/Pool. Expert partners will assist us as we study the research base, consider options, and design, develop, and implement our comprehensive system. We will secure additional experts with specific expertise as the workgroups identify emerging needs, and to release RFPs for major test development and implementation components. Experts include: Aspel, Teacher/Principal Effectiveness; Burdge and Clayton, Teacher Networks; Cameto and Haertel, Evidence-Centered Design; Kozleski, Teacher Evaluation Expert Panel; Roeber, Topol, Olson, Cost Effectiveness Studies;Ryan, Evaluation; Wiener, Curriculum Frameworks, Teacher Networks; Warren, cross-state dissemination and PD. See bios in CV Section.

Commitment to Diversity. We will create meaningful opportunities for persons from traditionally underrepresented groups, including those with disabilities, in the employment of project staff and experts, in the composition of our state members' widely varying demographic and cultural profiles, and involvement of teachers, parents, and others in stakeholder groups from design to implementation. We will provide the accommodations needed for full participation including interpreters for staff, partners, and stakeholders who have disability or English proficiency needs. We will ensure the project website will include relevant information and documents in a format that meets a government or industry-recognized standard for accessibility.

Chapter 4 – Management Plan

NCSC partners are a unified team with a decade-long track record as substantive collaborators for conceptually and practically complex and challenging projects. We have well-developed infrastructures on communication, teleconferencing, and networking, and are able to work in high tech and high touch partnerships, an absolute requirement for successful completion of this project. Our management system ensures coherence across our four goal/workgroup/task areas described in the Project Design section as well as four additional management tasks: (1)

assessment design and development; (2) curricular/instructional supports; (3) PD and training; (4) the evaluation process; (5) TAC and expert advisory function; (6) state transition planning; (7) public dissemination; (8) project management. Appendix H includes our organizational chart. The Project Design section describes workgroup activities, and Table 2 shows major milestones (see Appendix E for quarter by quarter timeline).

Table 2: Major Timeline Milestones

Note: each year WG 4 will conduct validity research on within year products/processes
Year 1: Define CCR for SSCD; Overall/state transition plans; Let/award technology RFP; Draft blueprints/PLDs; Produce/validate LPFs and prioritized LPs/entry points; Establish within-state COPs; Produce/conduct CCS training; Finalize ToA
Year 2: Let/award item writing, scoring, call center RFPs; Model items; Conduct content, universal design, bias reviews, and cognitive labs on items; Item bank developed; Develop/pilot skill sequences for ELA; Produce formative assmt and Progress Monitoring (PM) tools; Draft teacher content and teacher/principal effectiveness tools; Communication triage; Produce assmt admin training; Communication Summit
Year 3: Conduct sample field test; Finalize formative assmt and PM tools; Develop/pilot skill sequences for math; Field test tchr/princ tools; Accomm training/results interp training
Year 4: Finalize all components for census field test; Produce final documentation; Let/award standard setting RFP; Finalize PLDs/set final cut scores; Develop/pilot generalization skill sequences in ELA/math; Implement CCS Curricula by COPs; Validate effectiveness tools and develop training; Communication Summit; Finalize technical report; Final validity argument

Management team roles and responsibilities. The management team will monitor and evaluate attainment of goals, objectives, and timelines, identify barriers and solutions to problems

encountered by workgroups or individual collaborative members, and ensure that the research-to-practice efforts honor the contributions, insights, needs, and unique concerns of all collaborative members. They will meet via phone/web monthly, with one face-to-face annual meeting each year, in conjunction with a full project team/state annual face-to-face meeting; the state representatives from each workgroup will be determined at this meeting. The full project team/states will meet 2-4 times per year via distance technology as well. Project Director Quenemoen and Process Evaluator Towles-Reeves will meet by phone with each state partner quarterly, initially to jointly develop a state standards and assessment transition plan, then monitor progress, identify potential barriers, anticipate state unique needs as the work unfolds, and address state concerns throughout the project cycle. They will report back to the management team on state status and refer common issues to the management team or workgroups as appropriate. They will also work directly with workgroup leaders and the external evaluator to monitor and report status of goals, objectives, activities, and timelines, and work with the management team to smooth and integrate cross workgroup efforts. They will oversee development of RFPs and procurement processes, and will manage relationships with all vendors, working with an ad hoc committee of state partners highly skilled in procurement procedures and vendor management. The management team will also support and lead cross-GSEG and RTTA project collaboration with distance meeting technology, and will work to build common understandings across the projects.

State roles and responsibilities. Each state is committed to these basic requirements (see letters of support):1) Active participation in one or more topical area WGs; 2) Involvement of state stakeholders in development processes (e.g., item review, standard-setting); 3) Active participation in pilot and field testing of all components of the systems; 4) Participation in

validity and evaluative studies; and 5) Provision of communication and practice linkages to existing RTTA funded consortia. One state will represent each workgroup on the management team; each workgroup will select a state member to attend/oversee any ad hoc committee work such as LPF development, item review teams, and standard-setting; and at least one state representative will attend the annual face-to-face project meeting.

External evaluation plan. The external evaluator will help ensure that project activities are completed in a timely and quality manner and that proposed goals and objectives have been achieved. The Pacific Institute for Research and Evaluation (PIRE), a leading national evaluation and research firm (www.pire.org), will serve as external evaluator, with Dr. Brent Garrett as lead evaluator. PIRE will work in conjunction with the Management Team and the workgroups to ensure activities are coordinated with minimal burden to state partners. ***1) Types of data to be collected***: Three types of data will be collected to assess the degree, quality, and impact of the implementation of project activities, including data to assess the GSEG GPRA indicators (high quality, relevance, and usefulness) across all project activities and formative data to ensure fidelity of implementation to project design. It will include professional development evaluations, surveys, interviews, focus groups, etc. with project staff, state personnel, and expert partners, and summative impact data. Examples of summative data sources are validity evaluation reports, state-level assessment reports, and attainment of state transition plans. **2)** ***When data will be collected***: Formative and GPRA data designed to assess the degree and quality of implementation will be collected on an ongoing basis, and reported quarterly. Summative data will be collected, analyzed, and reported at the end of each project year to assess the impact NCSC has had on partner states. ***3) Methods***: A variety of methods will be used including document analyses, interviews, surveys, and focus groups. Other methods will be

identified by the Evaluation Workgroup. ***4) Instruments to be developed***: PIRE will work with the Evaluation Workgroup to develop research-based instruments that meet the specific needs of the project. Evaluation rubrics will be developed for analyzing state reports and validity evaluation reports. All instruments will be developed, tested, and implemented in accordance with standard evaluation protocols (Fowler, 2002; Dillman, 1999; Krueger & Casey, 2000). ***5) Data analysis*:** Quantitative and qualitative techniques will be used to analyze and summarize evaluation data. Frequencies, percentages, cross-tabulations, t-tests, and more sophisticated analyses, as needed, will be used to summarize quantitative data. Qualitative data will be analyzed through the exploration of important themes that evolve from the document analyses, surveys, interviews, focus groups, and other qualitative assessments. ***6) Reporting schedule*:** PIRE will work closely with the Evaluation Workgroup to develop a reporting system to meet the needs of the project and the funder. Preliminary formative data will be reported and shared with project management on a quarterly basis. An annual narrative report will be developed, and assistance provided to project management in the completion of the project's yearly Performance Reports. ***7) Using data to monitor progress and to provide accountability information:*** Quarterly formative reports will be developed and disseminated to monitor progress toward completing project activities. The Performance Report data will be used to provide accountability information. Other reports will be developed as needed or requested by the Management Team and in cooperation with the workgroups.