

**Race to the Top Assessment Program
Application for New Grants**

**Comprehensive Assessment Systems
CFDA Number: 84.395B**

**Submitted by Washington State,
on behalf of the SMARTER Balanced Assessment Consortium**



**CLOSING DATE: 06/23/2010
U.S. Department of Education
Washington, D.C. 20202
Form Expiration Date: 10/31/2010
OMB Control Number: 1810-0699
Paperwork Burden Statement**

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1810-0699. The time required to complete this information collection is estimated to average 502.25 hours per response, including the time to review instructions, search existing data resources, and gather the data needed, and complete and review the information collection. If you have comments or concerns regarding the status of your individual submission of this form, write directly to: James Butler, Race to the Top Assessment, Office of Elementary and Secondary Education, U.S. Department of Education, 400 Maryland Ave., S.W., Room 3C108, Washington, D.C. 20202-3118.

TABLE OF CONTENTS FOR APPLICATION (PART I.C.)	Page Number
Executive Summary	5
Application Assurances	7
Eligibility Requirements	12
Selection Criteria	16
(A)(1) Consortium Governance	17
(A)(2) Theory of Action	31
(A)(3) Assessment System Design	40
(A)(4) Assessment System Development	72
(A)(5) Research and Evaluation	88
(A)(6) Professional Capacity and Outreach	110
(A)(7) Technology Approach	117
(A)(8) Project Management	123
Competition Priorities	145
Competitive Preference Priority	146
References	162

TABLE OF CONTENTS FOR APPENDIX (PART I.C.)		
TITLE OF ATTACHMENT	Relevant Requirement, Selection Criterion, or Priority	Page/ Attachment Number
SBAC Organizational Structure	(A)(1)(b)(i)	X-1 / A1-1
Responsibilities of the Consortium’s Committees and Leadership	(A)(1)(b)(i)	X-2 / A1-2
Consortium Processes for Determining Committee/ Working Group Members and Electing Leaders	(A)(1)(b)(i)	X-3 / A1-3
SMARTER Balanced Assessment Consortium Memoranda of Understanding	(A)(1)(c)(i)	X-5 / A1-4
Overview of the Theory of Action	(A)(2)	X-6 / A2-1
Overview of Assessment System	(A)(3)(a)	X-7 / A3-1
Access by Design: A Theory of Action and Research	(A)(3)(b)(iii)	X-11 / A3-2
Sample Items	(A)(3)(c)(iv)	X-14 / A3-3
Reporting Characteristics by Assessment Type	(A)(3)(c)(vii)	X-39 / A3-4
Examples of Learning Progressions in ELA and Mathematics	(A)(4)(a)	X-40 / A4-1
Overview of Development Process for Summative Assessment	(A)(4)(a)	X-46 / A4-2
Goals for Enhanced Assessment Grant for Accommodations	(A)(4)(b)	X-47 / A4-3
Reporting Exemplar – Colorado Electronic Portal for Review of Performance Data	(A)(4)(d)	X-50 / A4-4a
Reporting Exemplar – Colorado Growth Model Sample District-Level Achievement and Growth Charts for Math, Reading, Writing	(A)(4)(d)	X-51 / A4-4b
Reporting Exemplar – Colorado’s Understanding Your Child’s Academic Progress	(A)(4)(d)	X-52 / A4-4c

Reporting Exemplar: Oregon’s User Guide for the Online Reporting System	(A)(4)(d)	X-56 / A4-4d
Reporting Exemplar: NAEP State Profile Data Base Report	(A)(4)(d)	X-92 / A4-4e
Overview of Professional Capacity Building	(A)(6)(a)	X-93 / A6-1
Overview of Development Process for Interim/ Benchmark Assessments	(A)(6)(a)	X-94 / A6-2
RFP for Grant Project Manager	(A)(8)(a)	X-95 / A8-1
Washington RFP for Project Management Partner	(A)(8)(a)	X-99 / A8-2
WestEd State Assessment Experience: Overview	(A)(8)(a)	X-124 / A8-3
Summary of Corporate Capacity of Interim Project Management Partner	(A)(8)(a)	X-125 / A8-4
Resumes for Interim Project Management Partner Key Staff	(A)(8)(a)	X-127 / A8-5
Budget Part A: Summary Table Part B: Level 1 Modules Part C: Level 2 Module	(A)(8)(c)	X-150 / A8-6 X-150 / A8-6 X-151 / A8-6 X-186 / A8-6
Survey of Operational Costs for SBAC States	(A)(8)(d)	X-189 / A8-7
IHE Letters of Intent	Competitive Preference Priority	X-190 / CPP-1

EXECUTIVE SUMMARY (PART I.D.)

The SMARTER Balanced Assessment Consortium’s (SBAC) vision for a new generation assessment system—one that includes a set of balanced components that can be adapted to meet students’ needs across participating States—is rooted in a concern for the valid, reliable, and fair assessment of the deep disciplinary understanding and higher-order thinking skills that are increasingly demanded by a knowledge-based global economy. This vision also is based on the belief that assessment must support ongoing improvements in instruction and promote meaningful learning experiences for students that lead to outcomes valued by all stakeholders.

The overarching goal of the SBAC is to ensure that all students leave high school prepared for postsecondary success in college or a career through a planned sequence of educational experiences and opportunities. To meet this goal, with support from institutions of higher education (IHEs) and workplace representatives, the Consortium will build upon the strong foundation in each participating State to create a high-quality, balanced multi-state assessment system based on the Common Core State Standards (CCSS). The role of the Consortium in this process is to guide the development and implementation of an assessment system that reshapes educational practice in participating States in strategic ways and leads to improved learning outcomes for students.

Following the principle of “responsible flexibility,” SBAC will provide options for customizable system components while also ensuring comparability of high-stakes summative test results across States. In addition, the Consortium is committed to creating a policy environment that fosters innovation while supporting the development of accountability systems that incentivize the right behaviors for students, teachers, and administrators and avoid inadvertently incentivizing behaviors that run counter to SBAC goals.

The comprehensive assessment system proposed by the Consortium calls for strategic use of a variety of item types and performance events to measure the full range of the CCSS and to ensure accurate assessment of all students, including students with disabilities, English learners, and low- and high-performing students. Specifically, SBAC proposes to implement a system that features the following:

- Common CCSS-based computer adaptive summative assessments that make use of technology-enhanced item types and teacher-developed and scored performance events;
- Computer adaptive interim/benchmark assessments—reflecting learning progressions or

content clusters—that provide more in-depth and/or mid-course information about what students know and can do in relation to the CCSS;

- Research-supported instructionally sensitive tools, processes, and practices developed by State educators that can be used formatively at the classroom level to improve teaching and increase learning;
- Focused ongoing support to teachers through professional development opportunities and exemplary instructional materials linked to the CCSS;
- Online reporting and tracking system that enables access to key types of information about student progress toward college- and career-readiness and about specific strengths and limitations in what students know and are able to do at each grade level; and
- Cross-State communications network to inform stakeholders about SBAC activities and ensure a common focus on the goal of college- and career-readiness for all students.

In a number of ways, innovative and efficient use of technology serves as the backbone of this balanced assessment system: (1) SBAC’s system capitalizes on the precision and efficiency of computer adaptive testing; (2) the expanded use of technology enables the Consortium’s goals of developing innovative and real-world item types that ensure measurement of student achievement across a wide performance continuum and provide efficiencies and enhancements for teacher and administrator professional development and capacity building at the local level; and (3) through use of an interoperable electronic platform and leveraging of cross-State resources, SBAC can deliver assessments and produce both standardized and customizable reports that are cost-effective, timely, and useful for a range of audiences in tracking and analyzing the progress towards college- and career-readiness of individual students, student subgroups, classrooms, schools, districts, and States.

In summary, SBAC’s proposed learning and assessment system is grounded in a sound Theory of Action. This system promotes research-supported instructional practice and incorporates a balanced set of technology-supported tools, innovative assessments, and state-of-the-art classroom support mechanisms that work coherently to support teaching and learning. Over time, with a purposeful governing structure and IHEs in participating States, this Consortium’s assessment system holds promise to effect the types of reform sought by the Race to the Top Assessment Program.

APPLICATION ASSURANCES (PART I.E.)

Race to the Top Assessment Program Comprehensive Assessment Systems Grant Application Assurances	
Legal Name of Applicant: State of Washington, Office of Superintendent of Public Instruction	Applicant's Mailing Address: PO Box 47200 Olympia, WA 98504-7200
Employer Identification Number: 91-6001112	Organizational DUNS: 808882898
Contact on Matters Involving this Application: Dr. Joseph L. Willhoft	Contact Position and Office: Assistant Superintendent – Assessment & Student Information
Contact Telephone: 360.725.6336	Contact Mailing Address: PO Box 47200 Olympia, WA 98504-7200
Contact E-mail Address: joe.willhoft@k12.wa.us	

SIGNATURE BLOCKS FOR APPLICANT

Directions: If one member State of the consortium is applying for a grant on behalf of the consortium, use the following signature block. The applicant State is the only State that must sign below. If the State has a president of the State Board of Education, then the signature of the State Board of Education President is applicable. However, if a State has no State Board of Education, then the signature is not applicable.

**APPLICANT SIGNATURE BLOCK for Race to the Top Fund Assessment Program
Comprehensive Assessment Systems Grant Application Assurances**

To the best of my knowledge and belief, all of the information and data in this application are true and correct.

I certify on behalf of the consortium that each member of the consortium has agreed to be bound by every statement and assurance in the application and that each Governing State is fully committed to the application and will support its implementation.

I further certify that I have read the application, am fully committed to it, and will support its implementation.

State Name:

WASHINGTON

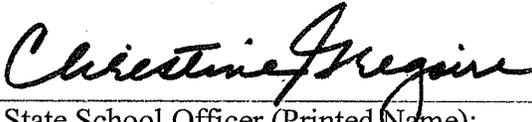
Governor or Authorized Representative of the Governor (Printed Name):

Christine O. Gregoire

Telephone:

(360) 902-4123

Signature of Governor or Authorized Representative of the Governor:



Date:

6/9/10

Chief State School Officer (Printed Name):

Randy I. Dorn

Telephone:

(360) 725-6004

Signature of the Chief State School Officer:



Date:

6/10/10

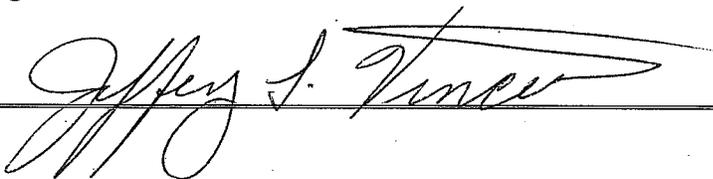
President of the State Board of Education, if applicable (Printed Name):

Jeff Vincent, Chair

Telephone:

(360) 725-6025

Signature of the President of the State Board of Education, if applicable:



Date:

6/15/10

ACCOUNTABILITY, TRANSPARENCY, REPORTING, PROCUREMENT AND OTHER ASSURANCES AND CERTIFICATIONS

Accountability, Transparency and Reporting Assurances

The applicant assures that it will comply with all of the accountability, transparency, and reporting requirements that apply to the Race to the Top Assessment program, including the following:

- For each year of the program, the applicant assures that it will comply with the requirements of the American Recovery and Reinvestment Act (ARRA) Division A, Section 14008, and other performance reporting that the Department may require.
- The applicant will cooperate with any U.S. Comptroller General evaluation of the uses of funds and the impact of funding on the progress made toward closing achievement gaps. (ARRA Division A, Section 14009)
- If the applicant uses funds for any infrastructure investment, the applicant will certify that the investment received the full review and vetting required by law and that the chief executive accepts responsibility that the investment is an appropriate use of taxpayer funds. This certification will include a description of the investment, the estimated total cost, and the amount of covered funds to be used. The certification will be posted on the applicant's website and linked to www.Recovery.gov. A State or local agency may not use funds under the ARRA for infrastructure investment funding unless this certification is made and posted. (ARRA Division A, Section 1511)
- The applicant will submit reports, within 10 days after the end of each calendar quarter, that contain the information required under section 1512(c) of the ARRA in accordance with any guidance issued by the Office of Management and Budget or the Department. (ARRA Division A, Section 1512(c))
- The applicant will cooperate with any appropriate Federal Inspector General's examination of records under the program. (ARRA Division A, Section 1515)

Procurement Assurance

The applicant assures or certifies the following:

- A competitive procurement process based on a "best value" selection will be used for tasks related to assessment design and development under the grant. All Federal and ARRA procurement requirements will be met under the grant.

Other Assurances and Certifications

The applicant assures or certifies the following:

- The applicant will comply with all applicable assurances in OMB Standard Forms 424B (Assurances for Non-Construction Programs), including the assurances relating to the legal authority to apply for assistance; access to records; conflict of interest; merit systems; nondiscrimination; Hatch Act provisions; labor standards; protection of human subjects;

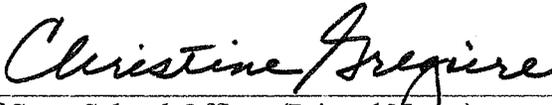
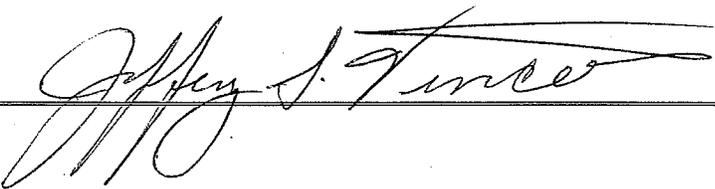
animal welfare; Single Audit Act; and the general agreement to comply with all applicable Federal laws, executive orders and regulations.

- With respect to the Certification regarding Lobbying (formerly in Department Form 80-0013), no Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the making or renewal of Federal grants under this program; the applicant will complete and submit Standard Form-LLL, "Disclosure Form to Report Lobbying," when required (34 C.F.R. Part 82, Appendix B).
- The applicant will comply with all of the operational and administrative provisions in Title XV and XIV of the ARRA. In using ARRA funds for infrastructure investment, recipients will comply with the requirement regarding Preferences for Quick Start Activities (ARRA Division A, Section 1602).
- Any local educational agency (LEA) receiving funding under this program will have on file with the State a set of assurances that meets the requirements of section 442 of the General Education Provisions Act (GEPA) (20 U.S.C. 1232e).
- Any LEA receiving funding under this program will have on file with the State (through either its Stabilization Fiscal Stabilization Fund application or another U.S. Department of Education Federal grant) a description of how the LEA will comply with the requirements of section 427 of GEPA (20 U.S.C. 1228a). The description must include information on the steps the LEA proposes to take to permit students, teachers, and other program beneficiaries to overcome barriers (including barriers based on gender, race, color, national origin, disability, and age) that impede access to, or participation in, the program.

The applicant and other entities will comply with the Education Department General Administrative Regulations (EDGAR), including the following provisions as applicable: 34 CFR Part 74–Administration of Grants and Agreements with Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations; 34 CFR Part 75–Direct Grant Programs; 34 CFR Part 77– Definitions that Apply to Department Regulations; 34 CFR Part 79–Intergovernmental review of Department of Education programs and activities; 34 CFR Part 80– Uniform Administrative Requirements for Grants and Cooperative Agreements to State and Local Governments, including the procurement provisions; 34 CFR Part 81– General Education Provisions Act–Enforcement; 34 CFR Part 82– New Restrictions on Lobbying; 34 CFR Part 84– Governmentwide Requirements for Drug-Free Workplace (Financial Assistance); 34 CFR Part 85–Governmentwide Debarment and Suspension (Nonprocurement); 34 CFR Part 86–Drug and alcohol abuse prevention; 34 CFR Part 97–Protection of human subjects; 34 CFR Part 98– Student rights in research, experimental programs, and testing; and 34 CFR Part 99– Family education rights and privacy.

**ACCOUNTABILITY, TRANSPARENCY, REPORTING AND PROCUREMENT
ASSURANCES SIGNATURE BLOCKS FOR APPLICANT**

Directions: If one member State of the consortium is applying for a grant on behalf of the consortium, use the following signature block. The applicant State is the only State that must sign below. If the State has a president of the State Board of Education, then the signature of the State Board of Education President is applicable. However, if a State has no State Board of Education, then the signature is not applicable.

ACCOUNTABILITY, TRANSPARENCY, REPORTING AND PROCUREMENT ASSURANCES SIGNATURE BLOCK for Race to the Top Fund Assessment Program Comprehensive Assessment Systems Grant Application Assurances	
State Name: WASHINGTON	
Governor or Authorized Representative of the Governor (Printed Name): Christine O. Gregoire	Telephone: (360) 902-4123
Signature of Governor or Authorized Representative of the Governor: 	Date: 6/9/10
Chief State School Officer (Printed Name): Randy I. Dorn	Telephone: (360) 725-6004
Signature of the Chief State School Officer: 	Date: 6/10/10
President of the State Board of Education, if applicable (Printed Name): Jeff Vincent, Chair	Telephone: (360) 725-6025
Signature of the President of the State Board of Education, if applicable: 	Date: 6/15/10

ELIGIBILITY REQUIREMENTS (PART I.G.)

Eligibility Requirement (1):

To be eligible to receive an award under this category, an eligible applicant must include a minimum of 15 States, of which at least 5 States must be governing States (as defined in the NIA).

Consortium of States (Governing States are identified with an asterisk.)	
1. Washington *	2. Missouri *
3. Connecticut *	4. Nevada *
5. Utah *	6. Idaho *
7. Maine *	8. Wisconsin *
9. North Carolina *	10. Oregon *
11. New Mexico *	12. Hawaii *
13. Vermont *	14. Kansas *
15. Michigan *	16. Montana *
17. West Virginia *	18. Ohio
19. Iowa	20. South Carolina
21. South Dakota	22. Colorado
23. North Dakota	24. Delaware
25. Alabama	26. Kentucky
27. New Hampshire	28. Pennsylvania
29. Oklahoma	30. New Jersey
31. Georgia	
<p>Washington State (#1) is a Governing State in addition to serving in the unique role of Lead Procurement State/Lead State for the Consortium.</p>	

Eligibility Requirement (2):

To be eligible to receive an award under this category, an eligible applicant must identify in its application a proposed project management partner and provide an assurance that the proposed project management partner is not partnered with any other eligible applicant applying for an award under this category. [Please see information on Contracting for Services in Part II.F.]

Consortium's proposed Project Management Partner:

WestEd – Interim partner for period not to exceed three months from date of notice of award.

Washington, as the fiscal agent for the Consortium, is in progress with the solicitation process for the primary Project Management Partner. The selection process and contracting implementation is scheduled for completion by October 1, 2010.

Contact information for proposed Project Management Partner:

WestEd (interim)
Stanley Rabinowitz, Ph.D.
Senior Program Director
Assessment and Standards Development Services
730 Harrison Street
San Francisco, CA 94107-1242
T: 415.615.3154
F: 415.565.3012

Check the box:

- ✓ The applicant assures that the proposed Project Management Partner is not partnered with other eligible applicants.

Per the Department's NIA for the Comprehensive Assessment System Grants, the Consortium views the role of the Project Management Partner as an integral part of the grant implementation and that the associated efforts of the Project Management Partner are part of the grant award itself. With the grant funding not available for actual work until October 1, 2010, the Consortium recognized the potential for project start-up delays as the Project Management Partner familiarized itself with the activities and objectives of SBAC. Concerns of this nature resulted in the Consortium Steering Committee's recognition of need for a transition plan and transition partner. The Consortium has implemented the option within the existing proposal project management contract with WestEd to extend the duration and appoint WestEd interim Project Management Partner for a period not to exceed three months from the date of notice of

an award. The purpose of this interim Project Management Partner contract is to explicitly meet the requirements of the NIA and to allow a smooth transition to immediate first steps as a funded Consortium while proceeding with the comprehensive procurement process required by the State of Washington.

Pursuant to the procurement rules of Washington State, a solicitation for a Project Management Partner is underway. The associated Request for Proposal (RFP) can be found on Washington’s Office of Superintendent of Public Instruction website. Viewers need to locate RFP No. 2010-07, *Comprehensive Assessment Systems Grant – Project Management Partner*, in the table of active solicitations. The estimated schedule of procurement activities is listed below.

Estimated Schedule of Procurement Activities	
Issue Request for Proposal	June 4, 2010
Letter of Intent to Propose Due	June 18, 2010
Question and Answer Period	June 7 – July 2, 2010
Pre-proposal Video Conference	June 29, 2010
Last Date for Questions Regarding RFP	July 2, 2010
Release Q&A plus RFP Addendum (as needed)	July 9, 2010
Proposals Due	July 30, 2010
Evaluate written proposals	August 2 – August 13, 2010
Announce “Apparent Successful Contractor”	August 16, 2010
Hold Debriefings (if requested)	August 17– August 19, 2010
Negotiate Contract	August 30 – September 10, 2010
File Contract with OFM	September 16, 2010
Begin Contract Work	October 1, 2010

Eligibility Requirement (3):

To be eligible to receive an award under this category, an eligible applicant must submit assurances from each State in the consortium that, to remain in the consortium, the State will adopt a common set of college- and career-ready standards (as defined in the NIA) no later than December 31, 2011, and common achievement standards (as defined in the NIA) no later than the 2014-2015 school year.

Check the applicable box:

- ✓ Signed assurances are included in Appendix A1-4.
- ✓ See page 3 and pages 14–15 in the MOU (located in Appendix A1-4).
- ✓ See page 16 in the MOU (located in Appendix A1-4) for chief procurement certification.

Assurances that States in the Consortium will adopt the Common Core State Standards by December 31, 2011, and common achievement standards by 2014–15 can be found in the MOU on page 3 (Appendix A1-4). Signature blocks binding States to the terms of the MOU can be found on page 14 for Advisory States and page 15 for Governing States. Certification by the chief procurement official of each State can be found on page 16.

SELECTION CRITERIA (PART I.H.)

(A)(1) Consortium Governance (up to 20 points)

The extent to which the consortium’s proposed governance structure will enable the successful design, development, and implementation of the proposed assessment system. In determining the extent to which the consortium’s proposed governance structure will enable the successful design, development, and implementation of the proposed assessment system, we will consider—

- (a) The consortium’s vision, goals, role, and key deliverables (*e.g.*, assessment components, scoring and moderation system, professional development activities), and the consistency of these with the consortium’s theory of action;
- (b) The consortium’s structure and operations, including—
 - (i) The organizational structure of the consortium and the differentiated roles that a member State may hold (*e.g.*, lead State, governing State (as defined in the NIA), advisory State); and
 - (ii) For each differentiated role, the rights and responsibilities (including the level of commitment to adopting and implementing the assessment system) associated with the role;
 - (iii) The consortium’s method and process (*e.g.*, consensus, majority) for making different types of decisions (*e.g.*, policy, operational);
 - (iv) The protocols by which the consortium will operate, including the protocols for member States to change roles or leave the consortium and for new member States to join the consortium;
 - (v) The consortium’s plan, including the process and timeline, for setting key policies and definitions for the proposed assessment system, including a common set of college- and career-ready standards (as defined in the NIA), a common set of performance level descriptors (as defined in the NIA), a common set of achievement standards (as defined in the NIA), common assessment administration procedures, common item release and test security policies, a common definition of “English learner,” and a common set of policies and procedures for accommodations (as defined in the NIA) and student participation; and
 - (vi) The consortium’s plan for managing funds received under this grant category;
- (c) The terms and conditions of the Memoranda of Understanding or other binding agreements executed by each member State, including—
 - (i) The consistency of the terms and conditions with the consortium’s governance structure and the State’s role in the consortium; and
 - (ii) The State’s commitment to and plan for identifying any existing barriers in State law, statute, regulation, or policy to implementing the proposed assessment system and to addressing any such barriers prior to full implementation of the summative assessment components of the system; and
- (d) The consortium’s procurement process, and evidence of each member State’s commitment to that process.

Required Tables and/or Attachments:

- ✓ *Summary Table for (A)(1)(b)(ii): States' Roles in the Consortium*
- ✓ *Summary Table for (A)(1)(b)(v): Consortium's Policy and Definition Time Line*

Optional:

- ✓ *For (A)(1)(b)(i): A visual model that conveys the consortium's organizational structure.*

(A)(1)(a) Consortium's Vision, Goal, Role, Key Deliverables

The SMARTER Balanced Assessment Consortium's (SBAC's) **vision** for a next generation assessment system—a system that is balanced and responsive to students' needs—is rooted in a commitment to the valid, reliable, and fair assessment of the deep disciplinary understanding and higher-order thinking skills that are pre-requisites for postsecondary success. It is also rooted in the belief that assessment must support ongoing improvements in instruction and learning and must be useful for all members of the educational enterprise: students, parents, teachers, school administrators, institutions of higher education (IHE) stakeholders, policymakers, and members of the public.

The overarching **goal** of SBAC is to ensure that, through improved teaching and increased student learning, all students leave high school prepared for postsecondary success in college or career. The Consortium recognizes that, if all high school graduates are to be prepared in this way, many aspects of the current education system must be improved, and all aspects must be better coordinated and more fully integrated. Thus, the **role** of the Consortium is to influence and support the development and implementation of learning and assessment systems to radically reshape the education systems in participating States so as to improve student outcomes.

Essential, though not sufficient, are development of a high-quality assessment system that strategically “balances” summative, interim, and formative components (Darling-Hammond & Pecheone, 2010; Rabinowitz, 2010); use of valid measurement across the full range of common rigorous academic standards, including assessment of the deep disciplinary understanding and higher-order thinking skills that are increasingly demanded by a knowledge-based economy; and establishment of clear, internationally benchmarked performance expectations (Schmidt, Wang, & McKnight, 2005). Other important elements that are outside SBAC's direct scope of work, but not outside its influence, are comprehensive pre-service and in-service teacher professional development and focused and valid systems of accountability. The Consortium proposes employing technology, research-supported distributed assessment components, and varied assessment types, while at the same time increasing educator assessment literacy and

involvement—all with the intention of maximizing the relevance and utility of the information produced from the various components of the assessment system.

Consortium **key deliverables** are

1. **A comprehensively designed assessment system** that includes the strategic use of a variety of item types (i.e., selected-response items, constructed-response items, technology-enhanced items, and performance events) to assess the full range of the Common Core State Standards (CCSS) with an emphasis on problem solving, analysis, synthesis, and critical thinking. This balanced system will incorporate a required Statewide summative assessment, with two types of optional assessments and tools used to inform instruction and help students understand where they are in their learning: (1) interim/benchmark (I/B) assessments used to track students’ learning progress at key points during the year, and (2) a variety of formative tools, processes, and practices for teachers to use to understand what students are and are not learning so as to adjust instruction accordingly.¹ This balanced system is designed to provide accurate assessment of all students—including students with disabilities (SWDs), English learners (ELs), and low- and high-performing students—for a variety of learning and accountability purposes. The summative component will be administered as a computer adaptive assessment and will include performance events. Evidence of student competencies will be linked to the summative system.
2. **An online adaptive test administration** with a secure item and performance event bank that includes psychometric attributes required to score the assessment in a comparable manner with other State members, and access to other applications determined to be essential to the implementation of the system.
3. **A consolidated reporting system** that enhances understanding of student progress toward college- and career-readiness for all members of the educational enterprise. The reporting system will be based on psychometrically sound scaling and equating procedures in order to provide reliable, valid, and fair scores that can be used to evaluate student achievement and growth. Reports will produce data for specified purposes, one of

¹ The interim/benchmark and formative components are described in the MOU as optional “formative/benchmark components,” which is a broader description for these system features.

which will be Title I accountability (e.g., school, teacher, and principal effectiveness). The achievement standards and achievement level descriptors will be internationally benchmarked.

4. **A system of professional development** focused on assessment literacy. To support curricular goals, including expected learning progressions, the Consortium will develop formative assessment tools related to curriculum and lesson development, as well as scoring and examination of student work. Because a key element of SBAC's professional learning approach for educators is to engage teachers directly in developing and scoring SBAC assessments, teachers and administrators will be asked to contribute to the item and performance event banks and participate in the moderated scoring process.

(A)(1)(b)(i) As of June 2010, 31 States make up the SMARTER Balanced Assessment Consortium. Together, they account for the Total State Membership, which, as shown in the Organizational Structure (see Appendix A1-1), includes the Lead Procurement State/Lead State, Governing States, and Advisory States.

The Lead Procurement State/Lead State (hereafter referred to as the Lead Procurement State) is Washington. Washington's unique role is described in section (A)(1)(b)(ii).

As of June 2010, 17 of the 31 member States are Governing States: They have fully committed to this Consortium alone and meet the qualifications specified in the Memorandum of Understanding (MOU); are members of only one Consortium applying for a grant in the Program; have an active role in policy decision-making for the Consortium; provide one representative to serve on the Steering Committee; provide representative(s) to serve on one or more working groups; approve the Steering Committee Members and the Executive Committee Members; and participate in the final decision-making of changes in governance and specific design elements.

As of June 2010, 14 of the 31 member States are Advisory States: They have not fully committed to any consortium but support the work of this Consortium; participate in all Consortium activities but do not have a vote unless the Steering Committee deems it beneficial to gather input on decisions or chooses to have the Total State Membership vote on an issue; may contribute to policy, logistical, and implementation discussions that are necessary to fully

operationalize the assessment system developed by the Consortium; and are encouraged to participate in the working groups.

All Consortium efforts will be overseen by the Steering Committee. A subset of this group is the Executive Committee, consisting of two Co-chairs, a representative from the Lead Procurement State, a representative from higher education, and one representative each from four Governing States. The Executive Committee will be accountable for overseeing development of the assessment system. The Steering Committee—a larger group composed of one representative from each Governing State in the Consortium—will support the Executive Committee and, to date, has been accountable for determining the broad picture of the assessment system. The Lead Procurement State is accountable for managing funds and all procurement on behalf of the Consortium. For additional details on committee responsibilities, see Appendix A1-2.

To maximize contributions and distribute workload efficiently and effectively, the Consortium has also created working groups, each charged with specific responsibilities, whose members are assigned to the group based on skills and expertise. To be sure all working groups are accomplishing their specific goals and contributing efficiently toward Consortium goals, Consortium leadership will review and revise group structures as appropriate. See Appendix A1-3 for a description of Consortium processes for determining committee/working group members and electing leaders.

Together, the Executive Committee, the Steering Committee, and the Lead Procurement State will work closely with the Project Management Partner and USED to ensure that Consortium efforts are coordinated efficiently and effectively to realize the design, development, and implementation of the assessment system, consistent with the Consortium's Theory of Action.

As illustrated in the Organizational Structure (Appendix A1-1), the Technical Advisory Committee (TAC) will provide expertise and advice to the Consortium in the areas of curriculum/instruction, assessment design, and technology. In addition, SBAC has secured commitments from the following organizations which have expressed their willingness to serve on the Policy Advisory Committee: American Association of School Administrators (AASA), American Federation of Teachers (AFT), American Youth Policy Forum (AYPF), Association for Career and Technical Education (ACTE), Mexican American Legal Defense & Educational Fund (MALDEF), National Association for the Education of Young Children (NAEYC), National Association of State Directors of Special Education (NASDSE), National Association

of Elementary School Principals (NAESP), National Association of Secondary School Principals (NASSP), National Association of State Boards of Education (NASBE), National Educators Association (NEA), National School Boards Association (NSBA), The James B. Hunt, Jr. Institute for Educational Leadership and Policy, Alliance for Excellence, American Association of State Colleges and Universities (AASCU), American Council on Education (ACE), and State Higher Education Executive Officers (SHEEO).

(A)(1)(b)(ii) Described below are the rights and responsibilities that pertain to all States in the Consortium, irrespective of role, followed by the rights and responsibilities that are based on a State's particular role in the Consortium.

Each Consortium member State, irrespective of role, is entitled to all key deliverables outlined above in (A)(1)(a). As for responsibilities, each member State is responsible for adopting the CCSS no later than December 31, 2011, and each State that is a member of the Consortium in 2014–15 will also be responsible for adopting common achievement standards and fully implementing Statewide, no later than the 2014–15 school year, the Consortium's summative assessment in grades 3–8 and high school, for both English language arts and mathematics. In addition, all member States are expected to adhere to the governance as outlined in the MOU; support decisions of the Consortium; follow agreed-upon time lines; participate in the decision-making process; and identify any barriers in State law, statute, regulation, or policy to implementing the assessment system and address any such barriers prior to full implementation of the summative assessment components of the system.

In addition to its general rights and responsibilities as a member State, each Governing State will be entitled to one representative on the Steering Committee, with one vote on all committee decisions. Each Governing State will be responsible for contributing significantly to policy, logistical, and implementation discussions that are necessary to fully operationalize the assessment system; participating significantly in Consortium-wide activities; participating in at least one working group; and approving individuals selected for representation and final decision-making in the following key areas: Steering Committee members, Executive Committee members, changes in governance and other official documents, and specific design elements.

In addition to its general rights and responsibilities as a member State, each Advisory State will be entitled to one vote on all polls of the Total State Membership. Each Advisory State also

will be responsible for contributing to policy, logistical, and implementation discussions that are necessary to fully operationalize the proposed assessment system; and participating in Consortium-wide activities.

The Lead Procurement State has the authority allocated to the Governing States and is also entitled to Washington's negotiated indirect rate for federal grants. The Lead Procurement State also has the responsibilities assigned to the Governing States and, in addition, is responsible for overseeing the management of funds, in collaboration with the Steering Committee and Executive Committee; overseeing all procurement on behalf of the Consortium; and providing one representative to serve on the Executive Committee.

All of these rights and responsibilities are summarized in the following table.

Summary Table for (A)(1)(b)(ii): States' Roles in the Consortium

Role Types of Member States	Description of the Rights and Responsibilities Associated with Role	Member States in this Role
<p>Total State Membership</p> <p>(Governing States, Advisory States, Lead Procurement State/Lead State)</p> <p>Note: The States in this role will be updated as the State membership changes.</p>	<p>All States in the Consortium, by the 2014–15 school year, will be entitled to the following:</p> <ol style="list-style-type: none"> 1. Comprehensively designed assessment system 2. Online test administration with a secure item and performance event bank 3. Consolidated reporting system 4. System of professional development <p>All States in the Consortium will be responsible for the following:</p> <ul style="list-style-type: none"> • Adopt the CCSS no later than December 31, 2011. <p>All States that are in the Consortium in 2014–15 will also be responsible for the following:</p> <ul style="list-style-type: none"> • Adopt common achievement standards no later than the 2014–15 school year, and • Fully implement the summative assessment in grades 3–8 and high school for both ELA and mathematics and no later than the 2014–15 school year. <p>In addition, States are expected to</p> <ul style="list-style-type: none"> • Adhere to the governance as outlined in the MOU, • Support Consortium decisions, • Follow agreed-upon time lines, • Participate in decision-making, and • Identify and implement a plan to address State barriers prior to full implementation of the summative assessment. 	<p>Washington Missouri Connecticut Nevada Utah Idaho Maine Wisconsin North Carolina Oregon New Mexico Hawaii Vermont Kansas Michigan Montana West Virginia Ohio Iowa South Carolina South Dakota Colorado North Dakota Delaware Alabama Kentucky New Hampshire Pennsylvania Oklahoma New Jersey Georgia</p>
<p>Governing State – Additional rights and responsibilities</p>	<p>Each Governing State is entitled to</p> <ul style="list-style-type: none"> • One representative on the Steering Committee, and • One vote on all decisions brought to the Steering Committee. <p>Each Governing State has the responsibility to</p>	<p>Washington Missouri Connecticut Nevada Utah Idaho Maine</p>

Role Types of Member States	Description of the Rights and Responsibilities Associated with Role	Member States in this Role
<p>Note: The States in this role will be updated as the State membership changes.</p>	<ul style="list-style-type: none"> • Contribute significantly to policy, logistical, and implementation discussions to fully operationalize the assessment system, • Participate significantly in Consortium-wide activities, • Participate in at least one working group, • Approve Steering and Executive Committee members, and • Approve final decision-making in changes in governance, design elements, and other official documents. 	<p>Wisconsin North Carolina Oregon New Mexico Hawaii Vermont Kansas Michigan Montana West Virginia</p>
<p>Advisory State – Additional rights and responsibilities</p> <p>Note: The States in this role will be updated as the State membership changes.</p>	<p>Each Advisory State is entitled to</p> <ul style="list-style-type: none"> • One vote when the Total State Membership is polled. <p>Each Advisory State has the responsibility to</p> <ul style="list-style-type: none"> • Contribute to policy, logistical, and implementation discussions to fully operationalize the assessment system, and • Participate in Consortium-wide activities. 	<p>Ohio Iowa South Carolina South Dakota Colorado North Dakota Delaware Alabama Kentucky New Hampshire Pennsylvania Oklahoma New Jersey Georgia</p>
<p>Lead Procurement State/Lead State – Additional rights and responsibilities</p>	<p>The Lead Procurement State has the authority allocated to Governing States and, in addition, to</p> <ul style="list-style-type: none"> • Receive the State’s negotiated indirect rate for federal grants. <p>The Lead Procurement State has the responsibilities assigned to Governing States, and, in addition, is responsible for</p> <ul style="list-style-type: none"> • Overseeing the management of funds, in collaboration with the Steering Committee and Executive Committee, • Overseeing all procurement on behalf of the Consortium, and • Providing one representative to serve on the Executive Committee. 	<p>Washington</p>

(A)(1)(b)(iii) Consensus will be the goal for all Consortium decisions. Major decisions for which there is no consensus will be resolved with a simple majority vote. The Steering Committee will determine what issues will be referred to the Total State Membership. Each State will have one vote when polls are conducted. If there is only a one- to three-vote difference, the issue will be re-examined to seek greater consensus. The Executive Committee will decide which decisions or issues are referred to the Steering Committee. The Consortium strives for fairness and transparency in all decisions, and this structure will ensure that polls are conducted knowledgeably and responsibly with strong State participation.

(A)(1)(b)(iv) Membership in the Consortium is assured when (1) a State declares its membership level and required MOU signatures are secured (from the State's Commissioner, State Superintendent, or Chief; Governor; and, if the State has one, from President/Chair of the State Board of Education); (2) the signed MOU is submitted to the Consortium Grant Project Manager (until June 23) or, after August 4, 2010, to the Project Management Partner; (3) the Advisory and Governing States agree and adhere to the requirements of the governance; (4) the chief procurement officer of the State seeking membership has reviewed the State's applicable procurement rules and provided assurance that it may participate in and make procurements through the Consortium; (5) the State is committed to identifying any existing barriers in State law, statute, regulation, or policy to implementing the proposed assessment system and to addressing any such barriers prior to full implementation of the summative assessment components of the system; and (6) the State agrees to support all decisions made by the Consortium. After the Consortium receives the grant award, any request for membership in the Consortium must be approved by the Consortium's Executive Committee. Upon approval, the Project Management Partner will then submit a change of membership to USED for approval. A State may begin participating in the decision-making process after receipt of the MOU.

Any State may exit the Consortium without cause, but must comply with the following exit process:

- A State requesting to exit the Consortium must submit in writing its request,
- The written explanation must include the statutory or policy reasons for the exit,
- The written request must be submitted to the Project Management Partner with the same signatures as required for the MOU,

- The Executive Committee will act upon the request within a week of receiving it, and
- Upon approval of the request, the Project Management Partner will then submit a change of membership to USED for approval.

A State may change its role within the Consortium from an Advisory State to a Governing State or from a Governing State to an Advisory State by submitting its requests and reasons in writing. The written request must be submitted to the Project Management Partner with the same signatures as required for the MOU. The Executive Committee will act upon the request within a week of receiving it and submit it to USED for approval.

As shown in the Summary Table for (A)(1)(b)(v), on October 1, 2010, the Consortium will initiate its support for all States in the Consortium to adopt the CCSS, with all States completing their adoption no later than December 31, 2011. Consistent with the Consortium's commitment to comparable and rigorous assessments, the Consortium will determine common performance level descriptors and achievement standards by August 31, 2014. SBAC will also create common assessment administration procedures and accommodations for English learners and accommodations for students with disabilities, respectively. Further, the Consortium will create common policies regarding item release, test security, definition of English learner, and participation for English learners and students with disabilities, respectively. The time line for this is presented in the table below.

Summary Table for (A)(1)(b)(v): Consortium’s Policy and Definition Time Line

Policy or Definition to be Adopted	Approx Date to Initiate	Approx Date to Adopt
State adoption of Common Core State Standards	October 1, 2010	December 31, 2011
Common set of performance level descriptors	October 1, 2010	August 31, 2014
Common set of achievement standards	October 1, 2010	August 31, 2014
Common assessment administration procedures	October 1, 2010	June 30, 2013
Common item release policy	October 1, 2010	June 30, 2013
Common test security policy	October 1, 2010	June 30, 2012
Common definition of “English learner”	October 1, 2010	June 30, 2012
Common policies and procedures for accommodations for English learners	October 1, 2010	June 30, 2012
Common policies and procedures for accommodations for students with disabilities	October 1, 2010	June 30, 2012
Common policies and procedures for student participation for English learners	October 1, 2010	June 30, 2013
Common policies and procedures for student participation for students with disabilities	October 1, 2010	June 30, 2013

The Consortium’s process and time line for setting other policies and definitions will be developed and determined by the Executive Committee and Steering Committee. The Committees and working groups will follow the decision-making process described in section (A)(1)(b)(iii).

(A)(1)(b)(vi) The Consortium’s plan for managing funds received under this grant category will be governed by the laws and rules of the State of Washington, as the Lead Procurement State, and in accordance with 34 CFR 80.36. Additionally, Washington is prepared to follow the guidelines for grant management associated with the American Recovery and Reinvestment Act of 2009 (ARRA), and it will be legally responsible for the use of grant funds and for ensuring that the Consortium carries out the project in accordance with federal requirements. Washington has already established an ARRA quarterly reporting system (also referred to as 1512 Reporting).

Per statute, Washington generates claims against grant funds based on qualifying reimbursable deliverables submitted on behalf of staff or clients, physical purchases, or contracted services. Washington's role as Lead Procurement State for the Consortium will ensure that monetary exchanges are executed against appropriate and qualifying reimbursable deliverables aligned to expenditure arrangements (i.e., contracts) made with vendors or contractors operating under "personal service contracts," whether individuals, private companies, government agencies, or educational institutions.

Washington is audited regularly by the State government, and has for the past five years been without an audit finding. As noted above, the State is prepared to meet the rigorous scrutiny associated with ARRA funding by leveraging its existing fiscal monitoring and control systems, including

- Washington's accounting practices, which are stipulated in the State Administrative and Accounting Manual (SAAM) managed by the State's Office of Financial Management. The SAAM (available upon request) provides details and administrative procedures required of all Washington State agencies, including the State's education agency, for the procurement of goods and services. Actions taken to manage the fiscal activities of the Consortium will, likewise, adhere to policies and procedures outlined in the SAAM.
- Comprehensive contracting rules to which Washington will adhere while serving as fiscal agent on behalf of the Consortium. These rules may be found in the Revised Code of Washington (RCW) 39.29 "Personal Service Contracts." Regulations and policies authorized by this RCW are established by the State's Office of Financial Management, and can be found in the SAAM.

(A)(1)(c)(i) The terms and conditions of the MOU are described throughout this section, (A)(1), and also can be found in Appendix A1-4. The MOU will become effective as of the date entered and will remain in force until the conclusion of the Program unless terminated earlier in writing by the Consortium.

(A)(1)(c)(ii) Each State agrees to identify existing barriers to implementing this proposed assessment system, in State laws, statutes, regulations, or policies, and to remove any such barriers before full implementation of the summative assessment components of this system.

Each State will use a common process to address barriers. The process includes identifying the issue or risk (if known); the statute, regulation, or policy related to the barrier; the governing body with the authority to remove the barrier; the approximate date to initiate action to remove the barrier; and a target date for removing the barrier. States may have chosen to include any known barriers in the Barriers Table at the time of signing the MOU (see MOUs in Appendix A1-4).

(A)(1)(d) The Consortium's procurement process will be dictated by the laws and rules of the State of Washington, as the Lead Procurement State. Prior to initiating contracts over \$20,000, bids must be acquired through a competitive solicitation process, defined as a formal process providing an equal and open opportunity to all qualified parties and culminating in a selection based on criteria delineated in the solicitation, which may include such factors as the consultant's fees or costs, ability, capacity, experience, reputation, responsiveness to time limitations, responsiveness to solicitation requirements, quality of previous performance, and compliance with statutes and rules relating to contracts or services.

The SAAM provides the following particular guidance: "Cost is always a factor but does not need to be the determinant factor. Agencies determine the weighted value of evaluation criteria and select a contractor and negotiate the contract based on these criteria." In sum, a Washington State agency has substantial discretion to build the criteria and scoring weights that it deems best to solicit the contract it wants. The limitations are intended to ensure that all responsible contractors are given equitable opportunity to win the solicitation. Each State's commitment to the Consortium's procurement process is evidenced by the signature blocks of each State's chief procurement official, provided in the MOUs in Appendix A1-4.

(A)(2) Theory of Action (up to 5 points)

The extent to which the eligible applicant’s theory of action is logical, coherent, and credible, and will result in improved student academic outcomes. In determining the extent to which the theory of action has these attributes, we will consider the description of, and rationale for—

- (a) Each component of the proposed assessment system and the relationship of the component to other components in the system;
- (b) How the assessment results produced by each component will be used;
- (c) How the assessments and assessment results will be incorporated into a coherent educational system (*i.e.*, a system that includes standards, assessments, curriculum, instruction, and professional development); and
- (d) How the educational system as a whole will improve student achievement and college- and career-readiness (as defined in the NIA).

The SMARTER Balanced Assessment Consortium (SBAC) supports the development and implementation of learning and assessment systems to radically reshape the education enterprise in participating States in order to improve student outcomes. SBAC believes that the current “drop from the sky” approaches to educational testing are ineffective for too many teachers and students. Through expanded use of technology and targeted professional development, the Consortium’s Theory of Action calls for full integration of the learning and assessment systems, leading to more informed decision-making and higher-quality instruction, and ultimately to increased numbers of students who are well prepared for college and careers.

The overarching goal of the SBAC is ***to ensure that all students leave high school prepared for postsecondary success in college or a career through increased student learning and improved teaching.*** Our approach is rooted in the belief that stronger learning will result from high-quality assessments that support ongoing improvements in instruction and learning, and that are educative for students, parents, teachers, school administrators, members of the public, and policymakers. Meeting this goal will require the reform and coordination of many elements across the education system, including, but not limited to, a quality assessment system that strategically “balances” summative, interim, and formative components (Darling-Hammond & Pecheone, 2010); provides valid measurement across the full range of common rigorous academic standards, including assessment of deep disciplinary understanding and higher-order

thinking skills that are increasingly demanded by a knowledge-based economy; and by the establishment of clear, internationally benchmarked performance expectations. Other elements that are outside SBAC’s direct scope of work, but not outside its influence, are comprehensive pre-service and in-service professional development and focused and valid systems of accountability.

Seven Principles Undergirding the Theory of Action

Our assessment proposal is shaped by a set of seven principles shared by both assessment systems in high-achieving nations and a number of high-achieving States in the U.S.

- (1) Assessments are grounded in a thoughtful, standards-based curriculum and are managed as part of an integrated system** of standards, curriculum, assessment, instruction, and teacher development. Curriculum and assessments are organized around a well-defined set of learning progressions along multiple dimensions within subject areas. Formative and interim/benchmark assessments and instructional supports are conceptualized in tandem with summative assessments—all of them linked to the standards and supported by a unified technology platform.
- (2) Assessments produce evidence of student performance** on challenging tasks that evaluate the Common Core State Standards (CCSS). Instruction and assessments seek to teach and evaluate knowledge and skills that generalize and can transfer to higher education and multiple work domains. They emphasize deep knowledge of core concepts and ideas within and across the disciplines—along with analysis, synthesis, problem solving, communication, and critical thinking—thereby requiring a focus on complex performances as well as on specific concepts, facts, and skills.
- (3) Teachers are integrally involved in the development and scoring of assessments.** While many assessment components are efficiently scored with computer assistance, teachers must also be involved in the formative and summative assessment systems so that they deeply understand and can teach in a manner that is consistent with the full intent of the standards, while becoming more skilled in their own assessment practices.
- (4) The development and implementation of the assessment system is a State-led effort with a transparent and inclusive governance structure.** Since December 2009, SBAC

has hosted weekly conference calls and several face-to-face meetings open to all States interested in establishing a Consortium of States for the development of assessments aligned to the CCSS. Those activities have resulted in a governance structure that has established a consensus decision-making model and clear leadership roles. Each State's commitment to our collaborative process and products will facilitate the development of our complex system and signal ongoing support for its implementation.

(5) Assessments are structured to continuously improve teaching and learning.

Assessment *as, of, and for* learning is designed to develop understanding of what learning standards are, what high-quality work looks like, what growth is occurring, and what is needed for student learning.

(6) Assessment, reporting, and accountability systems provide useful information on multiple measures that is educative for all stakeholders.

Reporting of assessment results is timely and meaningful—offering specific information about areas of performance so that teachers can follow up with targeted instruction, students can better target their own efforts, and administrators and policymakers can more fully understand what students know and can do, in order to guide curriculum and professional development decisions.

(7) Design and implementation strategies adhere to established professional standards.

The development of an integrated, balanced assessment system is an enormous undertaking, requiring commitment to established quality standards in order for the system to be credible, fair, and technically sound. SBAC is committed to developing an assessment system that meets all Critical Elements required by USED Peer Review, relying heavily on the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999) as its core resource for quality design. Other key sources of professional standards that will guide SBAC's work include a reasoning-from-evidence approach (e.g., see NRC, 2001; Mislevy, Almond, & Lukas, 2004); *Operational Best Practices in Large Scale Assessment* (ATP, CCSSO, in press); and the ANSI-endorsed *Student Evaluation Standards, Program Evaluation Standards, and Personnel Evaluation Standards* (JCSEE, 2002, 1994, 2008, respectively).

Components of the Theory of Action

Presented below are the components of the Consortium’s Theory of Action, including connections to other system components, the results to be produced, and some of the key related Consortium activities. A pictorial schematic of the SBAC Theory of Action is found in Appendix A2-1. While this figure presents the Theory of Action in a somewhat linear fashion, this is simply a limitation of representing a complex system in two dimensions and on a single page. The actual Theory of Action is much more recursive and multidimensional than graphically depicted.

Consortium and State policies and practices support high expectations and increased learning opportunities for students.

A major working assumption of the Consortium is that assessment reform must operate within the context of State policies and practices that can either support or hinder realization of the overall goal to have students graduate from high school as college- and career-ready. Thus, SBAC has committed to creating a policy environment that can support the innovative systems described in the design section of this proposal. Supportive policies would include the development of accountability systems that incentivize the right behaviors for administrators and teachers, and avoid inadvertently rewarding behaviors that would run counter to the learning goals. Another example is policy for provision of ongoing professional development structures and support for teachers.

The assessment system is aligned to a common set of State standards that clearly specify college, career, and grade-level expectations.

A State policy that is fundamental to SBAC’s Theory of Action is adoption of the Common Core State Standards (CCSS), which clearly specify college and career expectations as well as the knowledge and skills required at each grade level to meaningfully articulate progress toward these end-of-high-school expectations. These “fewer, higher, and deeper” standards—influenced by findings that high-achieving countries typically teach fewer topics more deeply—will serve as the basis for the comprehensive assessment system. And while it is critical that the assessment system validly reflects these standards, SBAC must interpret or translate these standards before they can be used effectively for assessment or instruction. Specific steps include the following.

1. Ensure that each member State adopts the CCSS by December 31, 2011.

2. Translate the standards into content/curricular frameworks, test maps, and item/performance event specifications to provide assessment specificity and to clarify the connections between instructional processes and assessment outcomes.

SBAC policies and standards are effectively communicated to districts and schools.

Enacting policies and having standards is not enough. A major lesson learned by SBAC member States is that clear and timely communication of policies and practices is essential for successful implementation of a new system. Effective communication is critical in the short term to signal change, and over the longer term to implement change. Specific steps include the following:

1. Develop a multimedia communications plan that is implemented by each member State to educate stakeholders about key aspects of college and career expectations.
2. Develop score reports that clearly communicate about the assessment system and the results to key stakeholder groups.

Teachers are provided with curriculum, instructional materials, rich professional development, and other supports and resources to effectively instruct students on the standards.

While effective communication with teachers is essential, the SBAC model calls for a fuller level of teacher engagement in an integrated learning and assessment system, which requires that teachers receive adequate supports and resources. This system component, central to the design of the SBAC system, encompasses many different teacher support features. Specific aspects include

1. Model curriculum and instructional modules that are aligned with the CCSS.
2. Training modules that help teachers focus their instruction on the CCSS and develop teaching practices that support more in-depth learning.
3. Training of teachers to use formative assessment tools and interim/benchmark assessments as well as to interpret results and use those results to determine next steps in instruction.
4. Teacher-moderated scoring of performance events as a professional development vehicle

to enhance teacher capacity to evaluate student work aligned to the standards.

5. Online interpretable score reports at the student and classroom level that clearly show strengths and weaknesses and can be tailored to fit individual needs and circumstances.

Technology provides increased access and opportunities for students to fully engage in the learning and assessment systems and supports the design, delivery, scoring, and reporting of the assessment system.

Innovative and efficient use of technology is the hallmark of the SBAC model. SBAC's Theory of Action posits that technology solutions for test delivery will provide students with increased access to the assessment and will yield more accurate measurement of their acquisition of knowledge and skills. For example, use of computer adaptive testing (CAT) methodologies will ensure that students across the full range of performance have an assessment experience that presents them with items that are best suited to their skill level. Average-, very low-, and very high-performing students will be more likely to stay engaged in the assessment because they will be responding to questions targeted to their skill level.

The computer delivery system broadens the availability of the accommodations while establishing a less restrictive testing environment for students with special needs. The system will also support several formalized accommodations. For example, text-to-speech and aural native language translations can be supported if students are tested in isolation, or if they have access to headphones. Refreshable Braille can also be supported with online tests.

Just as technology will support student access and engagement, it will also lead to more valid and timely reporting of assessment results, and lead to efficiencies and enhancements for professional development and resource tools. Specifically, SBAC will

1. Ensure that all students are provided with the technology needed for all aspects of the SBAC assessment system (summative, interim/benchmark, and formative).
2. Support research on how best to increase *access* for all students through the use of technology.
3. Use technology to efficiently deliver training programs, resources, score reports, data, etc., including interactive Web-based social networks designed for teacher use in the development and dissemination of effective curriculum and instructional practices.

4. Create innovative and real-world item types that rely on technology platforms.
5. Use adaptive item selection engines, drawing on a broad item pool, to ensure that accurate measures of student achievement are possible across a wide performance continuum without undue burden.
6. Establish accommodation protocols that capitalize on technological capabilities to support broader access to assessments for all students, including those most at risk.
7. Standardize member State accommodation policies through a coordinated Enhanced Assessment Grant.

A high-quality summative assessment system establishes high expectations and provides relevant information on achievement and growth to teachers, students, and others.

Assessments must be carefully structured to improve teaching and learning. This means establishing summative assessments that reflect the challenging CCSS content, emphasizing not just students' "knowing," but also "doing." SBAC envisions a summative assessment system composed of interactive selected-response and constructed-response items and simulations as well as teacher-developed performance events that measure the full range of student abilities on the CCSS. The incorporation of CAT is based on member States' positive experiences with this methodology (e.g., Oregon) and the many benefits it affords, such as precision of measurement and timely results (Kosty, McBride, Poggio, Wise, & Way, 2006; Lilley, Barker, & Britton, 2004; Rabinowitz, 2005). The summative assessment will accomplish the following:

1. Signal high expectations to students, parents, teachers, administrators, and policymakers.
2. Provide efficient, reliable, and valid information across the full range of achievement.
3. Engage IHEs at the high school level to ensure that assessments truly reflect a measure of readiness for college and careers.
4. Provide explicit measures of student progress toward college- and career-readiness through growth models and criterion-validity studies.
5. Promote policy alignment by establishing internationally benchmarked achievement standards that are common across Consortium States and that are comparable across multiple consortia.

Interim/benchmark (I/B) assessments and formative tools and strategies are integrated with the summative assessments to provide instructionally useful information to teachers,

students, and administrators.

While a rigorous summative assessment is essential, SBAC believes that it is insufficient to drive positive change in teaching and learning. Informed by the recent experiences in England and Hong Kong, SBAC posits that I/B and formative assessments are the other necessary assessment ingredients to drive teaching and learning (Darling-Hammond & Pechone, 2010). As such, I/B and formative assessments will be developed and implemented directly under the purview of the Consortium—not simply adopted from external sources. Grounded in cognitive development theory about how learning progresses across grades and competence develops over time (NRC, 2001; Pellegrino, 2006; Stiggins, 2002), the assessments will (a) work in concert with the summative assessment, (b) allow for more innovative and fine-grained measurement of student progress toward the CCSS (Shepard, Hammerness, Darling-Hammond, & Rust, 2005), and (c) provide diagnostic information that can help tailor instruction and guide students in their own learning efforts. Besides its close connection to the summative component, this component will also operate in tandem with the teacher resources and supports component as well as the teacher engagement component (see below). The main features that SBAC will incorporate into its comprehensive system include

1. I/B assessments on the same scale as the summative assessments to measure within-year student achievement and provide teachers and students with information on the degree to which students are on track to succeeding on the summative assessments.
2. Interpretative guides, using the publicly released I/B assessment items and performance events to illustrate how the SBAC assessments are manifestations of the CCSS.
3. Formative tools that teachers can use throughout the year to better understand where students are in their learning and determine any misconceptions, allowing for quick adjustment to instruction as well as differentiated instruction.

Teachers are engaged in the design, development, and scoring of assessment items and in the reporting of results.

The SBAC model envisages an integral role for teachers in an integrated learning and assessment system. This means teachers must be meaningfully engaged in all aspects of assessment. To that end, the SBAC model incorporates the following features:

1. Work with teachers and policy stakeholders to develop test maps that assess the full range

of the CCSS and that articulate within and across grade levels.

2. Involve teachers in specifying, writing, reviewing, and range finding test items/ performance events.
3. Use teacher-moderated scoring of performance events as a professional development vehicle to enhance teacher capacity to evaluate student work aligned to the standards.

Teachers, students, and administrators use information from instructionally useful assessments to improve teaching and learning.

Information from assessment results must be delivered in ways that are instructionally useful for schools and teachers as well as meaningful and actionable for students (Popham, 2006).

Making optimal use of technology, SBAC will

1. Fully involve teachers (and other end-users) in designing different score reports and web-enabled tools and services to maximize their communication value and usefulness.
2. Provide interactive reports and resources so that teachers fully understand performance for each student and the class as a whole.
3. Allow students to more fully engage in the learning process through ongoing interim/benchmark assessments that can be self-administered and reports that allow students to compare where they are to where they need to be.

In summary, the proposed SBAC learning and assessment system is grounded in a sound Theory of Action—taking advantage of current research and lessons from current practice—and incorporates a new generation of technology tools, innovative assessments, and state-of-the-art classroom support mechanisms to improve teacher and student capacity to meet the challenges in ensuring that all students are college- and career-ready.

(A)(3) Assessment System Design (up to 55 points)

The extent to which the design of the eligible applicant's proposed assessment system is innovative, feasible, and consistent with the theory of action. In determining the extent to which the design has these attributes, we will consider—

(a) The number and types of components (*e.g.*, through-course summative assessments (as defined in the NIA), end-of-year summative assessments, formative assessments, interim assessments in mathematics and in English language arts in the assessment system);

(b) For the assessment system as a whole—

(i) How the assessment system will measure student knowledge and skills against the full range of the college- and career-ready standards, including the standards against which student achievement has traditionally been difficult to measure; and provide an accurate measure of student achievement, including for high- and low-performing students, and an accurate measure of student growth over a full academic year or course;

(ii) How the assessment system will produce the required student performance data (*i.e.*, student achievement data and student growth data (both as defined in the NIA) that can be used to determine whether individual students are college- and career-ready (as defined in the NIA) or on track to being college- and career-ready (as defined in the NIA);

(iii) How the assessment system will be accessible to all students, including English learners and students with disabilities, and include appropriate accommodations (as defined in the NIA) for students with disabilities and English learners; and

(iv) How and when during the academic year different types of student data will be available to inform and guide instruction, interventions, and professional development; and

(c) For each component in mathematics and in English language arts in the assessment system--

(i) The types of data produced by the component, including student achievement data (as defined in the NIA), student growth data (as defined in the NIA), and other data;

(ii) The uses of the data produced by the component, including determining whether individual students are college- and career-ready (as defined in the NIA) or on track to being college- and career-ready (as defined in the NIA); informing determinations of school effectiveness for the purposes of accountability under Title I of the ESEA; informing determinations of individual principal and teacher effectiveness for the purposes of evaluation; informing determinations of principal and teacher professional development and support needs; informing teaching, learning, and program improvement; and other uses;

(iii) The frequency and timing of administration of the component, and the rationale for these;

(iv) The number and types of items (*e.g.*, performance events, selected response items, brief or extended constructed response items) and the distribution of item types within the component, including the extent to which the items will be varied and elicit complex

student demonstrations or applications of knowledge and skills (descriptions should include a concrete example of each item type proposed); and the rationale for using these item types and their distributions;

(v) The component’s administration mode (*e.g.*, paper-and-pencil, computer-based, or other electronic device), and the rationale for the mode;

(vi) The methods for scoring student performance on the component, the estimated turnaround times for scoring, and the rationale for these; and

(vii) The reports produced based on the component, and for each report, its intended use, target audience (*e.g.*, students, parents, teachers, administrators, policymakers), and the key data it presents.

Required Tables and/or Attachments:

✓ *Summary Table for (A)(3): English Language Arts*

✓ *Summary Table for (A)(3): Mathematics*

✓ *Examples of each item type proposed*

(A)(3)(a) The SMARTER Balanced Assessment Consortium is committed to developing an assessment system that purposefully balances summative, interim/benchmark (I/B),² and formative components and uses the information available from each component in a manner consistent with its design and purposes. We believe that this balance of assessment components will provide for a fully integrated system of learning and assessment that yields the necessary and actionable information to support quality instruction. The system will further support quality instruction by describing the knowledge and skills required for college- and career-readiness as laid out in the Common Core State Standards (CCSS)—an elaboration that will result from our work in developing an assessment system that measures deep disciplinary understanding and higher-order thinking skills. Our system of summative, I/B, and formative assessments will produce instructionally useful information available throughout the instructional year to help guide and support differentiated instruction. Teachers will be supported in the effective use of these assessment data through a well-developed system of professional development and professional learning groups—vital components of the overall design.

We provide an overview of our proposed assessment system in Appendix A3-1. Our assessment system capitalizes and expands on existing technologies to provide accurate measurement of individual student achievement and growth in student learning based on the CCSS. We are committed to a computer adaptive summative assessment for English language arts

² The I/B and formative components are described in the MOU as optional “formative/benchmark components,” which is a broader description for these system features.

(ELA) and mathematics to be administered in grades 3–8 and grade 11 within 12 weeks of the end of the instructional year. The adaptive summative assessment includes selected-response, technology-enhanced constructed-response, and extended constructed-response items. The computer adaptive approach provides maximally accurate assessment for each student to better determine whether students are on track for being college- and career-ready and to support the measurement of the full range of student abilities. The summative evaluation for each student will also include performance events that will provide a measure of the student’s ability to integrate knowledge and skills across multiple standards—a key component of college- and career-readiness. Performance events will be used to better measure capacities such as depth of understanding, research skills, and complex analysis, which cannot be adequately assessed with selected- or constructed-response items. At grades 3–8, students will engage in two rigorous performance events for ELA and mathematics. At the high school level, students will engage in up to six performance events by grade 11 for both ELA and mathematics. These events will be computer-delivered and typically will require one to two class periods to complete.

The second key component of SBAC’s assessment system is adaptive I/B assessments, built around learning progressions. Learning progressions describe “how learning typically unfolds” (Darling-Hammond & Pecheone, 2010, p. 9). Results from these assessments can be used by teachers to develop targeted instructional strategies aimed at addressing specific gaps in understanding to see where each student currently is situated on the learning continuum toward college- and career-readiness. These I/B assessments play a key role in our system because they provide more finely grained information about student progress toward college- and career-readiness than can be provided by the summative assessment. They can be administered at various points within the instructional year to support effective instructional decisions at the student level. This pool of items will contain the same types of items and performance events as the summative pool and will include all released summative items and events.

The third component of the Consortium’s assessment system is a set of formative tools, processes, and practices that support the needs of teachers on a daily basis as they help students learn and progress on their path toward college- and career-readiness. The goal of our formative assessment component is to enable full implementation of the learning and assessment system by helping teachers and administrators effectively use data from the summative and I/B assessments and to build their capacity to collect evidence during instruction that can be useful in diagnosing

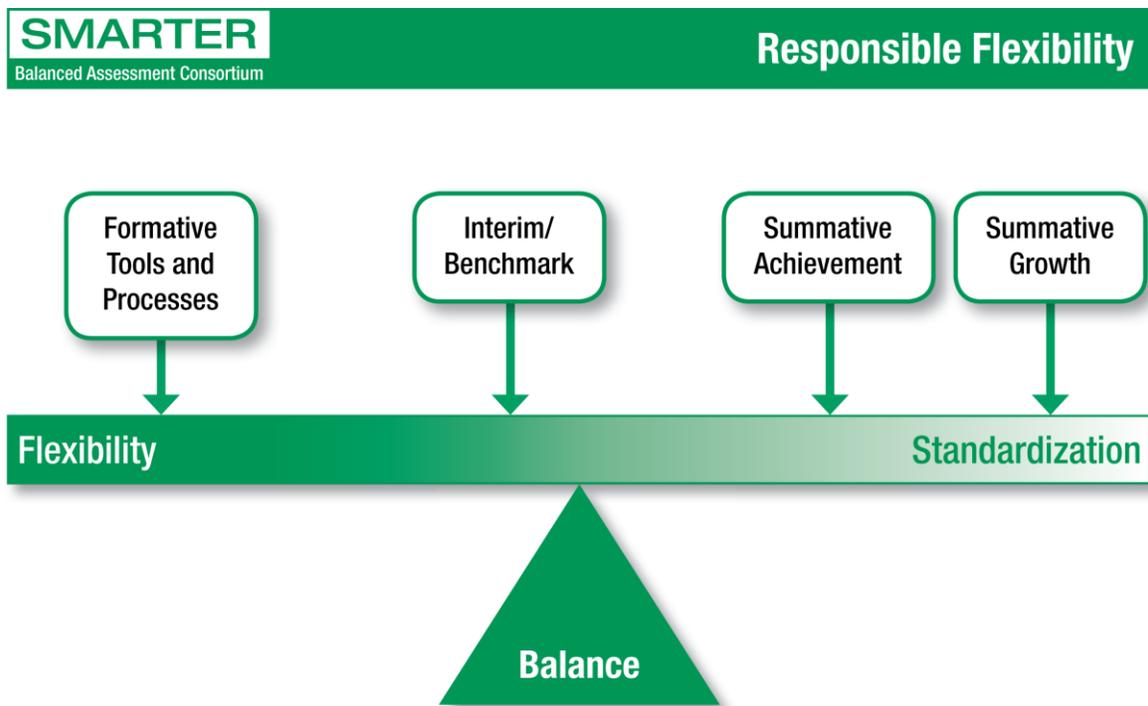
students' learning needs. To build these capacities, the Consortium will develop and disseminate research-supported tools, resources, and materials. It also will provide opportunities for training and collaboration through professional learning work groups to support the development of assessment literacy; an understanding of the content and performance expectations of the CCSS; development of model curriculum units; alignment of curriculum, instruction, and assessment; development of teacher capacity in the reasoning-from-evidence approach to item development supported by the Consortium; an understanding of the expectations of student performance through the effective use of scoring activities and released events with annotated student work; and an in-depth understanding of the learning process to sustain improved instructional practice. By combining these formative practices and tools with the summative and I/B assessments, we are developing a system for learning and assessment that will lead to more informed decision-making and will result in higher-quality instruction, and thus, higher levels of student achievement.

Further, SBAC is committed to developing an assessment system that reflects the principle of responsible flexibility. We want to provide States within the Consortium flexibility in implementing the components of our assessment system in a way that best meets their needs. Consequently, the Consortium is committed to a comprehensive research agenda to ensure that options for administration of the summative assessment adhere to three key principles—comparability, technical adequacy, and fairness. For example, the Consortium will be investigating the reliability and validity of offering States an optional distributed summative assessment as an alternative for States to the administration of the summative assessment within the fixed 12-week testing window. The distributed summative assessment will be developed based on content clusters to allow students to demonstrate mastery of content and skills throughout a course and at the most appropriate time for each student. The scores of these distributed assessments would be rolled up (along with the students' scores on the performance events) to make the overall decision about students' achievement with respect to college- and career-readiness. We believe that this “through course” option reflects the greatest flexibility for States and allows for differentiated instruction and assessment that we believe reflects the best in instructional and assessment practice. We also recognize the need to fully investigate this option to ensure that the results based on a distributed assessment are comparable to those based on the comprehensive summative assessment for all required purposes—student achievement and growth, school accountability, and, as a part of teacher and principal evaluation—meet the requirements for technical adequacy

and fairness.

SBAC recognizes that the purpose and intended use (with associated stakes) of each component play a large part in how flexibility must be weighed in relation to comparability and technical adequacy. Figure A3-1 demonstrates how we view the *balance* among these sometimes conflicting values. On the ends of this continuum, formative assessment tools must allow for classroom flexibility to be successful while the technical requirements of growth models for student and school accountability create strong limits on this dimension. I/B assessments and summative assessments used to measure student achievement fall in between. Ultimately, our research agenda (described in section (A)(5)) will determine how to weigh the various values explicit in our Theory of Action used to drive the development and implementation of our balanced assessment system.

Figure A3-1



(A)(3)(b)(i) A coherent, comprehensive assessment system consists of a set of strategically selected measures to assess students' knowledge and mastery of college- and career-readiness standards. The measures used within this system were selected based on their appropriateness for the construct to be assessed, and they will fit together as a set to evaluate the full range of

standards. Our assessment system will include computer adaptive selected-response items, technology-enhanced constructed-response items, and extended constructed-response items as well as standardized performance events at each grade level (3–8 and high school). The commitment of the Consortium to the use of performance events will be central to our ability to measure student knowledge and skills against the full range of the college- and career-ready standards. The performance components will reflect more ambitious events that can measure aspects of student performance that are reflected in the CCSS but have traditionally been difficult to measure on standardized assessments, including skills such as the use of relevant evidence and technology, thoughtful critique, and adaptive reasoning. Student performance on the adaptive summative assessment will be combined with the performance measures to provide the accurate measure of student achievement as required by the NIA.

SBAC is committed to a computer adaptive model because it represents a unique opportunity to create a large-scale assessment system that provides maximally accurate achievement results for each student. The adaptive assessments will be highly sensitive to the unique status of the learner and will sample content above and below grade level, as needed, to ensure the accurate assessment of an individual student’s progress toward meeting the expectation for college- and career-readiness. Recent authors have noted that computerized adaptive assessment can be particularly effective in measuring an individual student’s growth over time and is a model that is compatible with the vertically articulated content standards as reflected in the CCSS (Way et al., 2010).

The system is based on the annual administration of the summative assessment at grades 3–8 and 11 within a 12-week window at the close of the instructional year. These summative scores (including the results on the performance events) will provide for a common measure on which a stable measure of achievement and growth can be determined for each student. For States that wish to measure growth from grade 8 prior to grade 11, an adaptive summative assessment in grades 9 or 10 will be available.

(A)(3)(b)(ii) An important measure of the overall validity of the SBAC assessment system will be the extent to which summative results for each content area accurately measure whether students are on track or ready for college or career.

The Consortium’s adaptive summative assessment in combination with performance events will operationalize this expectation, and our research studies will validate their effectiveness. Our

evidence will cover the full range of a validity argument, drawing upon the Theory of Action as the foundation for the system's interpretive argument (Kane, 2007). We begin at the content level. Our proposed blueprint provides sufficient data across the clusters of the CCSS to measure achievement (i.e., obtained proficiency level) and growth (i.e., both progress toward meeting grade-level expectations and progress toward the grade 12 exit criteria). We will conduct independent alignment studies to ensure all items and events properly and fully assess the intended content.

Next, the technical analyses of our summative assessment will examine its ability to provide student achievement and growth data to measure college- and career-readiness. For example, the adaptive engine will sample items within grade level and above or below as necessary to provide precise measurement of the student's achievement level. Results of assessments, as translated by the vertically articulated content and achievement standards, will be expressed on the same common scale. Each student can be compared to the same "on-grade" standards regardless of whether or not off-grade items were used in his or her test.

Finally, we will conduct external validity studies to measure whether students who achieve mastery at a grade level (achievement score) or who are considered on a trajectory toward mastery (growth score) do indeed achieve that predicted level at the next grade and, eventually, by grade 12 and beyond.

Following the full system field testing in spring 2014, the Consortium will conduct standard setting for the adaptive summative assessments in grades 3–8 and grade 11 in ELA and mathematics. The method for standard setting will be determined by the Governing States in the Consortium in school year 2012–13, with an understanding that (a) standard setting for computer adaptive tests (CATs) and (b) use of a composite score that includes performance events will require that special issues be addressed prior to model selection. Key to the discussion is whether we will use an item-based approach such as the *Bookmark* method (Mitzel, Lewis, Patz, & Green, 2001) or a more holistic approach where actual student work samples are reviewed (e.g., *Body of Work*, Kingston, Kahl, Sweeney, & Bay, 2001). Neither approach alone is ideal for a system such as ours that features selected-response items, constructed-response items, and performance events. Regardless, the selected standard-setting process will be used to translate the CCSS into the performance standards that will be reported for Title I accountability purposes. Specifically, the purpose of this process is to set cut-scores (benchmarks) for:

- different levels of achievement (e.g., below basic, basic, on track to college- and career-ready, and advanced) for each of grades 3 through 8, and
- different levels of achievement (e.g., below basic, basic, college- and career-ready, and advanced) for grade 11.

Prior to or during standard setting, a set of achievement level descriptors (ALDs) for each achievement level in each grade (3–8 and 11) will be developed by a team of representatives from the Governing and Advisory States that will include teachers and representatives from IHEs and the workplace. The ALDs will serve as the framework that connects the content in the CCSS with summative item-pool/performance event content by describing what a student must know and be able to do at each achievement level.

Cut-scores set during August 2014 standard setting will be validated in July 2015 using spring 2015 operational data. The judges/participants will include local education agency (LEA) and state education agency (SEA) representatives who will review all evidence that the cut-scores set the previous year are reasonable and appropriate, given the actual findings from operational testing. The validated cut-scores will be presented to SBAC’s Steering Committee and Governing States for approval, then to the Total State Membership for review and approval in August 2015. The Consortium agrees to adopt common achievement standards across all Governing and Advisory States by the end of August 2015.

Given the desire and value of comparable standards for all U.S. students, SBAC will coordinate standard setting with any other consortia funded under this competitive process. This will include the development of comparable achievement level descriptors to guide the standard-setting process and linking studies that will ensure the ability to translate results (and accountability determinations) across the nation, regardless of to which Consortium a State belongs.

(A)(3)(b)(iii) The Consortium is deeply committed to ensuring that the intellectual integrity and full rigor of the academic content standards are maintained throughout all forms of assessments developed to serve this system and that all students, regardless of disability, language, or subgroup status, have adequate access to the content in the SBAC assessment system. Critical to supporting student achievement of the academic content to which the assessment system is aligned is appropriate student access to the grade-level content in the system. The SBAC system will ensure

meaningful accessibility to students; that is, valid forms of presentation, engagement, and response for students—to address access barriers including vision, hearing, motor, and other sensory, physical, cognitive, processing, and language needs of students—that enable students to fully demonstrate what they know and can do.

Generally, meaningful accessibility will be grounded in the upfront determination of the following three factors: the *targeted constructs* (a clear definition of the specific grade-level content skills and knowledge that the assessment is intended to measure); *unintended constructs* (identification and consideration of the possible influence of construct-irrelevant variables, such as nonessential complexity in text and visuals, and what the assessment is not intended to measure), and *student access profiles* (the particular sensory, physical, cognitive, processing, and/or language needs of students that can be addressed during testing in order to minimize construct irrelevance or the influence of unintended constructs). The upfront determination and careful consideration of these key factors throughout the assessment design and development process will support informed decisions that best ensure the purposeful design and systematic embedding of accessibility features in the system’s assessments, and that support student access to targeted grade-level constructs; do not alter the measure of the intended construct; improve the measure of the intended construct for all students, particularly students with disabilities and English learners; and support valid inferences of what students know and can do (e.g., Mislavy, Steinberg, & Almond, 2003; Almond et al., 2010). The SBAC assessment system will rely on experts from relevant fields (e.g., measurement, applied linguistics, special education, curriculum and instruction, technology) and draw upon the most recent research in areas that include cognitive load theory, readability, content relevance and difficulty, principles of universal design for assessment, and fairness for enhancing student access to targeted grade-level constructs and obtaining information that support valid inferences about what all students know and can do.

The general framework for ensuring meaningful access for all students engaging in the Consortium’s system will be informed by the Access by Design Model (see Appendix A3-2: *Access by Design: A Theory of Action and Research* [Fedorchak, 2010]). Consistent with the upfront consideration of key factors described previously, the Access by Design Model provides a framework that supports a systematic approach for purposefully engineering accessibility features into all assessment design blueprints and structures from the beginning and throughout the assessment development process. By doing so, the vast majority of student access needs (e.g.,

cognitive, processing, sensory, physical, language) are met by accessibility features (e.g., magnification, audio representation of graphic elements, linguistic simplification) designed into the structure and delivery of the assessment items and formats. This is carried out in a manner that ensures the integrity of the targeted grade-level constructs being measured is maintained. Additional “accommodations” requiring special documentation and procedures, or additional local resources for implementation, are thereby minimized.

Within the Access by Design framework, meaningful accessibility will be ensured according to the principles as outlined in Appendix A3-2. The principles focus in particular on access vis-à-vis the technology approach central to the SBAC assessment system.

(A)(3)(b)(iv) The balanced assessment system proposed by the Consortium will provide student-level data throughout the instructional year to inform and guide instruction, interventions, and professional development. A primary purpose of the assessment system (along with accountability) is to provide timely feedback to students, teachers, and principals. Whereas the summative assessment will serve as the primary measure to determine achievement and growth for accountability, the I/B assessments and formative tools and processes are designed to provide the meaningful, actionable data that can immediately impact instruction.

The computer-administered summative and I/B assessments can provide immediate results based on the selected-response and specific technology-enhanced item types that can be scored via the automated scoring feature built into the testing platform. The scale scores based on these items will be updated as results from the artificial intelligence (AI) scored items and performance events are available. Further, we anticipate that teacher scoring will be required for select dimensions of the performance events (e.g., inference) that may not be scored as reliably through AI technology. Once operational, we anticipate a two-week window for turnaround of items requiring teacher scoring.

We endorse the view of formative assessment as a means to collect evidence *during* the course of instruction about how students are learning so that instructional changes can be made to support student learning (Heritage, 2007). The Consortium will support the development of formative assessment tools that will be an integral part of day-to-day instruction and that will describe a student’s academic strengths and limitations, estimate the group’s knowledge level prior to instruction, check for misconceptions, provide evidence of progress toward learning goals, support

the evaluation of the effectiveness of instructional strategy, and identify areas for professional development.

(A)(3)(c)(i) The summative assessments in mathematics and English language arts will provide information about student progress toward college- and career-readiness in grades 3–8 and 11. This information will answer two important questions: (1) Has the student met or exceeded the achievement level for his/her current grade? and (2) Will the student meet the college-readiness achievement level by grade 12 if he/she continues on the current achievement trajectory? To accomplish both of these crucial accountability requirements, the Consortium will build a vertical scale and define vertically and horizontally articulated achievement standards in both ELA and mathematics. (A description of the procedure used to develop the vertical scale can be found in section (A)(5).) To assist in the interpretation of scores, results from the summative assessment will be reported as listed in the Summary Tables for section (A)(3). Where appropriate, information such as the scale scores and student progress will be able to be aggregated at the classroom, school, district, State, and Consortium levels. Additional data on student achievement and growth will be available thanks to our commitment to a truly “balanced” assessment system. Thus, the I/B component will allow more finely grained analyses of each student’s progress toward grade-level and college-readiness mastery. Our piloting plan for the I/B assessments will allow results to be placed on the same vertical scale as the summative assessment, greatly increasing teachers’ ability to transfer information from one to the other.

(A)(3)(c)(ii) We will produce data to determine whether every student is college- and career-ready or on track to being college- and career-ready by building a vertical scale across the 3–11 grade span and by developing vertically and horizontally articulated achievement standards.

Our Consortium anticipates that the model for determining school effectiveness will change from the current limited AYP model once ESEA is reauthorized. We also look forward to guidance from Congress and USED, via the reauthorization process, on the range of available options for using assessment data for “informing determinations of individual principal and teacher effectiveness for the purposes of evaluation; informing determinations of principal and teacher professional development and support needs; informing teaching, learning, and program improvement; and other uses.” Regardless of the model or models authorized, the Consortium’s

“Balanced System” approach is especially well suited for these purposes for the following reasons:

- Our summative assessment will be sufficiently reliable, valid, fair, and efficient to support high-stakes student, school, and teacher/principal accountability decisions;
- Our I/B assessments and formative support processes will pinpoint student needs and provide teachers and principals with the data they need to build curriculum and tailor instruction at the individual level;
- Our professional development model will provide teachers and principals the supports and strategies to implement data-driven reform;
- Our involvement of teachers and other education practitioners at all stages of development, implementation, and evaluation will create buy-in such that accountability decisions at all levels (student, teacher, school) will be seen as being as fair and equitable as possible; and
- Our reliance on ANSI-approved standards developed by the Joint Committee on Standards for Educational Evaluation in the areas of evaluating students, educational personnel, and programs (JCSEE, 2009).

We have conceptualized our score scale and achievement levels as broadly as possible so that they can accommodate a range of possible Federal and State accountability models. Our research agenda focuses on ensuring the technical adequacy of our data across these possible models and will ensure that comparability is achieved regardless of the model(s) selected.

(A)(3)(c)(iii) The assessment system has three major components that serve different primary purposes within the system. Thus, the frequency and time of administration vary by component. The adaptive summative component will be administered within the last 12 weeks of the instructional calendar. Students would be allowed two testing opportunities, and the student’s highest score would be used to determine achievement and growth. (The second testing opportunity is to provide students a second administration if there is evidence to suggest that the score associated with their first administration is invalid.) Performance events will be administered within an established testing window. These fixed administration windows are being implemented to facilitate comparability while providing States an adequate window to allow for student access to the computers required for administration. We will continue with this defined testing window until our research supports comparability decisions for assessments administered outside through the distributed model.

The I/B assessments allow for flexible administration at the discretion of the State, district, school, or teacher. Because these assessments will be tied to specific targeted knowledge and skills within a learning progression, the I/B item pool can be used to provide valuable information about student progress on the identified learning progressions before, during, or near the end of instruction.

We envision teachers implementing formative processes as *part of their instructional routine* (Resnick & Berger, 2010). As such, the frequency and timing would be throughout the instructional year and as part of daily instruction.

(A)(3)(c)(iv) The Summary Tables for (A)(3) outline the numbers and types of items proposed for the summative assessment. These numbers are based on the preliminary review of the CCSS and will be refined based on the final assessment framework and blueprint as outlined in section (A)(4)(a). The adaptive summative assessment will include selected-response items, technology-enhanced items, and extended constructed-response items and will be combined with the performance events to provide for the assessment of declarative and procedural knowledge along with the extended problem solving, inquiry, and synthesis expected for college- and career-readiness. The proposed distribution of item types is based on the number of CCSS, the preliminary analysis of their content, and the need to produce a scale that will include enough score points to reliably measure student abilities. Sample items for each item type included in the assessment system are presented in Appendix A3-3.

Whereas selected-response items have routinely been a part of assessments of student learning, their emphasis has too often been on low-level skills reflecting knowledge of discrete pieces of information that are not critical for subsequent student learning. The emphasis here will be on the development of items that reflect important knowledge and skills consistent with the expectations of the CCSS. Items can be developed to address knowledge and skills from more than one standard. The appropriate and judicious use of selected-response items provides for a cost-effective means to address content in terms of test development, administration, and scoring.

The Consortium is committed to making effective use of technology by including items that capitalize on the capabilities of the assessment platform. The effective use of technology can expand not only the nature of the content that can be presented but also the knowledge, skills, and processes that can be assessed (Quellmalz & Moody, 2004). Technology-enhanced items can take

advantage of drag-and-drop, hot spot, and simulation technologies along with the use of online tools to measure content that was previously not assessed or was assessed through constructed-response item formats requiring more elaborate scoring procedures. While most of the work with the technology-enhanced items has been done in the area of science, we envision the development of technology-enhanced items in both English language arts and mathematics. For example, a video presentation of a speech could be combined with a reading passage to provide students the opportunity to integrate reading and listening skills per the CCSS and evaluate content across diverse media.

Our assessment design calls for the use of extended constructed-response items. These items will be used to assess knowledge and skills not easily assessed with selected-response or technology-enhanced items. Constructed-response items allow students to demonstrate their use of complex thinking skills such as formulating comparisons or contrasts; proposing cause and effects; identifying patterns or conflicting points of view; categorizing, summarizing, or interpreting information; and developing generalizations, explanations, justifications, or evidence-based conclusions (Darling-Hammond & Pecheone, 2010). These complex thinking skills are consistent with the expectations for college- and career-readiness and will be included in both the English language arts and mathematics assessments.

The summative assessment system also includes the strategic use of performance events in both English language arts and mathematics. We anticipate two performance events per content area for grades 3–8 and up to six performance events in each content area by the end of grade 11. These events will evaluate the CCSS in ways that require more student-initiated planning, management of information and ideas, interaction with other materials and/or people, and production of more extended responses (e.g., oral presentations, exhibitions, product development, in addition to more extended written responses) that reveal additional abilities of students (Darling-Hammond & Pecheone, 2010) not captured by the other item types included in the summative assessment.

(A)(3)(c)(v) The Consortium is committed to the use of a computer adaptive summative assessment with performance events delivered via computer. We propose the computer adaptive system because of its ability to provide precision of measurement at the individual student level across the achievement continuum and for its efficiencies in scoring and reporting of results

(Kosty, McBride, Poggio, Wise, & Way, 2006; Lilley, Barker, & Britton, 2004; Rabinowitz, 2005). The computer-based platform also provides for immediate scoring of selected-response items and technology-enhanced items, the use of AI scoring applications for constructed-response items and performance events, and distributed scoring by teachers. For performance events, computer delivery provides a common platform that will enable students to produce responses in a variety of formats that can be digitized and uploaded for scoring. The computer assessment also reflects the appropriate use of technology as called for in the CCSS.

The Consortium will provide a paper-and-pencil option for a limited time (three years) to support States where required student access to computers for the test administration window remains a barrier. By definition, the paper-and-pencil forms will not be able to take advantage of all advances that SBAC will provide in the computer mode, but the integrity of the assessment will be maintained by adherence to the same test blueprint. The paper option will be offered as an accommodated form and will include selected-response items, constructed-response items, and performance events. The content addressed by the technology-enhanced items will be replaced with additional constructed-response items. To address security concerns, the paper option will be available in a limited window at the end of the school year. The costs associated with the administration and scoring of the paper-and-pencil assessments will be borne by the individual States requiring their use. Comparability between the paper option and the CAT will be established during standard setting.

(A)(3)(c)(vi) As shown in the Summary Table for (A)(3), the assessment system calls for the use of automated computer-based scoring of selected-response and technology-enhanced items. These methodologies are well developed and currently employed within State assessment programs (e.g., Oregon, Louisiana, Minnesota, and Utah). AI scoring will be employed for the constructed-response items and performance events. While we recognize that AI scoring continues to be an emerging technology, we believe that the methodology will continue to advance and that the benefits of the timely return of results warrant their continued development and use. By the time an item is administered on the operational assessment, the AI scoring methodology will have been implemented and evaluated for reliability, so the AI scoring can be used immediately upon the completion of the assessment administration. We also recognize that it is critically important to involve teachers in the scoring process to provide opportunities for them to internalize the standard

for student performance, to build school-level capacity for accurate scoring, and to ensure confidence in the AI scoring, as well as to provide the critical check on the accuracy of the AI-assigned score. Once the system becomes operational, we propose the use of a distributed teacher scoring model that is fully supported by the assessment platform. We project a two-week turnaround of teacher scores during the operational administration. For those select performance events that are found to tap student performances not effectively scored through AI technology, we will employ teachers as scorers and employ similar read-behind methodologies to ensure scorer accuracy. Teachers will play a greater role in the scoring of I/B and formative assessments.

(A)(3)(c)(vii) To help improve teaching and learning, SBAC's Theory of Action asserts that assessment reports must provide actionable information about students' progression toward college- and career-readiness. Through use of an interoperable electronic platform, SBAC will produce assessment reports that are more timely, relevant, and usable to stakeholders than those that are currently available to them. Our system will feature tools for producing preset and customizable reports that track and analyze the progress toward grade-level mastery and college- and career-readiness of individual students, classes, and subgroups of students (e.g., students with disabilities [SWDs], English learners [ELs], and students participating in a research intervention). We will provide timely user-friendly information to students, teachers, parents, administrators, and policymakers in traditional as well as innovative formats. For example, we will produce reports that show achievement level data (e.g., scale scores, growth measures) related to CCSS clusters and individual standards within the clusters. Interpretability of this information will be enhanced through innovative graphic displays (e.g., the *honeycomb* concept developed by Wireless Generation [Resnick and Berger, 2010] or the *ontology* model supported by CRESST [Baker & Herman, 2010]). In addition, in keeping with our Theory of Action, our score reports will follow the principles for reporting called for in the *Joint Standards* (1999) as well as recommendations from the research community (e.g., Hambleton, 2007). (See section (A)(4)(d) for more details about the overall design features as well as the approach and strategy for developing the reporting system.) Described below are the types of reports and design elements for reporting summative and I/B assessments. This information is also summarized in Appendix A3-4. For more information about the formative tools, processes, and practices, refer to section (A)(6).

Summative Assessment Reports

Accessible through the system portal, SBAC reporting will support both static report displays and dynamic customized reports based on data-mining tools. Summative assessment interpretive guides, graphic displays of data, and text focused on the particular needs of students, parents, teachers, school and district administrators, and key local, State, and national stakeholders will support the effective use of student performance results. They will be tailored to each major user group and allow for varying degrees of customization, depending on the needs of the particular user group. Access to student data and reporting tools will be credential dependent to ensure student privacy and protect the security of individual student results. Test results will be provided online and in paper form for three languages beyond English (i.e., Spanish and two languages determined based on results of the Home Language Survey). The major types of reports will include the following:

- *Performance summaries for each student on both standardized and performance measures.* This includes annual achievement scores on individual student achievement on the CCSS in English language arts and mathematics, including information on “readiness for college and career” (high school level) and “on track to be ready for college and career” (earlier grades) as well as indices of annual growth in student learning (e.g., Colorado Growth Model) that allow for normative comparisons of student gains. Score bands will be placed around each student test score based on the conditional standard error of measurement. Reports will provide item-level performance and sub-scores for content clusters from the ELA and mathematics tests, insofar as such results can be validly and reliably reported. Student scores on the performance components of the assessment will be reported separately as well as integrated statistically into the summary scale score and performance levels. Performance level descriptors for the summative assessment results as a whole will provide guidance in understanding the performance levels identified, as appropriate to grade level and content area. Student performance will be explained, as it is in the National Assessment of Educational Progress (NAEP) reports, with examples to enhance interpretation of results.
- *Class/school/district/State summaries.* Aggregated student data will be reported, providing quantitative and descriptive data to all stakeholders. Group data will be reported by teacher, school, district, Statewide, and across participating States. Also, it will be

disaggregated by designated subgroups, as identified in Federal and State regulations or by local-level interest. In reporting results for special populations (e.g., SWDs and ELs), the system will access a table or platform that clearly delineates which accommodations are available and which would lead to valid and non-valid scores.

Interim/Benchmark Assessment Reports

To maximize the usability of I/B assessment scores (as a predictor of performance on the summative assessment or as an early warning system to focus teacher attention on student and class learning deficits), results will be reported on the same scale as the summative assessments and mirror the summaries described above, with one important enhancement. Because I/B assessments are meant to drive instructional interventions more directly than their summative counterparts, I/B items will be built directly around the concept of learning progressions and the focus of each I/B assessment will be on fewer and deeper concepts—reports will be able to provide more direct student-level information at a finer grain size. The results will be of immediate value to students, parents, teachers, and school and district administrators as comparable evidence of student progress and program effectiveness throughout the school year—evidence not currently available in most States. Through credential-based access to the reporting suite (via the system portal), the full range of stakeholders can view and generate meaningful I/B assessment reports appropriate to their level of access.

Formative Assessment: Tools, Processes, and Practices

While formative assessments can take many forms, for the results to be put to diagnostic use, the assessment process should be integrated with ongoing instruction. The heavily contextualized nature of such assessment processes creates a challenge for reporting results in ways that are most directly useful to teachers and students.

SBAC will provide information in the form of professional agreements about classroom assessment tools, processes, and practices and what the results from them will look like. In practice, such reports will primarily take the form of records of student performances, including grades, teacher observation protocols and notations, classroom activities, evaluations using scoring guides, results from quizzes, and samples of student work, including methods of combining the variety of data into an evidence-based indicator of student growth and achievement.

Standardized protocols will be developed for the identification of evidence of student learning and for the interpretation of that evidence, and they will be shared throughout the system.

Electronic technology will provide the platform and the tools for teachers to share information about these processes and practices, between classrooms, through professional development, across grade levels, between schools and school districts, throughout the State, and across States. This component is unique in this assessment system in that classroom teachers take the lead in adding student performance information to the system.

By coordinating the ongoing results of the formative component with the periodic results of I/B assessments and the summary results of the summative assessment, this assessment system will address all of the essential needs for measuring student achievement and growth and student progress toward college- and career-readiness.

Summary Table for (A)(3): English Language Arts – Summative Reading

Grade	Type of Component	Types of data produced / how used	Frequency and timing of administration	Number and types of items, and distribution of item types ³	Administration mode	Scoring method and estimated turnaround time
3-8*	Comprehensive summative assessing CCSS	Scale score used for achievement or growth metric	Must be administered at least once annually; 1–2 testing opportunities per year.	30 Selected-response items (SR) 3 Extended constructed-response items (ECR) 7 Technology-enhanced items (TE) 1 Performance event	Computer adaptive: SR, ECR, and TE Computer delivered: Teacher-administered performance event	Computer adaptive: includes automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance event: combination of AI and teacher scoring with read-behinds to be completed within two weeks of administration)
HS*	Comprehensive summative assessing CCSS	Scale score used for achievement or growth metric	Must be administered at least once annually; 1–2 testing opportunities per year. Must be administered in grade 11 until research shows that scores from assessments administered in grades 9 or 10 are comparable to those from grade 11.	30 Selected-response items 3 Extended constructed-response items 7 Technology-enhanced items Up to 3 Performance events by the end of grade 11	Computer adaptive: SR, ECR, and TE Computer-delivered: Teacher-administered performance event	Computer adaptive: includes automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance events: combination of AI and teacher scoring with read-behinds to be completed within two weeks of administration

*Ten additional FT items required annually to build the item bank

³ Final distribution of items by grade to be determined by CCSS.

Report(s) produced, intended use, audience: Comprehensive Summative Reading Assessment

Individual student achievement and growth report: Student achievement and progress toward college- and career-readiness; intended to inform instructional effectiveness. Exemplars are provided in Appendices A4-4c and A4-4d.

Audience: Teachers, parents, students, administrators

Class, school, district, and State reports of student achievement: Student achievement and progress toward college- and career-readiness; intended to inform instructional and program effectiveness. Exemplars are provided in Appendices A4-4a, A4-4b, and A4-4d.

Audience: Teachers, administrators, parents, policymakers

Consortium-level report for Title I Accountability: Consortium-level student achievement toward college- and career-readiness for Federal reporting. An exemplar is provided in Appendix A4-4e.

Audience: Administrators, parents, policymakers, USED

Data Mining Tool: Supports dynamic subgroup disaggregation of assessment results by class, school, or district to allow public access to assessment results. Also supports secure-access data mining for teachers and authenticated administrators to inform instructional effectiveness for various subgroups and programs. Exemplars are provided in Appendices A4-4a, A4-4b, A4-4d, and A4-4e.

Audience: Policymakers, administrators, teachers, education stakeholders

Human Scoring Accuracy and Reliability Report: Inter-rater reliability and scoring validity (based on calibration papers) will be available on demand for the purpose of monitoring scoring validity and reliability.

Audience: Administrators

Summary Table for (A)(3): English Language Arts – Summative Writing, Listening and Speaking, and Language

Grade	Type of Component	Types of data produced / how used	Frequency and timing of administration	Number and types of items, and distribution of item types ⁴	Administration mode	Scoring method and estimated turnaround time
3-8*	Comprehensive summative assessing CCSS	Scale score used for achievement or growth metric	Must be administered at least once annually; 1–2 testing opportunities per year.	Writing/Language: 6 Selected-response items (SR) 6 Technology-enhanced items (TE) 2 On-demand writing prompts 1 Performance event (involving prewriting, drafts, and edits over time) Speaking/Listening: 2 Technology-enhanced items and 2 technology-enhanced constructed response (tied to writing performance event)	Computer adaptive: TE, CR items. Computer-delivered: Teacher-administered performance events	Computer adaptive: includes automated computer-scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be scored within two weeks of administration Performance events: combination of AI and teacher scoring with read-behinds to be completed within two weeks of administration
11*	Comprehensive summative assessing all CCSS	Scale score used for achievement or growth metric	Must be administered in grade 11 until research shows that scores from assessments administered in grades 9 or 10 are comparable to those from grade 11. Students have 1–2 testing opportunities.	Writing/Language: 6 Selected-response items (SR) 6 Technology-enhanced items (TE) 2 On-demand writing prompts Up to 3 performance events by the end of grade 11 Speaking/Listening: 2 Technology-enhanced items and 2 technology-enhanced constructed-response items	Computer adaptive: TE, CR items. Computer-delivered: Teacher-administered performance events	Computer adaptive: includes automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance events: combination of AI and teacher scoring with read-behinds to be completed within two weeks of administration

*Six additional FT items required annually to build the item bank

⁴ Final distribution of items by grade to be determined by CCSS.

Report(s) produced, intended use, audience: Summative Writing, Listening and Speaking, and Language

Individual student achievement and growth report: Student achievement and progress toward college- and career-readiness; intended to inform instructional effectiveness.

Audience: Teachers, parents, students, administrators

Class, school, district, and State reports of student achievement: Student achievement and progress toward college- and career-readiness; intended to inform instructional and program effectiveness.

Audience: Teachers, administrators, parents, policymakers

Consortium-level report for Title I Accountability: Consortium-level student achievement toward college- and career-readiness for Federal reporting.

Audience: Administrators, parents, policymakers, USED

Data Mining Tool: Supports dynamic subgroup disaggregation of assessment results by class, school, or district to allow public access to assessment results. Also supports secure-access data mining for teachers and authenticated administrators to inform instructional effectiveness for various subgroups and programs.

Audience: Policymakers, administrators, teachers, education stakeholders

Human Scoring Accuracy and Reliability Report: Inter-rater reliability and scoring validity (based on calibration papers) will be available on-demand for the purpose of monitoring scoring validity and reliability.

Audience: Administrators

Summary Table for (A)(3): Summative Mathematics

Grade	Type of Component	Types of data produced / how used	Frequency and timing of administration	Number and types of items, and distribution of item types ⁵	Administration mode	Scoring method and estimated turnaround time
3–8*	Comprehensive summative assessing all CCSS	Scale score used for achievement or growth metric	Must be administered at least once annually; 1–2 testing opportunities per year.	19 Selected-response items (SR) 3 Extended constructed-response items (ECR) 18 Technology-enhanced items (TE) 2 Performance events	Computer adaptive: SR, ECR, and TE items Computer-delivered: Teacher-administered performance events	Computer adaptive: includes automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance events: combination of AI and teacher scoring with read-behinds to be completed within two weeks of administration
HS*	Comprehensive summative assessing all CCSS	Scale score used for achievement or growth metric	The adaptive summative must be administered in grade 11 until research shows that scores from assessments administered in grades 9 or 10 are comparable to those from grade 11. Students have 1–2 testing opportunities on the adaptive portion of the summative assessment.	19 Selected-response items (SR) 3 Extended constructed-response items (ECR) 18 Technology-enhanced items (TE) Up to 6 performance events by the end of grade 11	Computer adaptive: SR, ECR, and TE items Computer-delivered: Teacher-administered performance events	Computer adaptive: includes automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance events: combination of AI and teacher scoring with read-behinds to be completed within two weeks of administration

*Four additional FT items required annually to build the item bank

⁵ Final distribution of items by grade to be determined by CCSS.

Report(s) produced, intended use, audience: Summative Mathematics

Individual student achievement and growth report: Student achievement and progress toward college- and career-readiness; intended to inform instructional effectiveness. Exemplars are provided in Appendices A4-4c and A4-4d.

Audience: Teachers, parents, students, administrators

Class, school, district, and State reports of student achievement: Student achievement and progress towards college- and career-readiness; intended to inform instructional and program effectiveness. Exemplars are provided in Appendices A4-4a, A4-4b, and A4-4d.

Audience: Teachers, administrators, parents, policymakers

Consortium-level report for Title I Accountability: Consortium-level student achievement toward college- and career-readiness for Federal reporting. An exemplar is provided in Appendix A4-4e.

Audience: Administrators, parents, policymakers, USED

Data Mining Tool: Supports dynamic subgroup disaggregation of assessment results by class, school, or district to allow public access to assessment results. Also supports secure-access data mining for teachers and authenticated administrators to inform instructional effectiveness for various subgroups and programs. Exemplars are provided in Appendices A4-4a, A4-4b, A4-4d, and A4-4e.

Audience: Policymakers, administrators, teachers, education stakeholders

Human Scoring Accuracy and Reliability Report: Inter-rater reliability and scoring validity (based on calibration papers) will be available on-demand for the purpose of monitoring scoring validity and reliability.

Audience: Administrators

Summary Table for (A)(3): English Language Arts – Interim/Benchmark Reading

Grade	Type of Component	Types of data produced / how used	Frequency and timing of administration	Number and types of items, and distribution of item types ⁶	Administration mode	Scoring method and estimated turnaround time
3–8	Interim/ Benchmark assessing CCSS	Scale score used for achievement or growth metric Cluster-level achievement and growth	Multiple testing opportunities per year	<u>Comprehensive Assessment:</u> 30 SR, 3 ECR, 7 TE 1 PE <u>Cluster Assessment:</u> 15 SR, 1 ECR, 3 TE 1 PE available following release from summative pool	Computer adaptive: SR, ECR, and TE Computer delivered: Teacher-administered performance event	Computer adaptive: automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance event: AI scoring with teacher score resolution
HS	Interim/ Benchmark assessing CCSS	Scale score used for achievement or growth metric Cluster-level achievement and growth	Multiple testing opportunities per year	<u>Comprehensive Assessment:</u> 30 SR, 3 ECR, 7 TE 1 PE <u>Cluster Assessment:</u> 15 SR, 1 ECR, 3 TE 1 PE available following release from summative pool	Computer adaptive: SR, ECR, and TE Computer-delivered: Teacher-administered performance event	Computer adaptive: automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance event: AI scoring with teacher score resolution

⁶ Final distribution of items by grade to be determined by CCSS.

Report(s) produced, intended use, audience: Interim/Benchmark Reading Assessment

Content Cluster Learning Progression Visual Report: Progress on content clusters of the Common Core State Standards; intended to inform instruction.

Audience: Teachers, parents, students, administrators

Individual student achievement and growth report: Student achievement and progress toward college- and career-readiness; intended to inform instructional effectiveness.

Audience: Teachers, parents, students, administrators

Class, school, district, and State reports of student achievement: Student achievement and progress toward college- and career-readiness; intended to inform instructional and program effectiveness.

Audience: Teachers, administrators, parents, policymakers

Consortium-level report for Title I Accountability: Consortium-level student achievement toward college- and career-readiness for Federal reporting.

Audience: Administrators, parents, policymakers, USED

Data Mining Tool: Supports dynamic subgroup disaggregation of assessment results by class, school, or district to allow public access to assessment results. Also supports secure-access data mining for teachers and authenticated administrators to inform instructional effectiveness for various subgroups and programs.

Audience: Policymakers, public, administrators, teachers, education stakeholders

Human Scoring Accuracy and Reliability Report: Inter-rater reliability and scoring validity (based on calibration papers) will be available on-demand for the purpose of monitoring scoring validity and reliability.

Audience: Administrators

Summary Table for (A)(3): English Language Arts – Interim/Benchmark Writing, Listening and Speaking, and Language

Grade	Type of Component	Types of data produced / how used	Frequency and timing of administration	Number and types of items, and distribution of item types ⁷	Administration mode	Scoring method and estimated turnaround time
3–8	Interim/ Benchmark assessing CCSS	Scale score used for achievement or growth metric Cluster-level achievement and growth	Multiple testing opportunities per year	<u>Comprehensive Assessment:</u> Writing/Language: 6 SR, 6 TE, 2 WP, 1 PE Speaking/Listening: 2 TE, 2 ECR <u>Cluster Assessment:</u> Writing/Language: 6 SR, 6 TE, 2 WPs, 1 PE Speaking/Listening: 2 TE, 2 ECR	Computer adaptive: TE, CR items Computer-delivered: Teacher administered performance events	Computer adaptive: automated computer-scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance event: AI scoring with teacher score resolution
HS	Interim/ Benchmark assessing CCSS	Scale score used for achievement or growth metric Cluster-level achievement and growth	Multiple testing opportunities per year	<u>Comprehensive Assessment:</u> Writing/Language: 6 SR, 6 TE, 2 WP, 1 PE Speaking/Listening: 2 TE, 2 ECR <u>Cluster Assessment:</u> Writing/Language: 6 SR, 6 TE, 2 WPs, 1 PE Speaking/Listening: 2 TE, 2 ECR	Computer adaptive: TE, CR items Computer-delivered: Teacher-administered performance events	Computer adaptive: automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance event: AI scoring with teacher score resolution

⁷ Final distribution of items by grade to be determined by CCSS.

Report(s) produced, intended use, audience: Interim/Benchmark Writing, Listening and Speaking, and Language Assessment

Content Cluster Learning Progression Visual Report: Progress on content clusters of the Common Core State Standards; intended to inform instruction.

Audience: Teachers, parents, students, administrators

Individual student achievement and growth report: Student achievement and progress toward college- and career-readiness; intended to inform instructional effectiveness.

Audience: Teachers, parents, students, administrators

Class, school, district, and State reports of student achievement: Student achievement and progress toward college- and career-readiness; intended to inform instructional and program effectiveness.

Audience: Teachers, administrators, parents, policymakers

Consortium-level report for Title I Accountability: Consortium-level student achievement toward college- and career-readiness for Federal reporting.

Audience: Administrators, parents, policymakers, USED

Data Mining Tool: Supports dynamic subgroup disaggregation of assessment results by class, school, or district to allow public access to assessment results. Also supports secure-access data mining for teachers and authenticated administrators to inform instructional effectiveness for various subgroups and programs.

Audience: Policymakers, administrators, teachers, education stakeholders

Human Scoring Accuracy and Reliability Report: Inter-rater reliability and scoring validity (based on calibration papers) will be available on-demand for the purpose of monitoring scoring validity and reliability.

Audience: Administrators

Summary Table for (A)(3): Interim/Benchmark Mathematics

Grade	Type of Component	Types of data produced / how used	Frequency and timing of administration	Number and types of items, and distribution of item types ⁸	Administration mode	Scoring method and estimated turnaround time
3–8	Interim/ Benchmark assessing CCSS	Scale score used for achievement or growth metric Cluster-level achievement and growth	Multiple testing opportunities per year	<u>Comprehensive Assessment:</u> 19 SR, 3 ECR, 18 TE 2 PE <u>Cluster Assessment:</u> 15 SR, 1 ECR, 3 TE 1 PE available following release from summative pool	Computer adaptive: SR, ECR, and TE Computer- delivered: Teacher- administered performance event	Computer adaptive: automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance event: AI scoring with teacher score resolution
HS	Interim/ Benchmark assessing CCSS	Scale score used for achievement or growth metric Cluster-level achievement and growth	Multiple testing opportunities per year	<u>Comprehensive Assessment:</u> 30 SR, 3 ECR, 7 TE 1 PE <u>Cluster Assessment:</u> 15 SR, 1 ECR, 3 TE 1 PE available following release from summative pool	Computer adaptive: SR, ECR, and TE Computer- delivered: Teacher- administered performance event	Computer adaptive: automated computer scoring of SR and TE items; 100% AI-scored ECR items; 10% teacher read-behinds to be completed within two weeks of administration Performance event: AI scoring with teacher score resolution

⁸ Final distribution of items by grade to be determined by CCSS.

Report(s) produced, intended use, audience: Interim/Benchmark Mathematics Assessment

Content Cluster Learning Progression Visual Report: Progress on content clusters of the Common Core State Standards; intended to inform instruction.

Audience: Teachers, parents, students, administrators

Individual student achievement and growth report: Student achievement and progress toward college- and career-readiness; intended to inform instructional effectiveness.

Audience: Teachers, parents, students, administrators

Class, school, district, and State reports of student achievement: Student achievement and progress toward college- and career-readiness; intended to inform instructional and program effectiveness.

Audience: Teachers, administrators, parents, policymakers

Consortium-level report for Title I Accountability: Consortium-level student achievement toward college- and career-readiness for Federal reporting.

Audience: Administrators, parents, policymakers, USED

Data Mining Tool: Supports dynamic subgroup disaggregation of assessment results by class, school, or district to allow public access to assessment results. Also supports secure-access data mining for teachers and authenticated administrators to inform instructional effectiveness for various subgroups and programs.

Audience: Policymakers, administrators, teachers, education stakeholders

Human Scoring Accuracy and Reliability Report: Inter-rater reliability and scoring validity (based on calibration papers) will be available on-demand for the purpose of monitoring scoring validity and reliability.

Audience: Administrators

Summary Table for (A)(3): Formative Assessment Tools, Processes, and Practices

Grade	Type of Component	Types of data produced / how used	Frequency and timing of administration	Number and types of items, and distribution of item types ⁹	Administration mode	Scoring method and estimated turnaround time
3–8, HS	Formative Assessment Tools, Processes, and Practices	Cluster-level achievement and growth	Teacher determined: as necessary to support teaching and learning	Teacher determined: a range of formal and informal assessments using the full range of item types as appropriate to learning objective	<p>Computer adaptive: SR, ECR, and TE</p> <p>Computer-delivered: Teacher-administered performance event and TE</p> <p>Classroom exercises: Teacher administered</p>	<p>Computer adaptive: automated computer scoring of SR, TE, ECR</p> <p>Performance event: AI scoring with teacher score resolution</p> <p>Classroom exercises: Teacher scored</p>
Report(s) produced, intended use, audience: Formative Assessment Tools, Processes, and Practices						
<p>To the extent that interim/benchmark assessment items are used as part of the formative assessment process, interim/benchmark reporting will be available.</p> <p>Reporting for formative classroom-embedded exercises will be teacher determined.</p>						

⁹ Final distribution of items by grade to be determined by CCSS.

(A)(4) Assessment System Development (up to 35 points)

The extent to which the eligible applicant's plan for developing the proposed assessment system will ensure that the assessment system is ready for wide-scale administration in a manner that is timely, cost-effective, and consistent with the proposed design and incorporates a process for ongoing feedback and improvement. In determining the extent to which the development plan has these attributes, we will consider—

(a) The approaches for developing assessment items (*e.g.*, evidence centered design, universal design¹⁰) and the rationale for using those approaches; the development phases and processes to be implemented consistent with the approaches; and the types of personnel involved in each development phase and process (*e.g.*, practitioners, content experts, assessment experts, experts in assessing English learners, experts in assessing students with disabilities, psychometricians, cognitive scientists, IHE representatives, career and technical education experts);

(b) The approach and strategy for designing and developing accommodations (as defined in the NIA), accommodation policies, and methods for standardizing the use of those accommodations for—

- (i) English learners; and
- (ii) Students with disabilities;

(c) The approach and strategy for ensuring scalable, accurate, and consistent scoring of items, including the approach and moderation system (as defined in the NIA) for any human-scored items that are part of the summative assessment components and the extent to which teachers are trained and involved in the scoring of assessments;

(d) The approach and strategy for developing the reporting system; and

(e) The overall approach to quality control; and the strategy for field testing assessment items, accommodations, scoring systems, and reporting systems, including, with respect to assessment items and accommodations, the use of representative sampling of all types of student populations, taking into particular account high- and low-performing students and different types of English learners and students with disabilities.

¹⁰ Universal design for learning” is used as that term is defined in section 103(24) of the HEA.

(A)(4)(a) SBAC is committed to the development of a coherent system in which summative assessments, interim/benchmark (I/B)¹¹ assessments, formative tools, professional development opportunities, and instructional resources work together to support student progress toward college and career readiness and high-quality instruction. Guided by the principles adopted by the Joint Committee on *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999), our development processes focus on maximizing the usefulness of each component of our proposed assessment system and the validity of interpretations drawn from results from each measure. We provide an overview of our proposed assessment system in Appendix A3-1.

System development starts with clear specification of the constructs to be measured, intended users and uses, and target student population, so that items and performance events comprising each assessment can be purposely designed to address the intended learning construct(s) in ways that are consistent with intended uses and maximally accessible for all students. Our approach brings together the last decade's work on the logic of assessment as a process of reasoning from evidence with principles of evidence-based design and model-based assessment (NRC, 2001; Mislevy & Riconscente, 2006; Baker, 2007). Its hallmarks are (1) clear specification of progressions of learning expectations, coordinated across summative, I/B, and formative assessment tools at increasingly more finely grained levels of detail; (2) test blueprints with clear rules for sampling the full domain of those learning expectations; (3) learning-based item and performance event design templates, including scoring rubrics that establish clear targets for teaching and learning; support coordinated sets of summative, interim, and formative tools; and provide a substantive basis for test comparability from year to year to complement psychometric indices; and (4) the potential for innovation and efficiency through technology-enhanced items and automated scoring.

It will be the charge of the Assessment Design Working Group to ensure that steps in the process of development for each component are transparent and lead to measures that are valid, fair, of high technical quality, and appropriate for the intended purposes. Guided by the Consortium's Technical Advisory Committee (TAC), the Research and Evaluation Working

¹¹ The interim/benchmark and formative components are described in the MOU as optional "formative/benchmark components," which is a broader description for these system features.

Group will monitor these steps and provide ongoing feedback to the Assessment Design Working Group and Steering Committee about the technical adequacy of the development processes and areas for improvement.

Building the assessment frameworks, test blueprints, and item templates

Linking standards and assessments, as well as making expectations clear for teaching and learning, requires that we first specify standards in ways that precisely identify what is to be assessed and how proficiency is defined relative to that content (i.e., at what levels of cognitive demand are students expected to know that content and how should they be able to apply it during testing?). For example, the CCSS for mathematics lay out the major domains of mathematics, and the clusters and standards for each grade constitute the content of what students are supposed to “know,” but we must also specify what students are expected to be able to “do” with the content. Building from the CCSS, we propose to convene key stakeholders and content specialists to develop assessment frameworks that precisely lay out the content and cognitive demands that define college- and career-readiness for each grade level. We then will systematically develop items and performance events that collectively provide evidence for valid conclusions about what a student knows and can do in relation to college- and career-readiness.

The vertical pathways to college- and career-readiness embedded in the CCSS across grades also enable us to hypothesize implicit learning progressions (i.e., empirically validated descriptions of how learning typically unfolds within a curricular domain [Darling-Hammond & Pecheone, 2010]) and what prior understandings are most essential for future performance.¹² Similarly, within grades, with the help of teachers and learning theorists, we will identify within-year learning progressions on which to base the summative and I/B assessments and formative tools. We plan to validate and strengthen these hypothesized progressions as one element of our planned research agenda (see section (A)(5)).

Once the comprehensive assessment framework that defines learning expectations is developed, assessment blueprints will be constructed that specify how items and performance events can work in various combinations to measure the full range of the CCSS. Performance event design templates then can be created to provide more concrete detail about the nature of test items and/or performance events that best measure different types of knowledge and skills. In this way, the content and cognitive demand are used in combination to describe and measure

¹² See Appendix A4-1 for examples of learning progressions in ELA and mathematics.

the development of competency within a domain (NRC, 2001).

Based on prior research and learning theory (Baker, 2007), we plan to develop reusable event design templates for generating a variety of selected- and constructed-response items and performance events. These templates will be designed to measure specific types of understanding (e.g., conceptual, procedural, or strategic knowledge) or skills (e.g., problem solving, reasoning, or argumentation). This type of reusable event design template or shell invokes a consistent set of principles, criteria, and expectations for defining appropriate content limits and generating events and scoring criteria to address particular cognitive demands that can be adapted for use across subject matter domains and grade levels. For example, using a common definition of problem solving, an appropriate event design might feature a non-routine, authentic problem (at various degrees of near or far transfer) that requires students to integrate core concepts and principles they have learned and apply reasoning to develop a novel solution strategy. In doing so, they follow the general path of identifying the problem within the novel context, identifying critical variables and their relationships, constructing a suitable representation of the problem, solving the problem, justifying their strategy, and/or communicating results (see, e.g., Mayer, 1992; NRC, 1999; Hiebert et al., 1996).

In the past, such authentic problem-solving tasks have been difficult for State assessment programs to implement in ways that are appropriate for large-scale assessment. Complex unstructured problems may have enabled high-competency students to show what they know and can do, but may have left low-ability students unable to respond. The remedy of using highly structured tasks may have provided access for low-ability students, but it does not reflect true problem solving. Over the past two decades, this picture has improved. The technology for building such items and tasks has evolved significantly, to the point where it may increase access and generalizability across tasks (Baker, 2007; Darling-Hammond & Pecheone, 2010). That progress allows our performance event designs, coupled with the use of technology, to offer a possible solution: technology-enhanced adaptive items and performance events can provide additional challenge to those students who are ready to move forward, while at the same time providing progressive scaffolding and hint systems to customize difficulty and diagnosis for other students.

Building in support for flexibility and accessibility, the templates will provide item developers with efficient, standard routines for item development and scoring, which ultimately

may provide capacity for automated development. Further, by supporting constancy and comparability in the nature of the learning assessed from year to year, this approach also will allow for better measurement of growth over time along key domains and/or learning progressions. Initially, we will develop detailed design templates for the summative performance events to support comparability from year to year. However, these templates also provide an easy link among summative, I/B, and formative functions and classroom instruction. That is, teachers can use the design templates to generate rich classroom assignments and curriculum-embedded assessment tasks.

Model for the development of the summative assessments

Building from the CCSS, the Consortium will develop cognitive models for the domains of ELA and mathematics that specify the content elements and relationships reflecting the sequence of learning that students would need to achieve college- and career-readiness. Implementation of these models will require close collaboration with IHEs, workplace representatives, cognitive scientists, classroom teachers, experts in assessing English learners and students with disabilities, and content experts in ELA and mathematics who have studied the CCSS and have deep understanding of the knowledge, skills, and cognitive processes that characterize the pathways in grades 3 through 8 and high school that lead to the college- and career-ready outcomes specified in the standards. Once the content of the CCSS has been studied, an assessment framework will be developed that specifies the item types that can provide the necessary evidence of student achievement.

The next critical step in the process is the development of item and performance event shells. These will provide constancy in the set of constructs measured by a set of items and events. They also will allow for better measurement of achievement and growth, development of equivalent tasks, and a framework to guide the development of scoring rubrics. The item/event shells will be complemented by the development of item specifications that will delineate the content limits to be reflected in the items. Once the assessment framework, item/event shells, and item specifications have been established, the test blueprint will be developed that will specify the number and types of items to be established in the summative item banks. An overview of the development process for the summative assessments can be found in Appendix A4-2.

Model for the development of the interim/benchmark assessments

The I/B assessments are designed to provide a measure of progress toward the summative goals and to help identify learning gaps and guide instruction (Rabinowitz, 2009). Consequently, I/B assessments will be developed to zero in on a student's current level of understanding in each learning progression in English language arts and mathematics. Learning progressions are empirically validated descriptions of how learning typically unfolds within a curricular domain or area of knowledge and skill (Darling-Hammond & Pecheone, 2010). By design, a set of items developed to measure learning progressions for the purposes of the I/B assessments will be deep in terms of content coverage for each content cluster in the CCSS—a larger number of items will be used to measure small, incremental differences in what students know and can do. Item “clusters” will be developed that can hone in on students' precise level of understanding of those linked pieces of knowledge and/or demonstration of skills that constitute a progression.

The I/B assessments will require the specification of the learning progressions as a first step in the development process. Once identified, these learning progressions will be mapped to the CCSS and the evidence-based model again will be applied to determine the knowledge and/or skills a student must demonstrate to show mastery of the steps in the learning progression. As with the summative assessment, an assessment framework, test blueprint, and item specifications will be developed to guide the development of the I/B assessment item bank.

Item development

The Assessment Design Working Group will work in conjunction with technical advisors and vendors to build from the item/event shells and item specifications in producing item development training materials. These training materials will serve as guides to item development but also as criteria for evaluating item alignment and quality. Initially, we will build both the summative and I/B assessment item banks by collecting existing items and performance events from States.¹³ All contributed items will be subjected to a centralized review for technical adequacy, appropriateness, and alignment that will involve extensive cross-State participation. The value of this centralized review will be to help to establish a common Consortium standard,

¹³ An assumption has been made that for the summative assessments 25% of the SR items and 15% of the CR items would be supplied from existing State item banks. For the I/B assessments, 30% of the SR items and 20% of the CR items would be contributed by the States. All performance events and technology-enhanced items will be developed during the grant period and will not be supplemented by State item banks.

clarifying expectations for test items, and helping to ensure uniform application of the alignment and quality criteria.

Once the alignment and quality control reviews have been completed, the distribution of items within the banks will be determined and targets for item development established. At first, we propose that regional item development meetings be convened to support the States as they develop items to complete the item bank. These meetings will include classroom teachers, IHE and workplace representatives, content experts, measurement experts, and specialists in assessing ELs and SWDs. We believe that conducting the development of items on a regional (i.e., multi-State) basis will move the Consortium forward with the implementation of a broader vision than might occur if development is situated within each State. After the initial drafting of items at these regional workshops, States will be responsible for finalizing and editing the items. States may choose to continue to use teachers and IHE representatives in this process or they may decide to retain the services of vendors to support their item development efforts.

Though States have extensive experience in developing SR and CR items, they do not bring the same depth and breadth of experience in developing technology-enhanced items and performance events. Consequently, we propose that States that have the greatest experience in these item types (e.g., Missouri, Ohio, Oregon, and Utah) take the lead in developing these items. These States are best positioned to move forward with development while research is conducted and monitored to support their efforts.

Once stakeholders have reviewed and approved the items a State has developed, items will be randomly distributed for review by at least two other States. Each State will apply common criteria in evaluating (1) the degree of alignment of each item to the CCSS (blind to the coding from other States) in terms of content and cognitive demand, and (2) the technical quality (e.g., clarity, accessibility, and developmental appropriateness) of each item. Proposed item edits will be tracked separately by each State. If consensus cannot be achieved, the item-originating State can choose to discard the item or move it to another State for review until two States have approved the item.

Prior to pilot and field testing, items will be reviewed for bias and sensitivity by a Consortium-wide panel that includes assessment-literate educators, representatives from IHEs and the workplace, and members of the community at large. Items identified as problematic will be sent to an expert review panel to determine whether to (1) reject the recommendation and

retain the item as submitted, (2) edit the item as recommended by the panel, or (3) discard the item. Items that are approved by the sensitivity panel will be eligible for pilot and field testing. States will receive reports regarding the degree to which their items were edited, the volume and percentage of items that progressed to the field test, and other information that may be helpful in refining the item development support materials, improving item writer quality, and providing additional motivation for high-quality item writing.

As a further measure of quality control, SBAC will submit 25% of the Consortium-approved items to an independent contractor for review of whether they adhere to quality standards and meet alignment criteria in terms of content and cognitive complexity. If the results based on the random sampling of 25% of the bank reflect strong alignment and adherence to quality standards, items will be approved for pilot and field testing and the review protocols will remain in place. If the results of the alignment study suggest that the item pool needs further study, the Steering Committee will collect recommendations for improving the cross-State review process.

Development for I/B assessments and formative assessment tools will follow similar processes. This model ensures equal attention to quality and access concerns during item development regardless of the “stakes” of the assessment.

Accommodations

(A)(4)(b)(i) and (ii) SBAC’s approach for item and performance event development, accommodations, accommodation policies, and accommodations use will be consistent with the principles of universal design for assessment and inclusive assessment practices, as defined by the National Center on Educational Outcomes (NCEO), and findings from recent research on accommodating students with disabilities (SWDs) (e.g., Russell, Goldberg, & O’Connor, 2003; Thompson, Thurlow, & Malouf, 2004) and English learners (ELs) (e.g., Kopriva, Emick, Hipolito-Delgado, & Cameron, 2007; Rivera, Acosta, & Willner, 2008). SBAC is committed to developing common understandings, guidelines, and/or policies and practices that will lead to more accurate reporting for ELs, SWDs, and Section 504 students and will ensure that students are served across States in an equitable manner. As a demonstration of the Consortium’s commitment to developing coherent accommodation policies and working together to achieve this common goal, twenty-three States within the Consortium supported the submission of an Enhanced Assessment Grant (EAG) to develop the required common accommodation policies. The specific goals of the EAG and proposed activities associated with each of these goals can be

found in Appendix A4-3.

Scoring

(A)(4)(c) The Consortium will build on the existing methods and capacities for automated scoring of selected-response and technology-enhanced CR items, and AI-scoring and hand-scoring of more complex CR items and performance events. The Consortium is committed to meeting or exceeding the industry standard for technical adequacy during all phases of the scoring process. We also are committed to the strategic use of teachers for scoring CRs and performance events, and envision teachers supporting the creation of scoring guides and selecting anchor and training papers. In the section that follows, we outline the process to be implemented to ensure accurate and consistent scoring of items. Section (A)(5) describes planned activities to ensure the ongoing monitoring of these scoring activities.

Development of scoring rubrics

Development of the scoring rubrics will be an integral part of the scoring process. Building on the reasoning-from-evidence analysis of the content domains and the item specifications, we will target the knowledge and skills to be assessed. Through the purposeful development of items and use of item and event templates, we will specify the CCSS-based knowledge or skills students are expected to demonstrate with each item. From this content specification, scoring rubrics will be developed. Consistent with our Theory of Action, we propose the active participation of teachers as well as IHE and workplace representatives in each participating State in all phases of the scoring process.

Selection of AI and teacher training materials

Once field testing has been completed, student work will be reviewed and exemplars selected to support AI scoring of constructed-response items and teacher-moderated scoring of more complex constructed-response items and performance events. Different models for AI scoring exist, including those that are based on natural language processing and those that analyze samples of student responses to produce a model of a rater's scoring behavior. The Assessment Design Working Group will collaborate with technical advisors in selecting the most appropriate model for meeting Consortium needs. Regardless of the model selected, it will be important to have sample student work to verify the AI score assigned or to train the scoring engine.

- *Selected-response items.* Selected-response items will be “machine-scored” as part of the

assessment platform. All answer keys will be verified by at least two content specialists. Key-check analyses will be conducted based on the field test responses to ensure score-key accuracy.

- ***Technology-enhanced items.*** Technology-enhanced items will be developed that include a variety of response formats including drag-and-drop, hot spot, drawing, graphing, and written responses. The majority of these will be scored automatically; the methodology used to score written responses is described in the following section. Drag-and-drop, drawing, and graphing responses will be developed to elicit a range of correct response options. As items are developed, scoring rubrics will be written that identify possible correct responses for each score point option. These rubrics will be compared to actual student responses collected during field testing. Field testing will require that a representative sample of students be included to ensure that the anchor papers reflect all possible ways to earn full and partial points on any given item. These model responses will be used to “train” the automated scoring system. Accuracy of the assigned scores will be validated by human raters as part of the automated-scoring training process. This methodology is in current use. It has been successfully applied to the scoring of items across different content domains and is used in high-stakes testing for medical licensure and architectural design (Bennett, 2004). Items that do not yield sufficiently high consistency between human and automated scoring (above 0.80) will not be considered for use on the summative assessment.
- ***Constructed-response items, performance events, and verbal responses.*** The scoring of more complex responses (constructed-response items, performance events, and verbal responses) will require the use of AI scoring. While AI scoring continues to be an emerging technology and a focus of ongoing research, the Consortium has confidence in the potential for these methodologies to strongly and appropriately support the Consortium’s goals. Methodologies exist to score the content and quality of written responses in terms of grammar, usage, mechanics, style, organization, and development (ETS, 2010). We believe that AI will continue to advance in the next three years, and that the time frame for expected new developments makes this a compelling scoring model. After conferring with technical advisors, we believe that during the time period between initial funding and operational administration of the Consortium’s summative assessment

in the spring of 2015, the AI scoring methodology will be firmly established and recognized as a valid and reliable scoring platform. Should this assumption not be met, we will develop backup plans that involve higher rates of human scoring to support or supplant the AI process.

The Consortium recognizes and values the professional development opportunity inherent in the use of teacher scorers. We value teacher scoring because of its potential to help teachers internalize the performance standards and buy into the scoring process. It also can help build school-level capacity in evaluating student work and be used to validate AI-assigned scores. For the summative assessments, teachers will read 10% of the CR items and 33% of the performance events. In recognition of the stakes associated with the summative use of the performance events, we have provided for increased teacher scoring to allow for the scoring of certain dimensions that may not be as reliably scored by AI technology (e.g., inference). For the I/B assessments, 10% of the constructed-response items will be scored by teachers following AI scoring. Teachers will also score individual responses to performance events that cannot be adequately scored by AI. For the summative assessments, teachers will not be allowed to score the work of students from their own State. Where there are discrepancies between the teacher-assigned and AI-assigned scores, responses will be blindly scored by additional rater(s) until agreement is reached on two scores.

Building from existing methods for training human readers, the Consortium is committed to the development of a robust training system that will include orientation to the item or event, discussion of the scoring guide, review of exemplar papers for each score point, scoring practice based on training papers including feedback on scoring accuracy, and completion of a qualifying set. Scorers must achieve at least 80% accuracy (within one score point) on the qualifying set before being allowed to score student work for the summative assessment. This training and moderation can be accomplished initially through regional scoring centers that can bring teachers together to orient them to the scoring process. As capacity is built, we believe that the SBAC web-based portal will support distributed training, calibration, and scoring.

While we envision that scoring of constructed-response items will be conducted with a single-content-domain rubric, each performance event will be scored with a multiple-content-domain rubric to provide for student work on the different dimensions of a task. By scoring with multiple-content-domain rubrics within a single performance event, we intend to capture the

richness of the tasks and provide for increased reliability with a larger number of score points per event. This multi-content domain rubric will provide a profile of performance across all events to be reported in addition to the score for each event and an overall composite score that meaningfully weights scores on the performance events with the score from the adaptive summative component. The relative contribution of performance events to the composite score will be determined following comprehensive review by the Consortium in collaboration with technical advisors and will be informed by ongoing research.

Ongoing monitoring of scoring

Ongoing scoring accuracy will be monitored through check sets and reader drift monitoring. Once scorer drift has been identified, the teachers will receive reader retraining, involving scoring leadership and recalibration as needed. Thresholds on quality indices will be set to trigger reader retraining, dismissal, and rereading of papers read by a retrained or dismissed reader.

Reporting

(A)(4)(d) SBAC is committed to developing an assessment system that reflects responsible flexibility. We seek to allow States to implement the components of our assessment system in a way that best meets their needs. For this reason, we propose to implement a reporting plan that affords an appropriate balance between standard and customized reports but also serves the various purposes of the reporting system. (See section (A)(3)(c)(vii) for a description of our specific reports for each component of our assessment system.)

Specifically, the reporting system will be developed to meet the following purposes: (a) managing an integrated assessment and accountability system that meets the needs of participating States; (b) providing “early warning” information to monitor curriculum, instruction, and learning; (c) making timely and informed improvements in curriculum and instruction; (d) supporting professional development; (e) providing meaningful achievement information to all stakeholders, including IHEs; and (f) providing comparable information about student achievement for accountability at the local, State, and Federal levels. As described previously, our system will feature tools for producing customizable reports that track and analyze the progress toward college- and career-readiness of individual students, classes, and student subgroups (e.g., SWDs and ELs). It will also employ the following design features:

- A common electronic platform that allows for better management of assessment data and significantly accelerates the speed with which assessment information is made available to stakeholders. The reporting module will have elements that are both dynamic (assessment results linked to units, lesson plans, and curricular resources) and static (results linked to accountability and growth models).
- A set of data analysis and report-generating tools, built on the model of the NAEP Data Explorer, allowing for the development of customized reports that display data through a variety of tables and graphic formats, and drill-down capacity during data analysis (see Appendix A4-4e for an example).
- In keeping with research-supported best practices, an array of trustworthy information about student achievement and growth, reporting results that are technically sound and consistent with the technical parameters and limitations of the data (see, e.g., Zenisky, Hambleton, & Sireci, 2009).
- A variety of reporting interfaces and reports differentiated by audience (student, teacher, parent, school or district administrator, SEA, or Consortium) with tiered security to limit access to authorized users.

The reporting system will provide both standard paper-based reports and more technologically advanced web-based data analysis tools.¹⁴ Base reporting of summative assessment results will be common across States for comparison purposes, but schools, districts, and States will have the capacity to conduct independent data analyses and research studies to answer the key accountability and student achievement questions relevant to their contexts. This innovative system will include links to model curriculum, instructional, and assessment resources aligned with State and local content, and professional development resources related to data use. Information on allowable accommodations and instructional supports will be accessible to teachers and test administrators as they prepare for instruction or assessment. District and school administrators and teachers will be trained in the interpretations of assessment results so as to become more informed users of assessment data.

Quality control measures

(A)(4)(e) SBAC’s approach to quality control (q/c) recognizes that producing error-free,

¹⁴ We provide a number of exemplars in Appendices A4-4a through A4-4e that suggest our current direction for this work.

technically sound, innovative assessments is a function of two key components: (1) development of interconnected systems for each phase of development, implementation, and evaluation, coupled with (2) a monitoring plan with built-in redundancies to ensure full compliance with these systems from all parties involved (member States, the support partners [e.g., Project Management Partner, vendors] and school staff and their students). SBAC will select the Project Management Partner and vendors, in large part, based on their inclusion of q/c plans and approaches as part of their proposals and their track record implementing such plans. Several organizations and authors have published compendia of quality control guidelines relevant to large-scale assessment, and we plan to draw heavily on these (e.g., see Allalouf, 2007; International Test Commission, 2005; and CCSSO, 2003).

Our experience is that errors are most likely to occur in the following four key assessment activities.

- *Item/form development*: Errors occur at this stage in two distinct ways. The first involves failure to adequately track item history from original drafting through editing, bias review, field testing, and production. Our item management system will maintain full item/performance event histories and will require management signoff at each step. Other errors occur from inadequate proofing at the production stage. At the “form” level, CAT-based assessments will require ongoing monitoring to ensure that the adaptive engine produces student-level data that meet the assessment blueprint. We will monitor this throughout the assessment window.
- *Scoring*: Both the human scoring and AI scoring approaches will use the same general q/c approaches—training and read-behinds. Human training will involve the selection of exemplars for use in training and calibration and a threshold of 80% scorer accuracy to qualify for scoring. All scoring (human and AI) will be checked with 10% read-behinds. Given the large number of participating States, this will result in the review of thousands of student work samples.
- *Equating*: Consistent with best State practices, SBAC will check our equating vendor by requiring third-party independent verification of all equating procedures and results.
- *Reporting*: Reporting shells will be carefully reviewed for accuracy and utility. As much as possible, we will track individual student responses all the way through the reporting stage.

Field testing

Field testing is an essential step to ensure technical adequacy and fairness to all student populations. SBAC will incorporate a strategy that begins with strategic piloting and ends with a comprehensive “all member State” representative field test to support

- development of item parameters to support the CAT design;
- review of CR items and performance events to ensure that the full range of responses are attainable and that each item is scorable based on its rubric; and
- inclusion of all student populations in the sample to ensure fairness and full access.

Specifically, we will include the following steps:

- *Small-scale pilot testing:* As items and performance events are developed, we will conduct small-scale tryouts to ensure they are functioning as intended. For selected-response (SR) items, convenience samples of approximately 500 students will be obtained to determine whether the correct answer and distractors are operating as intended. For CR items, we will examine whether each score point on the rubric is attainable. For performance events, the Consortium will study whether teacher and student supports are adequate and whether there are multiple scaffolded entry points for students across the achievement continuum. To purposefully support the development of accessible assessments for ELs and SWDs, prototype items and performance events will be piloted with these populations and evaluated through a method of micro-experimentation (within-subjects design) and revision, with the most promising variations further tested through a larger-scale field trial across the Consortium States. The effectiveness of these design variations will be evaluated by examining the contribution of specific task features and accommodations to the performance of students from special populations.
- *Field testing:* In 2013–14 and 2014–15, SBAC will conduct a full field test of all items and performance events in the adaptive summative pool that survive the content, bias, and sensitivity review processes. To ensure the validity of findings, all Consortium States

will be expected to participate and will be provided demographic targets that they must meet.¹⁵

- *Accommodations and the full range of student populations:* For SBAC, “all” truly means all. Our piloting and field test procedures will oversample SWDs, ELs, and urban and rural students to ensure that their assessment needs are met and they are not disadvantaged by any item or performance event. The Consortium States have submitted an Enhanced Assessment Grant (EAG) to fully study whether proposed accommodations for SWDs and ELs are sufficient. (See sections (A)(4)(a) and (A)(4)(b)(i) and (ii) for a full description of our strategy to ensure access for all students and Appendix A4-3 for the EAG grant proposal goals.)

¹⁵ Preliminary cut-scores will be determined based on review of field test data during this standard setting session. These cut-scores will be verified based on operational administration data collected in Spring 2015.

(A)(5) Research and Evaluation (up to 30 points)

The extent to which the eligible applicant's research and evaluation plan will ensure that the assessments developed are valid, reliable, and fair for their intended purposes and for all student subgroups. In determining the extent to which the research and evaluation plan has these attributes, we will consider—

- (a) The plan for identifying and employing psychometric techniques suitable to verify, as appropriate to each assessment component, its construct, consequential, and predictive validity; external validity; reliability; fairness; precision across the full performance continuum; and comparability within and across grade levels; and
- (b) The plan for determining whether the assessments are being implemented as designed and the theory of action is being realized, including whether the intended effects on individuals and institutions are being achieved.

Framework for Research and Evaluation Plan

The Consortium's research and evaluation plan is grounded in the principles adopted by the Joint Committee on Standards for Educational Evaluation in the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999). The Research and Evaluation Working Group will ensure that the assessment system has integrity and is characterized by high technical quality, that the measurement components are trustworthy and fair, and that results are useful for the purposes intended.

Our research and evaluation plan also follows a reasoning-from-evidence approach (NRC, 2001) in collecting and evaluating documentation to support claims of validity, ensuring that different types of evidence (including content, construct, predictive, and consequence-related validity and measurement reliability and precision) are collected on an ongoing basis during all phases of design, development, and implementation. This model requires close collaboration with the Consortium's Technical Advisory Committee (TAC) as well as IHEs, workplace representatives, and content experts in ELA and mathematics who have an in-depth understanding of the range of expectations in the Common Core State Standards (CCSS) and the college- and career-ready outcomes specified in those domains.¹⁶ Of particular importance will be validation of the cognitive model (i.e., construct representations, developmentally sound learning progressions over time, and benchmark learning targets at each grade) on which each

¹⁶ For example, Conley (personal communication, 2010) recommends that the Consortium focus on measuring a key set of cognitive strategies (problem formulation, research, interpretation, communication, and precision and accuracy) and self-management skills (time management, goal-setting, self-awareness, persistence, and study skills) that have been shown to be critical for success in college courses and technical certificate programs.

component is based and consideration of evidence that each item/performance event operates as intended across all student performance levels. All components of the assessment system (i.e., adaptive summative tests and performance events used to measure achievement and growth and optional interim/benchmark assessments and formative tools¹⁷) will be examined during these evaluation processes.

In addition, the research and evaluation plan will support the Consortium’s key pillars of balance, comparability, and flexibility. In keeping with SBAC’s Theory of Action, the research and evaluation agenda will support the use of assessment data for rational interpretations of student achievement and growth and for the improvement of teaching and learning. It will monitor key indicators of system performance at different levels; provide feedback to the Steering Committee, technical advisors, and participating States about coherence across system components; and inform Consortium-level decision-making about tradeoffs—particularly from a psychometric perspective—associated with system flexibility (e.g., distributed administration window, options for measuring growth and teacher effectiveness, on-demand access to item pool). Active and ongoing attention to test-based consequences (both intended and unintended, positive and negative) will be a priority. Through this process, the Research and Evaluation Working Group will ensure documentation of steps in implementing system components, collect evidence to inform refinements to the system during roll-out, and monitor overall system effectiveness.

Psychometric research and evaluation activities will be carried out for the high-stakes summative assessments (achievement and growth measures) and the optional interim/benchmark assessments. The Consortium is committed to the use of industry-standard psychometric techniques during all phases of system development, including planning, design and development, small-scale pilot testing, ongoing field testing, score scale development, operational administration, setting of performance/achievement standards, and post-administration data review. The TAC will provide expertise and advice to the Steering Committee and the Research and Evaluation Working Group through face-to-face meetings twice a year and electronic communication and conference calls, as needed.

¹⁷ The interim/benchmark and formative components are described in the MOU as optional “formative/benchmark components,” which is a broader description for these system features.

Measures of Achievement and Growth: SBAC’s Summative Assessment

All participating States will administer the Consortium’s summative assessments to students in grades 3–8 and high school. Data from the summative assessments will be used to measure annual achievement and growth, to inform evaluations of teacher and principal effectiveness, and for Title I accountability purposes.

(1) **Validity and Fairness.** The evaluation plan and research agenda will address whether the computer adaptive platform and innovative item pool together support the use of these assessments for high-stakes purposes. We will ensure that the underlying computer adaptive testing (CAT) engine exposes each student to an item set that measures the full range of the CCSS in a manner tailored to the student’s current achievement level and grade level expectations. In addition, we will monitor ongoing item/performance event pool development and examine the degree to which items and performance events address the full range (depth and breadth) of the CCSS; elicit specifically targeted knowledge, skills, and/or cognitive processes related to college- and career-readiness by using a reasoning-from-evidence approach during item/event development; are developmentally appropriate and support accessibility so all students can show what they know and can do during testing; and can be combined to create (a) a composite score that defines a meaningful weight for each component (adaptive score and performance events) in relation to the intent of the CCSS and (b) a detailed score profile for each student that describes that student’s progress toward college- and career-readiness on a number of dimensions.

During the development phase, through open-ended surveys, interviews, and focus groups, we will collect qualitative data about the items and performance events from students and teachers. We will monitor field testing of items and performance events and calibration using three-parameter logistic (3-PL) Item Response Theory (IRT) frameworks, with the addition of dimensionality analysis for defining scale granularity and multidimensional IRT where feasible and necessary. The Consortium will contract for independent alignment studies of items in the item pools at each grade. The research and evaluation team will use item-level analyses to examine the validity of claims of strong item-to-content and item-to-cognitive complexity match. The Consortium will be guided in item pool refreshment by findings from these studies and by the broader research agenda described at the end of this section.

Table A5-1 outlines the activities that will be used to monitor validity and fairness for the summative assessments, and Table A5-2 details the evaluation questions that will guide this work and planned data collection and analysis strategies.

Table A5-1. Specific Steps Related to Collection of Evidence of Validity and Fairness

Validity Topic	Beginning Steps (design & development)	Transition Steps (implementation)	Long-Term Steps (evaluation & research)
Content, Construct, and Predictive Validity	<ul style="list-style-type: none"> • Monitor documentation of design and development activities • Identify, <i>a priori</i>, plausible intended and unintended consequences • Ensure strategies are used to avoid unintentional introduction of construct-irrelevant variance • Oversee independent alignment studies • Review contingency plans • Ensure a broad representation from State educators and IHEs and workplace representatives in all activities • Conduct pre-equating analyses 	<ul style="list-style-type: none"> • Monitor documentation of implementation steps • Survey States and stakeholders regarding emerging unanticipated, unintended consequences • Monitor impact of weighting of performance events in composite score • Monitor fidelity of implementation of accommodations • Monitor fidelity of implementation to the Theory of Action • Observe test administrations • Monitor AI scoring (validity) and provide feedback for system improvement (i.e., formative) regarding scoring methods 	<ul style="list-style-type: none"> • Evaluate degree to which system components work together as intended • Evaluate consequences of testing (intended and unintended, both positive and negative) • Conduct utilization and cost-benefit analyses to determine whether findings support continued use of system components • Monitor impact over time of scale drift, changes in alignment, and scoring methods • Support IHE research • Conduct equating analyses to determine the comparability of scores over time
External Validity (Generalizability to College- and Career-Readiness)	<p>Include IHEs and representatives from the workplace in:</p> <ul style="list-style-type: none"> • Defining college- and career-readiness • Developing benchmarks in grade 3–8 in pathways to college- and career-readiness • Developing items for summative and I/B item/ event banks • Alignment activities 	<p>Include IHEs and representatives from the workplace in:</p> <ul style="list-style-type: none"> • Setting performance standards linked to college- and career-readiness • Monitoring emerging consequences 	<p>Include IHEs and representatives from the workplace in evaluating the predictive validity of performance standards in grades 3–8 (on track to college- and career-ready) and high school (college- and career-ready)</p>

Validity Topic	Beginning Steps (design & development)	Transition Steps (implementation)	Long-Term Steps (evaluation & research)
Fairness	<ul style="list-style-type: none"> • Ensure population specialists are involved in all activities • Monitor adherence to UD principles • Share trustworthy findings from studies (e.g., cognitive interviews with students) of different item types (accessibility, feasibility, comparability) with developers • Review accessibility requirements for all items and performance event types • Oversee development of test security guidelines 	<ul style="list-style-type: none"> • Monitor impact of testing at the subgroup level • Study fidelity of implementation of allowable accommodations during testing across States • Monitor emerging effects (intended and unintended) of test use • Monitor adherence to test security guidelines across States • Observe test administrations • Conduct stakeholder surveys and provide feedback for improving test accessibility for special populations 	<ul style="list-style-type: none"> • Conduct validity and reliability analyses at the subgroup level • Design protocols for ongoing monitoring of DIF and/or detecting evidence of bias • Monitor impact of college- and career-ready performance standards on special student populations over time • Support IHE research • Evaluate overall system security

(2) **Measurement reliability and precision.** By analyzing performance data during small-scale pilot testing, field testing, and operational administrations of the summative assessments, the Research and Evaluation Working Group will examine whether the items/events work as intended and the degree to which use of the item pools and adaptive engine work together to

- efficiently and effectively select items for each student based on that student’s prior responses and identify that student’s achievement level on the construct assessed;
- provide high and comparable degrees of precision in identifying students’ level of college- and career-readiness across a broad performance range; and
- support valid inferences about meeting achievement standards.

Further, the Research and Evaluation team will take the following steps to monitor the reliability of all scoring systems:

- Consult technical advisors in development of both achievement and growth scales and monitor characteristics over time;

- Consult technical advisors in decisions related to psychometric calibration, scaling, and equating models (e.g., multi-group calibration and linking);
- Conduct internal consistency analyses to estimate (1) population and sub-population reliabilities and (2) measurement error across the scales and at particular decision points and provide rationale for statistical indices used to evaluate disparate impact for subgroups;
- Collect data about classification accuracy *and* consistency based on achievement and growth measures; and
- Conduct optional transitional linking from existing State scales to SBAC scales.

(3) ***Item types and scoring reliability.*** The summative and I/B assessments will make use of technology-enhanced item types and performance events. During small-scale pilot testing, field testing, and operational administrations, the Research and Evaluation Working Group will monitor the

- reliability of automated and/or AI scoring of selected-response items, constructed-response items, technology-enhanced items, and performance events;¹⁸
- reliability of the scoring systems to ensure reliable data collection and to safeguard against technological problems in data collection; and
- development of scoring guides that detail the educational intent of each item type and how each works to collect information about students' levels of cognitive complexity/critical thinking skills.

Further, to support this work, we propose a research and evaluation process that will support

- implementation of quality control and test security measures to (a) maintain the integrity of scores over time and across students, schools, districts, and States; (b) monitor administration conditions; and (c) encourage ethical testing practices; and
- development of innovative test reports to ensure transparency with stakeholders about the psychometric strengths and limitations of each measure and facilitate appropriate interpretation of results and use of data.

¹⁸ To allow for rapid turnaround of results, the scoring process will capitalize on emerging technology as feasible. All items scored via artificial intelligence (AI) will be subjected to a systematic read-behind process.

Strategies for Measuring Growth

Growth models focus on changes in performance of individual students (and/or the aggregate of individual growth at the school or district level) over time and/or across performances. To support measures of growth, the Consortium will build vertical scales and vertically articulated grade- or grade-span-specific achievement scales and will be prepared to set standards for grades 3–8, using “on-track to college- and career-readiness” as the vertical articulation criterion.

However, unless a single growth model is required for accountability upon Title I reauthorization, the Consortium proposes to allow States flexibility in selecting the approach for modeling student growth that is most appropriate for their contexts. If required by the funder, a Consortium-wide growth model can be quickly implemented by applying any one of the existing models to all States’ data. But we propose instead, as described in the Research Agenda below, that the Consortium actively pursue empirical studies of the characteristics of the different models when used in conjunction with data from CAT so that any single Consortium-wide expectation for growth reporting can then be based on findings from emerging research. An anticipated result of the research and evaluation agenda will be the selection of the maximally valid measure for Consortium-wide adoption in the future. Allowing such flexibility is critical to implementing the Consortium’s Theory of Action. Its implications will be closely monitored by the Research and Evaluation Working Group.

An immediate adoption of a single growth model for use at the Consortium level is inadvisable for specific reasons. First, there is significant scholarly disagreement about the appropriateness of vertical scales for measuring student growth—with some researchers claiming that vertical scales can provide valid linear measures of student growth (e.g., Bond & Fox, 2001; McCall & Hauser, 2007) and others claiming that use of a vertical scale may compromise the validity of interpretations about student growth drawn from such measures (e.g. Betebenner, 2009; Dunn & Allen, 2009; Reckase, 2006; Schmidt, Houang, & McKnight, 2006; Martineau, 2006). In addition, each growth modeling approach is associated with specific strengths and limitations. Therefore, the Consortium proposes that States serve as laboratories to help SBAC examine the tradeoffs of various models in real-world contexts.

In addition, prior to implementation of a required model, consensus must be reached to make key decisions about what type of growth is best measured with CAT, how much growth will be considered sufficient at each grade level, and whether students with different starting points

should be expected to grow at the same rate (Gong, Perie, & Dunn, 2006). None of the measures of growth inherently defines how much growth is enough—each simply reports student growth without value judgments. While SBAC proposes to allow flexibility in choice of model at this time, we will provide research-supported guidance to States to inform decision-making about (a) placing values on specific ranges of growth to determine appropriate growth targets for individual students and (b) implementing measures of educator effectiveness as appropriate for the context. Importantly, however, this flexibility does not preclude Consortium-wide use of growth data for the purpose of informing teacher and principal evaluations, or other desirable aims.

Comparability

The Consortium recognizes that in proposing the implementation of an adaptive summative assessment across more than 30 States we are facing unique challenges. We will have many steps to oversee and many more contexts to monitor. High-stakes assessments used for accountability purposes by a Consortium necessitate greater attention to comparability across diverse States, especially given the types of potential flexibility (e.g., in administration options such as timing and use of translations) needed to support enactment of SBAC’s Theory of Action.

Responsible flexibility requires careful monitoring by the Research and Evaluation Working Group, in collaboration with the Consortium’s Steering Committee and TAC. The Consortium is committed to ongoing study to examine whether or how much flexibility compromises the validity of comparisons across schools, States, or consortia. At a minimum, with the flexibility described by the system design, there will be within-year and across-year studies of comparability that examine the percentages of students achieving the end-of-year performance standard at each grade level and across key subgroups in each State.

Ongoing research will be focused on determining the extent to which comparability holds, both across participating States and across consortia, for achievement scales, measures of student growth, preliminary measures of teacher and principal impact on student growth, test translations, and preliminary mode of administration. We will collect data to empirically study comparability over time, using such indicators as similarity of equating results, dimensionality of student scores, and relationships among content clusters. We will further examine cross-grade coherence in performance expectations and cut-scores, and support ongoing data sharing by

capitalizing on the open source data framework. We will monitor specific comparability issues related to construct-irrelevant variance, principles of universal design for assessment, accommodations, translations, and other inclusive practices.

Beyond determining the levels at which comparability claims can be validly made, we will also work with (a) professional development providers, to ensure that interpretive guidelines are disseminated and understood by State educators, and (b) representatives from IHEs and the workplace, to ensure that these guidelines are disseminated and understood by policymakers and end-users of college- and career-readiness indicators. In addition, we will work cooperatively with other consortia on possible scenarios that would allow meaningful cross-consortia comparisons. To further address concerns about comparability, the Consortium proposes that the Research and Evaluation Working Group work with the Report Working Group to ensure the development of guidelines for the types of cross-State performance comparisons that are more *and* less appropriate in this novel context.

Interim/Benchmark Assessments

Results from optional interim/benchmark (I/B) assessments are intended to provide more finely grained information about student progress toward college- and career-readiness than that obtained from the comprehensive summative assessments. To ensure that this component serves key purposes related to improved teaching and learning described in our Theory of Action, many decisions about use of the items—including administration window—will be determined on a State-by-State basis. The CCSS-aligned items in this non-secure pool can be grouped into customized blueprints that meet different needs at the State or local level. Because the intent of the I/B assessments is to directly support teaching and learning and these assessments will not be used for accountability purposes, the priorities for evaluation and research will emphasize ongoing research to improve the quality, trustworthiness, and usefulness of data from these assessments for instructional purposes.

(1) ***Validity and fairness.*** The I/B assessments will be developed with a foundation in cognitive theory and research about how students learn over time. For this reason, the Consortium’s short- and long-term research and evaluation priorities include examining the degree to which

- the I/B assessments are grounded in cognitive development theory about how

learning progresses across grades and competence develops over time; in that respect, the generalizability of learning progressions across various student populations (e.g., high vs. low achieving, students with disabilities [SWDs] and English learners (ELs), within and across States, will be of particular interest;

- the I/B assessments—in keeping with the Theory of Action—elicit specifically targeted knowledge, skills, and/or cognitive processes related to college- and career-readiness, per the Theory of Action, by using a reasoning-from-evidence approach during item/event development that is grounded in understanding about how students acquire competence or develop expertise in a content domain;
- the I/B item pool and banks address the full range (depth and breadth) of the CCSS;
- the items and performance events in the I/B pool—as with all items/events in the summative pool—are developmentally appropriate and support accessibility (i.e., follow universal design principles) so that all students can show what they know and can do during testing;¹⁹ and
- I/B item specifications or assessment frameworks are developed to guide the construction of the clusters of items/events that are appropriate for this pool.

During the development phase, through open-ended surveys, cognitive interviews, and focus groups, we will collect qualitative data about the items and performance events from students and teachers. On an ongoing basis, the research and evaluation team will examine the degree to which results from customized I/B assessments are appropriate for evaluating students' interim progress toward postsecondary readiness, predicting performance on the annual comprehensive summative assessment, diagnosing students' strengths and limitations, or guiding formative data use by teachers during instructional planning and implementation of classroom instruction.

(2) *Measurement reliability and precision.* The Research and Evaluation Working Group will examine (1) the degree to which the I/B pool of items and performance events is sufficiently large and diverse to measure small, incremental differences in what students know and can do and (2) the effectiveness of innovative item types such as technology-

¹⁹The Consortium is expected to enlist and train teams composed of independent content experts, IHE and workplace representatives, teachers, and population specialists in the assessment needs of English learners and students with disabilities to review all items in the interim/benchmark pool for appropriateness.

enhanced items and multi-faceted performance events in drilling down deeply into a student's level of understanding of those few key concepts or skills that define a particular learning progression. While the number of items or performance events presented to any one student during adaptive I/B testing and adaptive summative testing may be the same, to allow for the in-depth probing intended, the *types* of items that appear will be more varied, and the *range* of content assessed will be greatly reduced during I/B testing. These assumptions will need to be tested empirically and monitored to ensure technical adequacy for the purpose intended.

(3) *Item types and scoring reliability.* In addition to selected-response items, these measures will include a greater number of technology-enhanced items and performance events. In this way, the I/B item pool will serve as an opportunity to field test the feasibility of innovative item types that may also be appropriate for the summative item/event pool. Because the I/B assessments are designed to serve a different purpose than the summative assessments, the focus will be on

- exploring which scoring methods (fully automated, teacher-supported, or fully teacher-scored and moderated) allow for the most efficient and effective use of results;²⁰
- collecting inter-rater reliability data for teacher-scored items and performance events to be used formatively to improve scoring processes; and
- collaborating in the development of rubrics and scoring guides that may help teachers better understand how each item collected information about what students know and can do in relation to learning progressions and the pathways to college- and career-readiness.

Plan for evaluating the assessment system

To enable the completion of these responsibilities, the Consortium proposes to implement (1) short-term steps and contingency plans; (2) State-specific interim/transition plans, including linking studies to support the transition from old systems to new; and (3) long-term utilization, cost-benefit, and impact analyses. Initially, the evaluation process will inform the design, development, and implementation of the assessment components. Later we will collect data

²⁰ We can use this as an opportunity to test different scoring strategies.

related to important indicators of the effectiveness of the assessment system. Evaluation questions will stem from initial design considerations as well as from continual input during implementation of each component.

The Consortium's Research and Evaluation Working Group will oversee formative and summative evaluation processes. Each year the group will provide mid-term reports to the Steering Committee that may be used formatively to support strategic decision-making and advise refinement to course direction. It also will submit annual summative evaluation reports to the Steering Committee that will focus on emerging and/or chronic effects, the effectiveness of system functioning, and status updates on the research studies. Each participating State will be expected to use the same template for identifying and collecting data for this purpose and for documenting implementation steps. Each participating State also will be expected to cooperate with the evaluators during all data collection efforts, including cross-State research studies. The formative and summative reports to the Steering Committee will include consolidated comments and outcomes from all participating States.

Further, the Consortium's Research and Evaluation Working Group agrees to report and disseminate findings from evaluations both formally (e.g., in peer-reviewed journals or through invited or accepted presentations at national conferences) and informally (e.g., in newsletters, white papers, and policy briefs). The Consortium also will participate in any technical assistance activities conducted or facilitated by USED or its designees and work with the Department to develop a strategy to make student-level data available on an ongoing basis for cross-State or cross-consortia research activities.²¹ Finally, the Research and Evaluation Working Group will oversee maintenance of a document repository for materials such as technical reports, evaluation reports, research reports, business and legal documents, and State- and Consortium-wide documentation of processes, procedures, and protocols.

Table A5-2 provides an overview of the types of evidence that will be collected on short-term, interim, and long-term bases.

²¹ The Consortium will work with the Department and while following all guidelines in the Family Educational Rights and Privacy Act (FERPA) and 34 CFR Part 99, as well as State and local requirements regarding privacy.

Table A5-2. Description of Data Collection Strategies and Evaluation Questions, by Phase

	Beginning Steps (2010–12): Development Phase	Transition Steps (2010 through 2014–15): Initial Implementation Phases	Long-Term Steps (2012–20)
Focus of Evaluation During this Phase	To monitor and document steps taken during development to ensure that one or more planned components (1) will measure the full range of CCSS in ELA and mathematics, including those standards that traditionally have been difficult to measure; (2) are valid for the purpose of measuring individual student achievement and growth over a full academic year or course at the subgroup, ²² classroom, school, LEA, State, and Consortium levels; (3) are appropriate for <i>all</i> students, including ELs and SWDs; (4) are valid and reliable	To monitor implementation fidelity, including scoring and reporting functions	To evaluate if and to what degree the system is effective, i.e., the intended effects are realized as described in the Theory of Action, the integration of balanced assessment is supporting educator decision-making in participating States, the system overall has had a positive impact on teaching and learning, and unanticipated and unintended consequences (positive and negative) that emerge over time are considered during decision-making about continued use of each system component

²² Eligible applicants receiving funds under this competition must aggregate data using the student subgroups in section 1111(b)(3)(C)(xiii) of ESEA (i.e., by gender, by each major racial and ethnic group, by English proficiency status, by migrant status, by students with disabilities as compared to nondisabled students, and by economically disadvantaged students as compared to students who are not economically disadvantaged, except that such aggregation is not required in a case in which the number of students in a subgroup is insufficient to yield statistically reliable information or the results would reveal personally identifiable information about an individual student). When using the term “subgroup” throughout this application and the NIA, we mean these student subgroups.

	Beginning Steps (2010–12): Development Phase	Transition Steps (2010 through 2014–15): Initial Implementation Phases	Long-Term Steps (2012–20)
	for the purposes of determining whether individual students in grades 3–8 are <i>on track to being college- and career-ready</i> ²³ and whether high school students are on track to being college- and career-ready and/or <i>are college- and career-ready</i> ; ²⁴ (5) will yield data that describe the full continuum of student performance, including the highest- and lowest-achieving students, based on Consortium, defined achievement standards and performance level descriptors ²⁵ for college- and		

²³That is, the student is performing at or above grade level such that the student will be college- and career-ready (as defined below) by the time of high school graduation, as demonstrated by an assessment score that meets or exceeds the Consortium’s achievement standard for the student’s grade level on a summative assessment in mathematics or English language arts.

²⁴That is, the student is prepared for success, without remediation, in credit-bearing entry-level courses in an IHE, as demonstrated by an assessment score that meets or exceeds the Consortium’s achievement standard for the final high school summative assessment in English language arts or mathematics. In addition, the Consortium expects students to be ready to enter and advance in a job or succeed in advanced training for a high-skill, high-wage job; able to read, comprehend, interpret, analyze and locate complex technical materials; use mathematics to plan, set priorities, and solve problems in the workplace; and pass a State-approved industry certification or licensure exam in the field (SREB, 2009).

²⁵ That is, a statement or description of a set of knowledge and skills exemplifying a level of performance associated with a standard of performance.

	Beginning Steps (2010–12): Development Phase	Transition Steps (2010 through 2014–15): Initial Implementation Phases	Long-Term Steps (2012–20)
	career-readiness; (6) determine school effectiveness for purposes of Title I accountability; (7) inform decision-making about individual principal and teacher effectiveness for purposes of evaluation; and (8) inform decision-making about principal and teacher professional development and support needs		
Data Collection Strategies During this Phase	<ul style="list-style-type: none"> • cognitive interviews to try out different item types with students • small-scale pilot studies • field testing with strategic cross-State sampling • equating studies • expert validation of new item/event types and new types of reporting scales (incremental achievement and critical thinking scales) 	<ul style="list-style-type: none"> • cross-State collaboration on transition plans • focus group discussions with teachers in participating States • observation during testing • reliability analyses, post-operational DIF analyses, and examination of error 	<ul style="list-style-type: none"> • document analyses (e.g., proficiency level descriptors, guidelines for test use, scoring protocols, score reports) • examination of performance data at the subgroup, school, and State levels • sharing of data from IHEs and employers (e.g., results from placement exams or performance assessments) • feasibility analyses and cost-benefit analyses in which tradeoffs associated with continuing current practices are weighed • experimental or quasi-experimental design studies to empirically study system effects such as

	Beginning Steps (2010–12): Development Phase	Transition Steps (2010 through 2014–15): Initial Implementation Phases	Long-Term Steps (2012–20)
		<p>rates, omission rates, and distractor analyses</p> <ul style="list-style-type: none"> • collaboration and data sharing with IHEs, including focused field trials 	<p>systematic differences in curriculum or instruction across States or changes in student motivation or teachers’ practices over time</p> <ul style="list-style-type: none"> • surveys and focus groups with cross-State samples of teachers, school and district administrators, parents, and other stakeholders
Evaluation Questions Asked During this Phase	<p>Are test purpose and intended uses clearly specified?</p> <p>Does the validity argument adequately connect test scores and score-based interpretations? Is there evidence of an underlying plan linking theory of action and outcomes?</p> <p>Does the pool of items measure the full range of the Common Core State Standards? Was evidence collected to show that the items aligned to a certain standard capture the depth and breadth of the content/construct intended by each standard?</p> <p>Do students have the opportunity to learn tested content?</p>	<p>Are components being implemented as intended?</p> <p>Did flaws in the conceptualization of the test (i.e., test design) emerge?</p> <p>Was evidence collected that the measures are appropriate for the intended population? Were steps taken to eliminate items showing potential bias?</p> <p>Was evidence of alignment to CCSS documented?</p> <p>Were scoring processes monitored and data (e.g.,</p>	<p>Has new evidence emerged that calls into question current interpretation of test scores? Are students with passing scores on the high school adaptive summative assessment ready for college or careers?</p> <p>Has there been a decrease in the number of students enrolling in college who need remedial classes or an increase in the number of students earning postsecondary education or training credits early through dual enrollment? What is the success rate of students in classes considered non-remedial?</p> <p>What percentage of students are misidentified by the assessment system (as either career/college-ready or not), as measured by a postsecondary criterion (e.g., placement exam shows need for remedial college coursework or additional on-the-job training)?</p> <p>What percentage of students pass State-approved employer certification exams?</p>

	Beginning Steps (2010–12): Development Phase	Transition Steps (2010 through 2014–15): Initial Implementation Phases	Long-Term Steps (2012–20)
	<p>Were proactive or preventative steps (e.g., quality control and test security, bias/sensitivity reviews, professional development for test proctors) taken to avoid all plausible unintended consequences or persistent side effects? Were plausible unintended consequences such as score inflation or impact on student achievement in non-tested content areas (e.g., science and social studies) considered?</p> <p>How will scoring be moderated across schools and over time?</p>	<p>reliability estimates) collected to support continued use of those scoring methods for those item types and testing purposes?</p> <p>How were teachers trained for teacher-moderated scoring? How were they calibrated and recalibrated, if needed?</p> <p>Are results useful to SEAs for accountability purposes? Are results useful to LEAs for monitoring annual achievement and growth? Who else used the results?</p> <p>Are professional development opportunities available to teachers and administrators?</p> <p>Was sufficient validity evidence for the</p>	<p>How do results from the summative assessment compare with results from NAEP preparedness measures? With ACT or SAT college readiness benchmarks? With Advanced Placement test results?</p> <p>What is the relationship between scores on the assessments and local grades?</p> <p>Are the assessments cost effective? Do they drain resources without adding significant value? Is there sufficient justification to continue the assessment without significant modification?</p> <p>Do the positive consequences —both intended and unintended—of using a measure outweigh the negative? Are the benefits greater than the costs?</p> <p>Have important changes occurred recently in instructional practice?</p> <p>Because consequences generalize from the student level to other levels, was a differential impact noted at the subgroup, school, LEA, and SEA levels? Has evidence of differential impact across student subgroups emerged over time?</p> <p>Has the achievement/performance gap narrowed?</p>

	Beginning Steps (2010–12): Development Phase	Transition Steps (2010 through 2014–15): Initial Implementation Phases	Long-Term Steps (2012–20)
		<p>performance events collected and documented, including their alignment and/or transferability to real-world situations? Was the range of administration conditions considered?</p> <p>Were uniform accommodation guidelines followed?</p>	<p>Has anything in the testing context changed recently? Are changes in policy or practice compromising the original intent of the system or component?</p> <p>Does emerging research raise any concerns about current practices or call into question past assumptions? Have any issues relating to the ethics and justice of the consequences of the assessment emerged?</p> <p>Were unintended uses of test results uncovered? Has clear evidence of test misuse emerged?</p> <p>Did unintended positive consequences emerge from test use? Did unintended negative consequences emerge from test use?</p>

Research Agenda

In this section, we describe the Consortium’s ambitious research agenda that will allow us to systematically examine empirical questions—raised, prioritized, and vetted by the Steering Committee and participating States—that are critical to informing decision-making about next steps in the development and implementation of a truly innovative and effective comprehensive assessment system.

The planned research agenda is intended to be sufficiently robust to support gradual infusion of theory- and research-supported innovation and continuous system improvement. Based on consultation with State and national experts on practical and theoretical issues, this evaluation plan will include aggressive programs of research in the following areas:

Component and System Development

- The development and validation of a distributed summative assessment system (through-course assessments), including (a) achievement scaling on the complete assessment scale based solely on the unit assessed, (b) predicting the probability of success on a complete assessment, (c) studying the construct measured (retention vs. acquisition), and (d) comparability of prescribed-window summative (non-distributed) versus distributed summative assessment results.
- Research on growth modeling (in conjunction with IHEs), such as (a) learning progression development and validation, (b) tying growth measures to learning progressions, (c) deciding how much growth should be expected at each grade (what is “good enough?”), (d) tying growth measures to formative tools (for targeting professional development), (e) differential functioning of multiple growth measures, (f) growth measures for distributed and/or interim/benchmark assessment, (g) growth as measure of gains in acquisition or retention, (h) growth measures of higher-order thinking skills, and (i) feasibility and reliability of advanced automated scoring.
- Characteristics (or specifications) of innovative technology-based performance tasks or events, especially those that measure the integration of academic content and technical skills (e.g., iterative problem solving with multiple solutions, sustained literacy or oral

communication tasks, extended inquiry using web-based resources and media, simulations)²⁶ and timing of administration.

- Systems for performance events that may be developed by the Consortium but sustained by the States.
- The development, validation, and implementation of a scale designed to measure higher-order thinking skills, application-based learning, and the ability to transfer knowledge and skills to unique contexts.

Comparability, Consequential Validity, and Score Interpretation

- Appropriateness and feasibility of using interim/benchmark measures for purposes of (a) measuring growth in high school, when annual summative measures are not required, and (b) informing teacher or principal evaluations when annual summative measures are not required.
- Usefulness of interim/benchmark assessments in supporting teaching and learning, either in single administration format (may include end-of-course assessments if items are secure/restricted access) or distributed administration format (curriculum-embedded, on-demand access).
- Standard setting at all grades tied to postsecondary readiness that (a) seeks to identify methods for setting cut-scores on an adaptive high school assessment linked to college- and career-readiness that will be considered sufficiently rigorous to be acceptable to IHEs and career training institutions as a basis for admission and placement decisions, and (b) allows for periodic reviews of cut-scores to assure that what is labeled “on track” in one grade continues to predict “on track” performance in the next grade level, from grade to grade through the high school test.
- Impact of different performance event weighting strategies on a composite score.
- Study of precision and usefulness of different measurement models for capturing educator impact (teacher effectiveness) on student achievement and growth.
- Dimensionality and its impact on achievement and growth inferences.
- Comparability issues within a system, with flexible participation options within and across

²⁶ For additional ideas and examples, see Quellmalz, Timms, & Buckley, 2009.

grade levels, jurisdictions, and consortia, given (a) options for various achievement scales, (b) options for various measures of student growth and associated reporting tools, (c) options for various measures of teacher and principal impact on student growth, (d) options for mode of administration (at least through 2016–17), (e) variation in item types, and (f) inclusive practices (e.g., accommodation, translation).

- Comparability of summative translated item set and English-only set; also implications of providing translations only for most populous subgroups.

Quality Control, Measurement Reliability and Precision

- For adaptive testing, evidence of precision across the continuum of achievement, with attention to (a) population-wide point-estimates of reliability, (b) population-wide conditional (score-specific) estimates of reliability, (c) individual student and subgroup reliabilities (e.g., conditional standard error measures), (d) the minimum thresholds for precision in reporting, and (e) reasonable stopping rules for situations in which the desired precision cannot be reached (e.g., inestimable or nearly inestimable achievement levels from [near] perfect response or [near] totally incorrect response strings).
- Reliability of AI scoring of items and performance events on the summative assessments.
- Reliability of teacher-moderated scoring of performance events both over time and across events, students, classrooms, and States.

(A)(6) Professional Capacity and Outreach (up to 15 points)

The extent to which the eligible applicant’s plan for implementing the proposed assessment system is feasible, cost-effective, and consistent with the theory of action. In determining the extent to which the implementation plan has these attributes, we will consider—

- (a) The plan for supporting teachers and administrators in implementing the assessment system and for developing, in an ongoing manner, the professional capacity to use the assessments and results to inform and improve instructional practice; and
- (b) The strategy and plan for informing the public and key stakeholders (including legislators and policymakers) in each member State about the assessment system and for building support for the system from the public and those stakeholders.

Building the professional capacity of teachers and school administrators is central to effecting the level of reform called for in our Theory of Action. For that reason, SBAC proposes a comprehensive plan to involve these two groups in all phases of the development and implementation of our balanced assessment system and to support them in identifying and/or developing and using the types of research-supported tools that have been shown to be essential to improved teaching and learning. Because Consortium efforts will not be successful without strong support from educators as well as a broad set of State stakeholders, we also are committed to developing a comprehensive communication network. Effective communication about the Common Core State Standards (CCSS) and the SBAC assessment system is essential to building a supportive environment for change. For this reason, capacity building and outreach efforts will extend from Consortium States to their districts, schools, and classrooms to promote explicit and meaningful connections among the CCSS, instruction, assessment, and effective data use. See Appendix A6-1 for an overview of SBAC’s proposed comprehensive plan for capacity building.

Essential Conditions for and Features of SBAC Professional Capacity Building

To ensure effective implementation of the proposed assessment system, a wide range of capacity-building activities are planned that are tailored to the needs of different stakeholder groups. These activities are designed to develop skills in collecting, interpreting, and using data among all end-users, including State policymakers, school and district leaders, teachers, parents, and students. We will support the development of research-based tools and resources that address needs at different levels (e.g., State, district, school, and classroom).

While capacity building is critical to successful implementation of the SBAC assessment

system, the Consortium also recognizes the limitations of what can be accomplished directly through its own resources. It is neither feasible nor appropriate for the Consortium to provide training for every administrator and teacher in each member State. Instead, our approach will be for the Professional Capacity and Outreach Working Group to collaborate with existing professional development networks in each participating State to ensure that capacity building penetrates the State, regional, and local levels. The networks will focus on building school and district capacity by helping them develop resources and trainings to support teachers, teacher leaders, and administrators at the local level. This approach serves three underlying goals: (1) it models a *sustainable structure* that supports teaching, learning, and reliable scoring of performance assessments and development of formative tools and processes; (2) it fosters *interdependence* and opportunities for shared learning among teachers and school leaders across States; and (3) it provides access to *authentic learning experiences* related to effective data use.

This approach to capacity building brings the added benefits of efficiency and cost effectiveness. Each participating State already has a network of district-level instructional coaches, teacher leaders, and/or trainers, etc., that can serve as the foundation for capacity building related to SBAC goals. In collaboration with representatives from institutions of higher education (IHEs) and the workplace, the Consortium will work with these existing networks to develop the materials and tools on which local professional development efforts can build.

Drawing upon guidance from ESEA on what constitutes high-quality professional learning, SBAC's capacity-building efforts will use "*a comprehensive, sustained and intensive approach to improving teachers' and principals' effectiveness in raising student achievement*" (NSDC, 2009). By taking advantage of existing State and local networks, many professional development activities can be site-based and all can be ongoing and embedded in the instructional context. Regularly scheduled technology-supported meetings (e.g., webinars or videoconferences), a key component of this plan, are intended to ensure that a consistent message is communicated to teachers and principals across districts and States. Also, in the name of consistency, SBAC will work with participating States to develop common frameworks for assessment, guidelines for effective formative assessment practices, and research-supported exemplars of curricular/instructional materials to support teachers' professional growth. Recommended readings, focused group discussions, use of online tools, and sharing of annotated examples of best practices and exercises are among the types of activities that will be available to help

educators develop expertise within the context of local practice.

Essential Steps in SBAC Professional Capacity Building

SBAC’s professional capacity-building efforts will begin immediately upon project funding. As the CCSS are the fundamental building blocks of the assessment system, our first priority will be to monitor ongoing State- and district-level efforts to “unpack” these standards at each grade level. As a first step, the Professional Capacity and Outreach Working Group will collect, review, and share exemplary resources from participating States. Web-based discussions will focus on the concepts, skills, and knowledge embedded in the CCSS, the learning pathways students must follow to reach valued college- and career-ready outcomes, and strategies for developing CCSS-aligned curriculum and instruction. SBAC will work with State teachers, technical advisors, and representatives from IHEs and the workplace to develop curriculum frameworks that address the full range of the CCSS and that clearly articulate what students must learn within and across grade levels. For the interim/benchmark assessments, the Professional Capacity and Outreach Working Group also will support the curriculum alignment to the learning progressions and the CCSS domains of ELA and mathematics (see Appendix A6-2).

Other capacity-building efforts with the primary aim of empowering instructional improvement will focus on collaboration in developing and disseminating resources focused on the following set of essential topics:

- ***Assessment literacy*** (i.e., the purpose of each component in a balanced assessment system [adaptive summative assessments that include performance events, optional interim/benchmark assessments, and formative tools²⁷]; assumptions associated with computer adaptive testing [CAT]; appropriate interpretation of results from each measure or tool and comparisons across schools, States, and consortia; and effective use of achievement and growth data)
- ***Alignment of curriculum, instruction, and standards*** (i.e., strategies for ensuring alignment among the CCSS, curriculum, instructional practices, and assessments)
- ***Formative assessment to guide instruction*** (i.e., the development of formative tools, processes, and practices; collecting and using student feedback to improve instruction; and making instructional adjustments or refinements based on the results of formative

²⁷ The interim/benchmark and formative components are described in the MOU as optional “formative/benchmark components,” which is a broader description for these system features.

data)

- **Constructed-response item and performance event development and scoring** (i.e., ways in which teachers can contribute to the development of constructed-response items and performance events for the summative and interim/benchmark assessments, and applying SBAC rubrics and protocols during scoring of items and events)
- **Individual-, group-, and organization-level learning** (i.e., supporting ongoing dialogue about strategies for improving student learning and facilitating professional development opportunities for pre-service and in-service teachers)

Professional Capacity Building in Development and Scoring of Summative and Interim/Benchmark Assessments and Formative Tools and Processes

A key element of SBAC’s professional learning approach for educators is to seek to engage teachers directly in developing and scoring SBAC assessments. Contributing to the item and performance event banks and participating in the moderated scoring process has been found to afford teachers a unique professional learning opportunity (Darling-Hammond & Pecheone, 2010). Planned efforts are summarized in Table A6-1 below.

Table A6-1: Strategies for Building Capacity

COMPONENT	STRATEGIES FOR BUILDING CAPACITY
Summative assessments	States will be encouraged to enlist teacher involvement in developing summative assessment items and performance events and in scoring constructed-response items and performance events. In all States, teachers will continue to review items and events prior to field testing.
Interim/benchmark (I/B) assessments	The use of curriculum-embedded I/B assessments that are based on learning progressions will provide teachers with models of high-quality tools that yield more finely grained information about what students know and can do. As with the summative measures, SBAC’s design calls for teachers to contribute to the I/B item and performance event banks and to review all items prior to operational use. Teachers also will be actively involved in scoring the constructed-response items and performance events.
Formative assessment tools and processes	Integration of formative tools, processes, and practices into Consortium classrooms poses the greatest professional development opportunity—and challenge—for teachers and instructional leaders. SBAC will provide models for local development of formative tools and processes, including professional learning modules that showcase best practices in data use for formative purposes in relation to the CCSS. Guidance will be developed that can be adapted to support differentiated instruction and a range of

	classroom practices. Web-based dialogues among teachers will provide opportunities to discuss approaches for collecting and using student feedback to improve instruction.
--	--

Use of Technology to Support Professional Capacity Building

Creative and adaptive use of technology to support professional development efforts will be a hallmark of the SBAC approach. Fundamental to the integrated assessment and learning system will be the system portal. The system portal will be the single entry point for educators, students, and parents to interact with the various aspects of the system, as appropriate, including the assessment delivery platform, the formative assessment resource clearinghouse, the distributed hand-scoring system, interactive dynamic reports, and an educator dashboard of student performance information. SBAC will take advantage of a variety of technologies (e.g., webcasts, issue-focused chat rooms, virtual classrooms, clearinghouses with live links to recordings and documents) and media (print, online) to disseminate resources and information, and to interact with users in promoting understanding of how these resources are related to improved instruction. Through SBAC’s educator dashboard—a secure online system—educators will be able to

- Download and view various assessment reports, scoring rubrics, and longitudinal data, and develop customized reports;
- Explore instructional strategies and recommendations for interventions related to specific assessment results for individual students or student subgroups;
- View and download vetted resources, including model instructional units, formative tools and rubrics, and sample performance events—all aligned to the CCSS;
- Access applications that facilitate and provide guidance on the design and scoring of items to ensure ongoing educator engagement in test use and validation; and
- Network with other teachers to share information, resources, and tools, and to engage in dialogue about curriculum, instruction, and assessment.

In collaboration with teachers and expert advisors, SBAC will develop and share guidelines to facilitate the selection and/or creation of useful resources and a process for soliciting and vetting resources before possible inclusion in the SBAC dashboard’s clearinghouse. Trained “peer reviewers” in participating States will review sets of materials or documents electronically

and apply specific criteria in judging the quality and usefulness of those resources for meeting existing or emerging Consortium needs.

In addition to an educator dashboard, SBAC will develop online tools and resources for students and parents. These may include enrichment activities that help extend current units of instruction, and guidelines for exploring in greater depth I/B performance data or results from formative tools that may be used diagnostically to better understand students' strengths and limitations. Engaging students and their families in using instructional resources is a key element in creating learner-centered school environments. Outreach efforts will support access to reports, student assessment information, and instructional resources by authorized users in a web-based environment.

Outreach and Communications Plan

Per our Theory of Action, a key priority of SBAC member States is to provide clear, accurate, and timely communication of standards, policies, and practices. This is essential for successful implementation of the new system within and across States. Effective communication is critical in the short term to signal change, and over the longer term to build continuing support for reform. SBAC is committed to transparency and clarity in communicating to all stakeholders (e.g., legislators, policymakers, IHEs, the workplace, community members, parents, educators, and students) the principles of the Consortium, the purposes of each assessment component in a balanced system, and the practices and intended outcomes of this assessment system.

The Consortium's Professional Capacity and Outreach Working Group will collaborate with the Steering Committee and technical advisors to develop and implement an outreach and communications plan for informing and updating the public and key stakeholders in each participating State about SBAC implementation. Among the issues covered in the plan are how to communicate descriptions of the assessment system, its purpose, and its components; sample and released test items and performance events; scoring rubrics for constructed-response items and performance events; achievement-level descriptors; and assessment results. It is important to note that the plan also will need to address effective dissemination of a variety of reports.

The Professional Capacity and Outreach Working Group will work with and through existing State communication mechanisms to ensure that essential information gets to stakeholders and the general public in a timely fashion. This collaboration will ensure that each participating State

has a communication plan and a set of communication tools to use with local and State users of data from SBAC assessments. These tools will include

- A timeline for public release of informational materials;
- An overview of communication processes and products (e.g., arrangements for media coverage, talking points, and illustrations needed for all spokespersons within the Consortium);
- Protocol for national press conferences, including State and Consortium contacts for multi-state media releases;
- Materials for public consumption, such as templates for PowerPoint presentations, brochures, pamphlets, information letters, and newsletters; and
- Access to an SBAC-developed and -managed website that provides communication materials, technical reports, and important notices about Consortium projects/initiatives.

Finally, the Professional Capacity and Outreach Working Group will work with existing regional and Federal networks, such as the Comprehensive Centers and Regional Educational Laboratories, as well as the private sector (e.g., service providers and test publishers), so that they are also informed about SBAC's work and, as appropriate, can be engaged in our outreach efforts.

(A)(7) Technology Approach (up to 10 points)

The extent to which the eligible applicant is using technology effectively to improve the quality, accessibility, cost-effectiveness, and efficiency of the proposed assessment system. In determining the extent to which the eligible applicant is using technology effectively, we will consider—

- (a) The description of, and rationale for—
 - (i) The ways in which technology will be used in assessment design, development, administration, scoring, and reporting;
 - (ii) The types of technology to be used (including whether the technology is existing and commercially-available or is being newly developed); and
 - (iii) How other States or organizations can re-use in a cost effective manner any technology platforms and technology components developed under this grant; and
- (b) How technology-related implementation or deployment barriers will be addressed (*e.g.*, issues relating to local access to internet-based assessments).

SBAC’s Theory of Action highlights technology as the major vehicle for supporting the integration of the learning and assessment systems. ***Technology will be utilized to support all aspects of the system***, including developing, administering, and scoring of the adaptive summative assessments, performance events, and adaptive interim/benchmark assessments; developing and using formative tools, processes, and practices; and accessing the ***SBAC system portal*** that will house the public website, an educator dashboard, a clearinghouse of instructional resources, and item writing and scoring modules. Per the NIA, all assessment items will be developed to an industry-recognized open-licensed interoperability standard approved by USED during the grant period, without non-standard extensions or additions. The interoperable design will support a consistent system of assessment delivery across Consortium States, provide item portability as needed, ensure continuity of universal item design features across multiple platforms and uses, and provide competition among vendors for future work. An ***open-source technology solution*** will be used as a means to promote continued collaboration, innovation, and low cost of ownership. Recognizing the current limitations in available open-source software, the assessment system software will be developed using a combination of existing and newly developed open-source software and proprietary software.

System portal. Fundamental to the integrated assessment and learning system will be the system portal. The system portal will be the single entry point for educators, students, and

parents to interact with the various aspects of the system, as appropriate, including the assessment delivery platform, formative practices and tools resource clearinghouse, distributed hand-scoring system, interactive dynamic reports, and an educator dashboard of student performance information. The portal will be designed to promote continued expansion and innovation by allowing the integration of external resources. User-defined roles and views will allow the portal to serve as a central location for stakeholders such as administrators, educators, parents/guardians, and students while ensuring that security and confidentiality are properly maintained. For example, students will have access to their individualized results and relevant classroom and project materials, as well as the ability to communicate and share information with teachers and other students, and the ability to submit work for local or distributed scoring. This will enable students and parents to monitor progress toward college- and career-readiness and allow students to collaborate with peers and educators for continuous improvement.

For teachers, the portal will strengthen connections among standards, curriculum, assessment, instruction, student learning, and teacher development. The portal will be the hub for both the educator resource clearinghouse and educator dashboard, which will be fully integrated to allow educators to move seamlessly from student performance profile information found in the dashboard to linked resource materials that support targeted instructional change. These technology-supported connections are essential to making student achievement evidence actionable for teachers and fostering a culture of continuous improvement for students, parents, and teachers. Training for teachers to use the online formative tools, processes, and practices will also be technology-supported via a variety of methods, including videoconferences, webinars, and online training modules that can be accessed by educators at any time.

Technology to support assessment development and implementation. SBAC will use best practices in application design, development, and implementation to create an innovative and flexible online system with a robust set of features that allow States the stability of a Consortium-developed system and the flexibility to implement State-specific approved variations. To that end, the SBAC's Technical Advisory Committee (TAC) will leverage the lessons learned from Consortium members and ensure that best practices of online assessment design and architecture are reflected early in the requirements-gathering and technology-development processes. The purview of the technical advisors will include issues related to user-interface design of the system, software development standards, system development and

enhancement roadmap, database architecture, system testing, security, and all other technological issues necessary for continued growth in system efficiency and effectiveness in improving teaching and learning.

SBAC will design the system to be browser- and operating-system-independent, thereby minimizing user issues and increasing access to the testing platform by all schools. User-interface designs and information architecture will undergo exhaustive use-case scenario testing by an independent third party and will be enhanced to maximize system usability.

Item development will be supported through an online authoring environment, which provides teacher training; allows item submission, feedback, and review; and supports the item approval process. The item development process and design requirements will be documented prior to generating system architecture or code specifications. Additionally, the item development system will warehouse existing items provided by Consortium States and support evaluation of alignment to the CCSS and suitability for the summative or interim/benchmark assessments.

Item designs will incorporate specific “feature tagging” systems that will be used to render items in a consistent manner that meets student access and communication needs.

Documentation will allow full interoperability across different delivery platforms in the various Consortium States (e.g., Accessible Portable Item Protocol [APIP]). The item authoring system will be structured to facilitate the development of items that follow the principles of universal design for assessment.

SBAC will use a computer adaptive testing (CAT) engine to deliver summative and interim/benchmark assessments and to score (and/or distribute for scoring) student responses. The CAT engine will be fully integrated with the Consortium item bank to ensure that CAT forms meet content and other psychometric constraints. Assessment item banks will be partitioned into a secure item bank for the summative assessment and a non-secure bank for the interim/benchmark assessments. The assessment platform will support the online delivery of a variety of item types, including traditional selected-response items, constructed-response items, curriculum-embedded performance events, and technology-enhanced items (modeled after assessments currently used by the U.S. military, the architecture licensure exam, and NAEP). The use of technology-enhanced item types will target content standards that, in the past, have

been difficult to measure, such as deep knowledge and/or application of knowledge and skills.

The assessment system will support multiple opportunities for the summative assessment administration window for grades 3 through 8 and high school. It will store students' results for multiple assessment administrations to support defined reporting rules. Banking of student performance event responses and scores will be enabled by a database designed to maximize data integration across the various components of the assessment system (e.g., scoring system and reporting engine) to allow performance event responses to be combined with CAT responses to yield the overall summative assessment score.

The assessment system scoring engine will manage all student response scoring: standard machine scoring (selected-response items), automated scoring (constructed-response items and performance events), distributed teacher-moderated scoring (constructed-response items and performance events), and local scoring (specific I/B assessment performance events). As advancements are made with automated scoring solutions during the life of this project, they will be integrated into the system. Where technology solutions do not yet exist to score student responses collected via this system, the system will facilitate the distributed human scoring of those responses by appropriate individuals or groups of individuals (e.g., teachers score student responses on classroom assessments; trained/certified individuals score student responses on large-scale assessments). The distributed human scoring system will manage scorer qualifying activities, response scoring assignments, and scorer reliability and validity auditing.

Summative and I/B assessment results will be accessible via the assessment system portal and directly integrated with the student performance profile. Student assessment results will flow directly from the scoring and psychometric systems to the reporting system that manages the full suite of reports; thus, reporting will be dramatically faster and less prone to manual integration errors. The latest business-intelligence-based reporting will be used to serve the full continuum of reporting needs, from the more traditional static reports of annual achievement in English language arts (ELA) and mathematics to dynamic reporting that allows data-mining and vivid representations of a student's achievement along a learning progression framework for I/B assessment results. The identity management system will enable a host of stakeholders to access customizable reports; however, the specific suite of reporting tools available will be user-credential-dependent. The reporting suite interface and connections to formative materials will

be designed and piloted with teacher input using rapid iterative design and usability testing processes.

Efficiencies, accessibility, and security of technology solutions. Technology solutions will be selected to maximize each State’s ability to deliver services and tools in a cost-effective manner Statewide, regardless of school size, location, or funding status. Equity of student access across small, rural, urban, and other areas will be a high priority. The system will require minimal local IT staff involvement to install, operate, or update any software applications, test engine, or related test administration components. For example, the test engine and its administration ancillaries will have self-diagnostic tools that can determine if the workstation or network configuration is prepared correctly for operation. Additionally, the system will be designed to allow States to integrate their information management system to support administration and management of student performance data.

The SBAC TAC will approve and update security protocols that use state-of-the-art security methods for encryption of test and personal data (e.g., 128-bit SSL encryption). The strict security protocols of encrypted databases will protect the security and confidentiality of student information in adherence to current Family Educational Rights and Privacy Act (FERPA) regulations. Disaster recovery and backup strategies will also be designed into the system to mitigate the risk of data loss. An industry-standard identity management system will define the information available and the rights a user has when logged into the system. Additionally, test session security (secure browser, session control, administrator control, password-required login) will be designed to maximize test security while minimizing local burden.

Cost-effective reuse will be facilitated through a system design process that prioritizes the use of open-source and interoperable standards, alternative scoring models, performance event sharing, hosting options, infrastructure guidelines, and administrative efficiency. A critical component of the ongoing system development process will include code sharing and feedback from an expanding, unrestricted user-base. This process will ensure that the system remains current and continually grows to meet the challenge of improving teaching and learning. Upon completion of system development, the Consortium’s legal counsel will create a public license (e.g., gnu.org, Mozilla.org/mpl) defining this product as a free open-source software application.

Identifying and mitigating against barriers. Design of the assessment system will leverage the collective wisdom of the Consortium and the member States’ experience addressing

challenging implementation and deployment issues. Therefore, as part of the initial assessment system development process, the Consortium will formalize the collection of lessons learned across Consortium States through a Technology Implementation Task Force. The charge of this task force will be to work with States currently administering online assessments to identify risks and to develop mitigation and/or avoidance strategies to address the risks. These risks, challenges, and solutions will be documented in a formal report to provide guidelines for technology implementation best practices to ensure that lessons learned are reflected in areas like system architecture and the graphic user interface.

Development of the assessment and learning system will be considered a dynamic process of evaluating how technological solutions that are in place meet current needs of the assessment system design (i.e., improving teaching and learning). The TAC will conduct periodic needs assessments and make recommendations for enhancements to the development roadmap. Evaluating the suitability of existing tools and systems currently in development (e.g., the CCSSO data platform) or in use by States (e.g., the Item Bank system in development by the State of Michigan) will be integral to ensuring that Federal investment dollars are maximized.

(A)(8) Project Management (up to 30 points)

The extent to which the eligible applicant’s project management plan will result in implementation of the proposed assessment system on time, within budget, and in a manner that is financially sustainable over time. In determining the extent to which the project management plan has these attributes, we will consider—

- (a) The quality, qualifications, and role of the project management partner, as evidenced by its mission, date of founding, size, experience (including past success in implementing similar projects), and key personnel assigned to this project (including their names, curricula vitae, roles, percent of time dedicated to this project, and experience in managing similar projects);
- (b) The project workplan and timeline, including, for each key deliverable (e.g., assessment component, scoring and moderation system, professional development activities), the major milestones, deadlines, and entities responsible for execution; and the approach to identifying, managing, and mitigating risks associated with the project;
- (c) The extent to which the eligible applicant’s budget—
 - (i) Clearly identifies Level 1 budget modules (as defined in the NIA) and any Level 2 budget modules (as defined in the NIA);
 - (ii) Is adequate to support the development of an assessment system that meets the requirements of the absolute priority; and
 - (iii) Includes costs that are reasonable in relation to the objectives, design, and significance of the proposed project and the number of students to be served; and
- (d) For each member State, the estimated costs for the ongoing administration, maintenance, and enhancement of operational assessments in the proposed assessment system and a plan for how the State will fund the assessment system over time (including by allocating to the assessment system funds for existing State or local assessments that will be replaced by assessments in the system).

Required Tables and/or Attachments:

- ✓ *Summary Table for (A)(8)(a): Key Project Management Personnel*
- ✓ *Attachments for (A)(8)(a): Curricula Vita of Key Personnel*
- ✓ *Summary Table for (A)(8)(b): Project Work Plan and Time Line*
- ✓ *Attachments for (A)(8)(c): Budget (See Part I.J)*

(A)(8)(a) The SMARTER Balanced Assessment Consortium (SBAC) is committed to an open procurement process for identifying a Project Management Partner who best meets the following criteria: (1) capacity to effectively and efficiently complete the scope of work; (2) proven track record of success in projects with a similar scope; and (3) economy (i.e., the ability to produce the highest-quality deliverables for the fewest resources possible). SBAC used these criteria in procuring the services of an external agent to support the management and development of this proposal. The Request for Proposals (RFP) that was used to recruit that agent is attached as

Appendix A8-1. The process and selection criteria proved so effective in this first procurement that we are committed to following a similar open, competitive procedure for identifying the Project Management Partner to support SBAC in developing its proposed assessment system. Appendix A8-2 is the RFP let by the State of Washington to procure the services of a successor Project Management Partner, and demonstrates the Consortium's aim to meet the intent of, as well as the requirements described in, the NIA, while abiding by the State of Washington's procurement laws, and continuing to adhere to the values of openness and transparency that are the keystones of this Consortium.

Approach to Project Management

Project management comprises a careful balance of time, resources, and quality. As is clear in our Theory of Action, quality is not optional. The system must meet the highest standards of best practice and meet the needs of our stakeholders so that we can truly enact change. The RFP for the Project Management Partner (Appendix A8-2) serves as an example of the Consortium's approach to ensuring quality by incorporating balance into each of its contracts and Service Level Agreements. As for resources, we will ensure an ambitious development approach while also staying within our budget. Our plan is based on the assumption that the work will be supported solely by the award from USED, even though we will seek additional support from private foundations. The Consortium will control costs through a system of checks and balances as well as by including explicit "not to exceed" parameters in each of its contracts. Time, then, truly is the only variable in our project management approach that will have any flexibility, as long as the assessment is operational by 2014–15.

As a safeguard to ensure the timely completion of deliverables, the Consortium will use a multi-tiered risk management strategy for identifying barriers to completion. First, the working group structure will distribute the work into manageable components so that State assessment experts as well as technical advisors will incrementally monitor progress and escalate issues to the Project Management Partner and the Executive Committee when appropriate. Similarly, as described in the governance structure, the Executive Committee will include technical and policy assessment experts who will prioritize the urgency of problems or concerns, based on their own experience and the advice of the Project Management Partner. The Project Management Partner will disseminate weekly project status updates, requiring systematic tracking that will determine

whether any deliverable is at risk. Consortium members will be able to review these updates and notify the Project Management Partner or members of the Executive Committee if they disagree with the status characterization. Finally, as the Lead Procurement State, Washington, will periodically review the progress of the entire project from a compliance perspective, and as a member of the Executive Committee, will be able to directly identify any issues that would appear to interfere with the timely completion of deliverables.

The primary risk management strategy will be to create comprehensive work plans as soon as possible to ensure that sufficient time and resources are allocated to complete the work. A secondary strategy will be frontloading tasks in the project plan as far ahead as possible to ensure that there is sufficient time to move dates back as necessary to meet the operational administration requirement in 2014–15. The timelines will be reviewed by the quality control vendor to ensure that the full body of work is appropriately described.

Acquiring Project Management Services

Per Section 2.2 of Appendix A8-2 (*Estimated Schedule of Procurement Activities*), the State of Washington has set the timeline for identifying and placing under contract the permanent Project Management Partner by October 1, 2010. The RFP details a comprehensive solicitation that focuses on adherence to protocol, fidelity to the solicitation specifications, and a thorough review process enabling the selection of a “best fit” candidate. Adherence to protocol will ensure fairness for prospective vendors, while fidelity to the specifications will ensure quality proposals are submitted that meet the technical requirements of the Consortium. With a thorough review process detailing the criteria proposals will be evaluated by Consortium members, the Department can have confidence that the best qualified candidate will be selected to assist in the implementation of a grant award.

Since the NIA stipulates that the Project Management Partner is part of the grant implementation plan, the Consortium has inferred that the associated efforts of the Project Management Partner are part of the grant award itself, which would not be available for actual work until October 1, 2010. Given this situation, the Consortium would experience project start-up delays as the Project Management Partner familiarizes itself with the activities and objectives of SBAC. Such concerns resulted in the Consortium Steering Committee’s recognition of a need for a transition plan and transition partner. Thus, when developing the RFP for the Grant Project

Manager (see Appendix A8-1), the need for a short-term extension of services beyond the immediate deliverable (proposal submission) was introduced as a contractual clause, allowing SBAC the right to extend the scope of work through subsequent amendment as a means of supporting SBAC through a project transition phase.

The Consortium has implemented this option by extending its contract with WestEd in a role as interim Project Management Partner for a period not to exceed three months from the date of notice of an award. The purpose of this interim Project Management Partner contract is to explicitly meet the requirements of the NIA and to allow a smooth transition to immediate first steps as a funded Consortium while proceeding with the comprehensive procurement process required by the State of Washington. In the event that WestEd is not selected as the Project Management Partner through this competitive process, all parties will remain committed to ensuring a complete and seamless handoff of responsibilities to the organization assuming this role on a permanent basis.

Interim Project Management Partner

WestEd is uniquely qualified to undertake the role of interim Project Management Partner for SBAC. The agency's experience and expertise in the development and management of comprehensive assessment systems has prepared it for this important work. The knowledge and skills WestEd brings to the Consortium include

- **Expertise in large-scale project management and progress tracking.** The proposed Project Management Partner team is experienced in planning and managing projects with multiple partners and activities. Team members understand how to communicate effectively and efficiently, collect and synthesize information from diverse sources, and track project progress—even when time lines are short and stakes are high.
- **Experience working with States on standards and assessment issues.** WestEd has worked directly with multiple SEAs and most of the assessment contractors on large-scale assessment issues. The assembled team knows what types of information SEAs need to support informed decision-making (e.g., how proposed changes may affect struggling learners, students with disabilities [SWDs], and English learners [ELs]) and team members are well versed in bringing diverse groups of constituencies together to achieve common goals. In Appendix A8-3, there is an overview of the states in which

WestEd staff have provided standards and assessment related services. Most recently, staff have been actively involved in supporting states as they conduct crosswalks between the CCSS and applicable state standards, and are supporting states during the transition to instruction based on the CCSS.

- **Experience working with State consortia.** In addition to their current role as the manager of the proposal process for SBAC, Dr. Stanley Rabinowitz, the proposed interim Project Management Partner senior advisor, has served for the past several years as a member of the New England Common Assessment Program TAC, learning about the characteristics of a successful state assessment consortium. Additionally, the Assessment and Accountability Comprehensive Center directed by Dr. Rabinowitz supports a group that includes 16 regional comprehensive centers, USED, and many States working together to meet important assessment goals.
- **Experience with all aspects of “next generation” assessments.** The proposed WestEd team understands and endorses the balanced assessment model. This team has deep, current assessment knowledge and leadership experience in guiding states on the development, implementation, and evaluation of comprehensive, state-of-the-art assessment systems. WestEd also advises other entities that seek its experience and expertise in the research, policy, technical, and logistical aspects of building and sustaining a reliable, valid, and feasible, yet forward-looking assessment system. For example, WestEd is the lead advisor to the Bill and Melinda Gates Foundation College Readiness Assessment Pilot.

More broadly, WestEd brings institutional capacity to serve in the role of Project Management Partner. WestEd is a preeminent educational research, development, and service organization with over 500 employees and 16 offices nationwide. WestEd has been a leader in moving research into practice by conducting research and development (R&D) programs, projects, and evaluations; by providing training and technical assistance; and by working with policymakers and practitioners at state and local levels to carry out large-scale school improvement and innovative change efforts.

In carrying out this work, WestEd has built solid working relationships with education and community organizations at all levels, playing key roles in facilitating the efforts of others and in initiating important new improvement ventures. Over the past 42 years, WestEd and its two

predecessors, Far West Laboratory for Educational Research and Development (FWL) and Southwest Regional Laboratory (SWRL), have carried out nearly 2,000 successful projects representing major contributions to the nation’s R&D resources. At any given time, WestEd has from 450 to 700 active contracts and grants. Current work extends beyond the Western Region to include most states in the nation and an increasing number of other countries. More information regarding WestEd’s breadth of experience and capacity can be found in Appendix A8-4.

Key Staff for Proposed Interim Project Management Partner

Key staff for this project are listed in the Summary Table for (A)(8)(a) below. Selected staff resumes are provided in Appendix A8-5. All named staff will be available to begin work immediately upon grant award announcement.

Summary Table for (A)(8)(a): Key Project Management Personnel

Key personnel of Interim Project Management Partner	Role Assigned	Percent of Time Dedicated to Project	Vita Attached
Dr. Stanley Rabinowitz	Senior Advisor	0.25	✓
Dr. Joanne Jensen	Project Director	0.50	✓
Dr. Carole Gallagher	Research Director	0.50	✓
Dr. Carol Whang	Project Coordinator	0.50	✓
Dr. Christyan Mitchell	Director of Technology Implementation	0.30	✓
Dr. Elizabeth Berkes	Project Support	0.50	
Patricia Armstrong	Project Support— Mathematics	0.50	
Beverly Nedrow	Project Support—English language arts	0.50	

(A)(8)(b) A System of Checks and Balances

The critical first step is to develop a work plan and schedule that includes the high-level requirements for meeting major goals. This process will define the milestones and absolute requirements from which a more detailed development schedule for development will emerge. Effective project management will require development of detailed schedules for each of the assessment system components, which will be combined into a fully integrated master schedule. The process of combining each assessment component into the master schedule increases the awareness of touch-points, hand-offs, and dependencies that may otherwise go unnoticed. Understanding how each project component will interact and relate affords project management staff the ability to better anticipate potential risks and plan for contingencies.

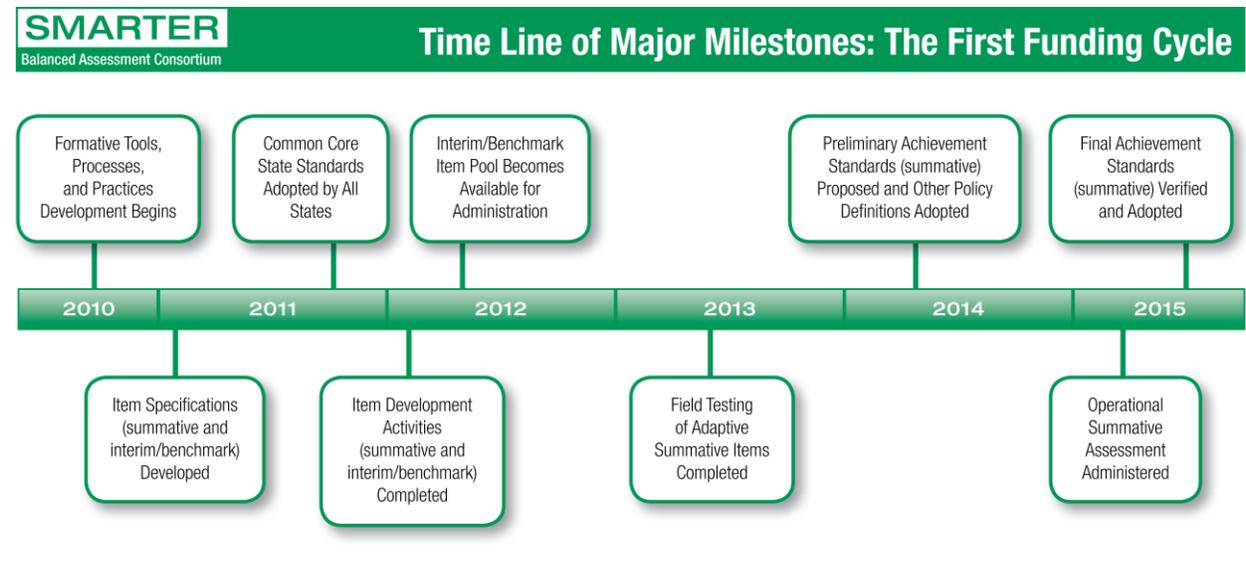
After the master schedule is developed, daily project management reports will be shared with the appropriate SBAC staff/leadership, including daily stoplight-status reports. The stoplight-status reports provide a high-level status indicator for each assessment component—indicating, for each assessment component, whether it is considered “green” (on-schedule, with no anticipated risks), “yellow” (on-schedule, with medium risk of moving off-schedule), or “red” (off-schedule or on-schedule with high risk of moving off-schedule). Any schedule variances (i.e., yellow or red) will be reported with strategies for course correction, the estimated likelihood that corrective action will be effective, and possible mitigation strategies if course correction fails.

As part of the project schedule development process, the Project Management Partner will work with participating States to identify implementation barriers, risks, and possible solutions or mitigation strategies. The key to the success of a project of this complexity will be contingency planning from the outset (e.g., the budget indicated in part (c) of the Project Management section includes a paper-form contingency plan in the event of adaptive computer system failure).

Consistent with the values driving use of ARRA funds, the Consortium will ensure that expenditures are tightly monitored through a series of checks and balances. As described in the budget proposal, Washington State will allocate personnel to serve as project director, project manager, and fiscal analyst. Their role will be to monitor the progress of the grant and the expenditures from a compliance perspective to ensure that the grant deliverables and

expenditures are consistent with the agreement signed by USED. The Project Management Partner, as described above, will plan, monitor, and report on the work as necessary to achieve the deliverables as described in the agreement.

Finally, the Consortium will hire an external third party to conduct quality control of project plans, applications, and documentation as necessary to maintain the confidence of USED, the Consortium’s stakeholders, and U.S. taxpayers. For purposes of this submission, the time line presented in the Summary Table for (A)(8)(b) reflects an aggressive implementation plan with this intent in mind. The Project Management Partner will provide a revised and fully fleshed out comprehensive project plan based on input from additional technical experts who are able to assist in the estimation of the duration and contingent nature of tasks. The major project milestone activities and deliverables are shown in the Summary Table for (A)(8)(b): Project Work Plan and Time Line. As an overview, these milestones are shown in the time line below.



Summary Table for (A)(8)(b): Project Work Plan and Time Line

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
Adopt the Common Core State Standards, which are college- and career-ready standards, and to which the Consortium's assessment system will be aligned		10/1/2010	12/31/2011	Each State
Release RFP for item writing specifications, protocols, and training materials	Contract for consultant to write RFP	10/1/2010	11/1/2010	Washington (Lead Procurement State)
	RFP released	11/1/2010	12/1/2010	
Select contractor(s) for test specifications, item writing specifications, protocols, and training materials		10/1/2010	4/1/2011	Executive Committee
Contractor(s) produces test specifications, item writing specifications, protocols, and training materials	Draft item writing protocols and training materials	4/1/2011	7/31/2011	Contractor
	Pilot item writing protocols and training materials	4/1/2011	6/30/2011	Contractor with Governing States
	Establish item writing goals based on test specifications	4/1/2011	8/31/2011	Contractor
	Train State/State delegates for item writing	8/31/2011	12/1/2011	Contractor with Governing States

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
Item Development	Evaluate Consortium-submitted items and conduct gap analysis	12/6/2010	4/1/2011	Contractor with Governing States
	Distributed item writing	12/15/2011	3/1/2012	Governing States
	Conduct Content Panel Validation Review I	3/15/2012	3/20/2012	Contractor
	Targeted small-scale pilot assessment	4/1/2012	4/07/2012	Contractor with Governing States
	Revise item writing protocols and training	4/20/2012	5/1/2012	Contractor
	Distributed item writing	6/30/2012	9/3/2012	Governing States
	Conduct Content Panel Validation Review II	12/15/2012	12/20/2012	Contractor with Total State Membership
	Item bias review	12/15/2012	12/20/2012	Contractor with Total State Membership
Identify vendor(s) for Test Development, Delivery, Scoring, and Reporting Application	Contract for consultant to write RFP for test delivery application	10/1/2010	11/1/2010	Executive Committee
	RFP released	11/1/2010	12/1/2010	Washington (Lead Procurement State)
	Select vendor(s) for test development, delivery, scoring, and reporting application	12/1/2010	4/1/2011	Executive Committee

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
Vendor(s) produces Test Development, Delivery, Scoring, and Reporting Application	Create application specifications	4/1/2011	8/1/2011	Vendor
	Unit test and system test application	4/1/2011	8/1/2012	Vendor
	Pilot item authoring system	4/1/2011	8/1/2011	Vendor and Governing States
	Pilot delivery, scoring, and reporting application	4/1/2011	12/1/2012	Vendor and Governing States
	Revise application in preparation for field test	4/1/2011	3/1/2013	Vendor
Field test items using State existing online applications and/or stand-alone implementation of new online system as operational pilot	Assessment administration training	2/1/2013	2/22/2013	Vendor and Governing States
	Embedded and stand-alone field test	3/1/2013	6/1/2013	Vendor and Governing States
	Field test analysis file complete	7/1/2013	7/1/2013	Contractor
Field Test Item Performance Review (Data Review)	Data review meetings	8/15/2013	8/20/2013	Contractor
Adoption of Achievement Standards	Representative stakeholders participate in achievement standard setting	8/31/2014	8/31/2014	Stakeholders from Total State Membership
	Executive Committee reviews the standard-setting process and achievement standards and approves the standards	8/31/2014	11/1/2014	Executive Committee

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
	Executive Committee recommends achievement standards to Steering Committee based on standard setting for their approval	11/1/2014	11/30/2014	Executive Committee
	Standard-setting verification using operational results and makes final recommendations for approval	7/1/2015	7/20/2015	Vendor and Executive Committee
	Governing States vote to approve achievement standards	7/20/2015	8/1/2015	Governing States
	Total State Membership adopts achievement standards	8/1/2015	9/1/2015	Total State Membership
Test Administration	Operational assessment is available	3/1/2015	6/1/2015	Vendor and Total State Membership
	Fully implement Statewide the Consortium summative assessment in grades 3–8 and high school for both English language arts and mathematics	10/1/2014	6/1/2015	Total State Membership
	Operational results reported	TBD	TBD	Vendor
Financial Planning	Governance/Finance Working Group identify ranges of options regarding Consortium financing options	10/1/2010	10/1/2011	Governance/ Finance Working Group with consultants

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
	Governance/Finance Working Group revises operational cost estimates using format established through finance options	10/1/2011	6/1/2012	Governance/ Finance Workgroup with consultants
	Steering Committee selects finance option(s) for Governing vote based on Executive Committee review	6/1/2012	1/30/2013	Steering Committee
	Governing States vote on financial plan	1/30/2013	6/1/2013	Governing States
Identify vendor(s) for Curriculum and Formative Materials	Contract for consultant to write RFP for curriculum and formative material development	10/1/2010	11/1/2010	Executive Committee
	Release RFP	11/1/2010	12/1/2010	Washington (Lead Procurement State)
	Select vendor(s)	12/1/2010	4/1/2011	Executive Committee
Summarize definitions of ELs, State policies, practices, and research on accommodations to inform educators, researchers, and other stakeholders across the States in the SMARTER Balanced Assessment Consortium	Create a report summarizing definitions, criteria and guidelines for identifying and including ELs in content assessments. The report will provide a comprehensive list of accommodations, the frequency of use of each accommodation, and analyses of potential responsiveness of these accommodations for ELs. The report also will identify differences in policies, practices,	10/1/2010	6/30/2011	Assessment Design Working Group

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
	accommodations, and rationales for the use of specific accommodations across the Consortium States			
Produce a definition, inclusion policies, and a set of accommodations appropriate for EL students that are commonly accepted by educators, researchers, and other stakeholders across the States in the SMARTER Balanced Assessment Consortium	Conduct 3-4 update and discussion Webinars that include Consortium SEA teams representing assessment, Title III, special education, and curriculum; (2) hold a final meeting with a representative from each State to attain consensus on allowable accommodations for the Consortium	7/1/2011	10/31/2011	Assessment Design Working Group
	Summarize and disseminate the consensus process and agreements reached with Consortium States. Dissemination strategies will include on-line materials, conference presentations, and other appropriate means that might be identified	11/1/2011	12/31/2011	Assessment Design Working Group
Summarize definitions of SWDs and 504 students, State policies, practices, and research on accommodations to inform educators, researchers, and other stakeholders across Consortium States	Create a report summarizing the definition, criteria, and guidelines for identifying and including SWD and 504 students in regular content assessments. The report will provide a comprehensive list of accommodations, the frequency of use of each accommodation and analyses of potential responsiveness of each of the accommodations for SWD and 504 students. The report also will identify	10/1/2010	6/30/2011	Assessment Design Working Group

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
	differences in policies, practices, accommodations, and rationales for the use of specific accommodation across the Consortium States			
Produce a definition, inclusion policies, practices, and a set of accommodations appropriate for SWD and 504 students that are commonly accepted by educators, researchers, and other stakeholders across the Consortium States	Conduct 3-4 update and discussion Webinars including Consortium SEA teams representing assessment, Title III, special education and curriculum; (2) hold a meeting with a representative from each State to attain consensus on allowable accommodations for the Consortium	7/1/2011	10/31/2011	Assessment Design Working Group
	Summarize and disseminate the consensus process and agreements reached with Consortium States. Dissemination strategies will include online materials, conference presentations, and other appropriate means that might be identified	11/1/2011	12/31/2011	Assessment Design Working Group
Identify the characteristics of the accommodation needs of EL, SWD, and 504 students in a technology-based assessment environment that is commonly accepted by educators, researchers, and other stakeholders across the States in the SMARTER	Conduct an in-depth study of States currently using a high level of computer-delivered tests and the types of accommodations available to students (highlighting, increased typeface, avatar voices, etc.). Create a report summarizing the findings of the in-depth study of the current use of technology in assessment by States, including the specific	11/1/2010	3/1/2011	Assessment Design Working Group

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
Balanced Consortium	types of technology used			
	Conduct an in-depth study examining the technology available from vendors. Summarize the accommodations possible through assessment technologies currently available	2/1/2011	6/30/2011	Assessment Design Working Group
	Determine the barriers to the use of technology for access to assessments and suggest potential solutions. Create a report explicating the barriers to the use of technology for assessment and the recommended solutions	6/1/2011	9/31/2011	Assessment Design Working Group
	Conduct a case study of the Oregon assessment system and the use of technology within the system for EL, SWD, and 504 students. Synthesize and report findings describing the use of technology in the Oregon assessment system so that it might be used to inform the SMARTER Balanced Assessment Consortium decisions	9/1/2011	3/1/2012	Assessment Design Working Group
To monitor fidelity in application of inclusion and accommodations guidelines and/or policies by States, an inclusion and accommodations manual	Create and disseminate an accommodations manual. The EL section will address definition, inclusion and accommodation of EL in content area assessments and will	12/1/2011	9/30/2012	Assessment Design Working Group

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
will be provided that can be used to ensure consistent and streamlined inclusion and accommodation policies across the Consortium of States	include a common definition of an EL, recommended inclusion and accommodation guidelines, and valid and effective accommodations for EL. The SWD/504 section of the manual will address definition, inclusion and accommodation of SWD and 504 students in regular content area assessments and will include recommended inclusion and accommodation guidelines, and valid and effective accommodations for SWD and 504 students			
Systematically review accommodations for the SMARTER Balanced Assessment Consortium	Recommend a detailed process (in writing) to be used by the SMARTER Balanced Assessment Consortium in annual review of inclusion and accommodations guidelines and/or policies	7/1/2011	10/1/2011	Assessment Design Working Group
	Create an online standard template for use in evaluating accommodation use for peer review	10/1/2011	12/31/2011	Assessment Design Working Group
	Establish a research agenda the Consortium can use to study, monitor, evaluate, and improve the validity of accommodations	10/1/2011	9/30/2012	Assessment Design Working Group
Technical Advisory Meetings	Quarterly Meetings	TBD	TBD	Executive Committee

Major Milestones	Associated Tasks	Start Date	End Date	Responsible Entity
SBAC Planning and Management Meetings	Two meetings per year	TBD	TBD	Executive Committee

(A)(8)(c) The Consortium intends to create a comprehensive and balanced assessment system that includes summative, interim/benchmark, and formative assessments. Online computer adaptive summative assessments in ELA and mathematics administered in grades 3–8 and at least once in high school will be created. The design of the summative assessment is described in Section (A)(3). In addition, the Consortium will provide online adaptive interim/benchmark assessments for grades 3–8 and high school that will provide for more in-depth assessment of what students know and can do in ELA and mathematics based on smaller clusters of content aligned to the Common Core State Standards and to describe incrementally the degree to which students are on track to be college- and career-ready. The design of the interim/benchmark assessment is also described in full in Section (A)(3). Finally, as part of the Level 1 budget request, the Consortium will provide formative assessment tools and resources and professional development support for teachers in elementary, middle, and high school as part of the integrated system of teaching and learning described in the Theory of Action (Section (A)(2) and in the assessment design (Section (A)(3)).

The Consortium’s main deliverables include a comprehensively designed assessment system, a consolidated reporting system, online test administration with a secure item and performance event bank, and an extensive system of professional development tools and resources.

The table below indicates the Level 1 budget estimates by working group.

	Project Management & Governance	Assessment Design	System Design	Research & Evaluation	Professional Capacity & Outreach	Technology	Higher Ed Engagement	Total
Total Funds Requested	\$10,435,922	\$97,950,884	\$428,693	\$5,008,550	\$7,550,650	\$27,074,143	\$1,538,977	\$149,987,819

As depicted in the table above, almost 85% of the requested funds go to support two key elements of the SBAC assessment system—the development of the summative and interim/benchmark assessments (\$98 million) (described in Section (A)(3) and in the detailed budget narrative section of the document in Appendix A8-6) and the technology required to deliver the computerized adaptive assessments and other systems integral to the effective implementation of the Consortium’s vision (\$27 million). The development of the assessment

and technology systems will be facilitated by vendors working closely with States and the Consortium in a collaborative process under contractual arrangements with the Consortium.

The Consortium plans on developing two key technology systems as part of its assessment system—a computerized assessment delivery platform to deliver the actual computer adapted assessments (\$15 million) and a centralized data repository, reporting, and professional development delivery system (\$7.5 million) where all student responses and professional development materials will be housed and all test results and other information will be generated and reported. Additionally, the Consortium plans on enhancing and using a Michigan-developed web-based item authoring and item banking system where teachers and other contributors can develop and archive new items for the assessment system (\$2 million) and have numerous item characteristics and metadata stored for use in populating item pools for the test delivery system. This process, using an existing State system and enhancing it for use by the Consortium, is a significant cost-savings approach.

Key to a successful implementation of the SBAC vision is the Consortium project management and governance (\$10 million) bringing together the technology, assessments, and other key elements into a coherent, unified, and balanced assessment system. The Consortium has created and implemented a comprehensive management structure consisting of an Executive Committee, a Steering Committee, and Governing States, with Washington as Lead Procurement State. Washington is accountable for managing the funds and all procurement on behalf of the Consortium. The Executive Committee will oversee the working groups and the Technical Advisory Committee (TAC) and ensure that State-specific issues are taken into account in the development of the assessment system. The Project Management Partner (\$8 million), which will be retained before the project begins, will play a key role in the effort to ensure that Consortium efforts are coordinated efficiently and effectively to realize the design, development, and implementation consistent with the Consortium’s vision, goals, and Theory of Action.

The Consortium has a robust research plan that will be directed by the Executive Committee and managed by the Research and Evaluation Working Group to ensure that the assessment system has integrity and is characterized by responsible flexibility, high technical quality, trustworthy and fair measurement components, and results that are useful for the purposes intended. The Consortium will contract with outside research specialists (\$5 million) to assist with the research and evaluation studies outlined in section (A)(5).

Building the professional capacity of teachers and administrators is central to the Consortium's Theory of Action; therefore, SBAC has a comprehensive plan to involve teachers and administrators at all phases of development and implementation of the balanced assessment system and to provide the types of tools and supports research has shown to be essential to change behaviors of teachers and improve the learning of students in the classroom. The Consortium also knows that reform will not be successful without strong support outside of schools, so it is committed to implementing a comprehensive program for communicating with a broad set of stakeholders. Professional capacity building and outreach efforts will extend from Consortium member States to districts, schools, and classrooms to promote explicit and meaningful connections between the CCSS, college- and career-readiness, instruction, assessment, and effective data use. The Consortium has budgeted \$5.125 million to contract with an organization that will develop instructional and curriculum resources for educators and \$1.5 million to work with an outside communications firm to help develop the strategy and plan for informing key stakeholders and legislatures about the assessment system and for building support for the system from the public and those stakeholders.

The Consortium recognizes the importance of working collaboratively with higher education in order to achieve the goal of better preparing students for college- and career-readiness. Achieving this goal will benefit higher education over time because students will enter IHE systems having met clear college-ready standards that are common across the Consortium States. The Consortium intends to collaborate with higher education in the creation of student achievement standards that will define college-ready. This collaboration will ensure that high school graduates will be able to track their readiness for college and careers throughout high school and will reduce the need for remediation within higher education. The Consortium expects this collaboration will ultimately result in students being on track to college- and career-ready and college- and career-ready. The Consortium will create an advisory group that will include members of the higher education community to achieve its goals of better preparing students to be college- and career-ready. This work is budgeted at \$1.5 million.

For further details regarding the SBAC assessment system budget, including an explanation of the model and assumptions used to develop the actual assessment costs, please refer to the detailed budget narrative located in Appendix A8-6.

(A)(8)(d) The projected costs for the ongoing administration, maintenance and enhancements of the SBAC summative and interim/benchmark assessments are \$19.81 and \$7.50 per pupil, respectively. Of the 31 States in the Consortium that provided data on their current assessment spending for ELA and mathematics (Appendix A8-7), the average cost (in today's dollars) per pupil cost is \$31. Of those States, 25 reported per pupil costs above the projected SBAC per pupil costs for the summative assessment and six reported costs below. Five States reported assessment costs that included interim/benchmark or formative assessments along with their summative assessment data, but only one (Michigan – \$21) reported a per pupil assessment cost close to that of the proposed SBAC summative assessment per student figure. Thus, the system will be very sustainable for the large majority of member States in the Consortium. Even without these six States, SBAC exceeds the requirement of having at least five Governing States and a minimum of 15 member (Advisory) States.

It is expected that all States in the Consortium will eliminate their current ELA and mathematics summative assessments in favor of the new SBAC assessments and allocate all funding for those portions of their current assessment systems to the new system. Additionally, the Consortium will aggressively pursue ways to decrease the costs of its new assessment system by increased use of cost savings technologies and devising procurement strategies to maximize the value received for different, key elements of the assessment system. SBAC will also be seeking assistance from foundations to help fund parts of its assessment system and will work with member States to devise a cost allocation methodology that will allow all States to participate in the Consortium at a cost no greater than that which each State pays for its current summative assessment system. It is expected that all States that have expressed a desire to join SBAC will be able to fund the ongoing operational cost for the summative assessment.

COMPETITION PRIORITIES (PART I.I)

Competitive Preference Priority: Collaboration and Alignment with Higher Education
(up to 20 points)

The Department gives eligible applicants competitive preference points based on the extent to which they have promoted collaboration and alignment between member States' public elementary and secondary education systems and their public IHEs (as defined in section 101(a) of the Higher Education Act of 1965, as amended (HEA)) or systems of those IHEs. Eligible applicants addressing this priority must provide, for each IHE or IHE system, a letter of intent that—

- (a) Commits the IHE or IHE system to participate with the consortium in the design and development of the consortium's final high school summative assessments in mathematics and English language arts in order to ensure that the assessments measure college readiness;
- (b) Commits the IHE or IHE system to implement policies, once the final high school summative assessments are implemented, that exempt from remedial courses and place into credit-bearing college courses any student who meets the consortium-adopted achievement standard (as defined in the NIA) for each assessment and any other placement requirement established by the IHE or IHE system; and
- (c) Is signed by the State's higher education executive officer (if the State has one) and the president or head of each participating IHE or IHE system.

All letters of intent must provide the total number of direct matriculation students (as defined in the NIA) in the partner IHE or IHE system in the 2008-2009 school year. An eligible applicant must also provide the total number of direct matriculation students (as defined in the NIA) in public IHEs in the consortium's member States.

The Department will award up to 20 competitive preference points based on the strength of commitment demonstrated in the letters of intent and on the percentage of direct matriculation students in public IHEs in the member States who are direct matriculation students in the partner IHEs or IHE systems. To receive full competitive preference points under this priority, eligible applicants must provide letters of intent that demonstrate strong commitment from each partner IHE or IHE system and that represent at least 30 percent of direct matriculation students in public IHEs in member States. No points will be awarded for letters of intent that represent fewer than 10 percent of direct matriculation students in public IHEs in member States.

Tables and/or Attachments:

- ✓ *Summary Table for Competitive Preference Priority*
- ✓ *Attachments: Letters of Intent*

The SMARTER Balanced Assessment Consortium recognizes that any effort to graduate all high school students ready for success in college must be collaborative, involving elementary and secondary education systems and their higher education systems. Thus, while much of the Consortium effort focuses on changes at the elementary and secondary levels, the Consortium also plans to work closely with higher education. Chiefly, SBAC intends to partner with IHEs to establish specific achievement standards for each summative high school assessment—performance standards that, when met, will be accepted by IHEs as signifying a student’s readiness to take on credit-bearing college-level work without first going through remediation. The establishment of such achievement standards will ensure that, throughout high school, students will be able to track their progress toward college readiness based on the results of their summative assessments.

To date, the Consortium has taken the following steps to promote collaboration with States’ public IHEs or IHE systems:

1. The Consortium understands the value of higher education representation on the Consortium leadership group and, thus, will include one representative from the higher education community to serve on the Executive Committee, as described in the Consortium Processes for Determining Committee/Working Group Members and Electing Leaders (see Appendix A1-3).
2. In addition, the Consortium has organized two different working groups that will focus on higher education issues. An Institutions of Higher Education policy advisory group will focus mainly on defining Consortium policies around the high school assessment and other issues related to preparing students for college. The Collaboration with Higher Education Working Group has focused around engaging higher education associations in the work of the Consortium. These two groups are shown in the Organizational Structure (see Appendix A1-1). Consortium leadership will ensure that the involvement of IHE representatives across the Executive Committee, the policy advisory group, and the working group is coordinated efficiently and effectively to address the design and development of the high school assessment, the determination of college-readiness, the

setting of achievement standards, and the creation of professional development supports, consistent with the Consortium's vision, goals, and Theory of Action.

3. The Consortium has convened meetings with higher education stakeholders to orient them to the Consortium, this grant application, and the IHEs' role in shaping the high school assessment and in determining college-readiness.
 - The Consortium has secured the support of the State Higher Education Executive Officers (SHEEO). Paul Lingenfelter, President of SHEEO, has encouraged SHEEO members to work with the Consortium.
 - The Consortium has also secured the support of the National Association of System Heads and the American Council on Education, and both have also encouraged their State members to send a letter of support.
4. The Consortium has received Letters of Intent from 162 public IHE/IHE systems across 30 States in the Consortium (described in further detail below).²⁸

Letters of Intent

As demonstrated by the signed Letters of Intent (see Appendix CPP-1), each IHE or IHE system has (1) agreed to commit the IHE or IHE system to participate with the Consortium in the design and development of the Consortium's final high school summative assessments in English language arts and mathematics in order to ensure that the assessments measure college-readiness; (2) agreed to commit the IHE or IHE system to implement policies, once the final high school summative assessments are implemented, that exempt from remedial courses and place into credit-bearing college courses any student who meets the Consortium-adopted achievement standard for each assessment and any other placement requirement established by the IHE or IHE system; and (3) provided the total number of its direct matriculation students for the 2008–09 school year. All Letters of Intent have been signed by the State's higher education executive officer (if the State has one) and/or the president or head of each participating IHE or IHE system.

As noted above, each IHE or IHE system that signed the Letter of Intent was asked to provide data for the number of its direct matriculation students for 2008–09. The data source for these numbers is the State or the actual IHE or IHE system. In some cases, the data source is the National Center for Educational Statistics Integrated Postsecondary Education Data System

²⁸ The Consortium has also secured letters of commitment from 13 private IHEs across four States.

(IPEDS). These numbers are reported in the Summary Table for the Competitive Preference Priority. Although some States also provided the total number of direct matriculation students in their State, those numbers were only used for the Summary Table if the State total was greater than the State total from IPEDS. Because IPEDS data are based on first-time, degree-seeking students in fall 2008, this number is most likely larger than the direct matriculation number, as it includes all first-time students, not just those that matriculated within two years of graduating high school. Thus, the data source for the total direct matriculation students in the State is either the total provided by the State or IPEDS, whichever is larger. By reporting the larger number, our estimate is very likely a conservative estimate, and the actual percentage is slightly higher. IPEDS data were selected as the source because IPEDS provides numbers formally reported for each IHE in the United States.

As shown in the Summary Table for the Competitive Preference Priority, the total number of direct matriculation students in public IHEs in 2008–09 across the Consortium’s member States is 995,538. Of this total number, the Consortium has secured commitments from IHEs with 739,377 direct matriculation students in 2008–09.²⁹ The total number of direct matriculation students in the Consortium’s participating IHEs represents 74% of the total number of direct matriculation students across all States in the Consortium.

²⁹ Some States also secured support from private IHEs. These Letters of Intent are included in Appendix CPP-1; however, no matriculation numbers are reported for these private institutions in the Summary Table.

Summary Table for Competitive Preference Priority

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
Washington ³⁰	Washington Higher Education Coordinating Board	Y	Y	Y	15,734	33,177	47%
Missouri ³¹	Mineral Area Community College	Y	Y	Y	523	39,325	46%
	Jefferson College	Y	Y	Y	908		
	University of Central Missouri	Y	Y	Y	1,183		
	State Fair Community College	Y	Y	Y	515		
	Linn State Technical College	Y	Y	Y	364		
	Ozarks Technical Community College	Y	Y	Y	515		
	Missouri State University	Y	Y	Y	1,918		
	St. Louis Community College System	Y	Y	Y	2,722		
	University of Missouri – Columbia & Kansas City	Y	Y	Y	2,399		
	University of Missouri – St. Louis	Y	Y	Y	1,242		
	Missouri Western State University	Y	Y	Y	525		
	Missouri Southern State University	Y	Y	Y	822		

³⁰ The data source for the number of direct matriculation students in the IHE system and the total direct matriculation students in the State is the NCES IPEDS.

³¹ The data source for the number of direct matriculation students in the IHEs is the Missouri Department of Higher Education. The data source for the total direct matriculation students in the State is the NCES IPEDS.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	Southeast Missouri State University	Y	Y	Y	298		
	Metropolitan Community College	Y	Y	Y	4,020		
Missouri (private)	Central Methodist University	Y	Y	Y	N/A	N/A	N/A
	Westminster College	Y	Y	Y			
	Avila University	Y	Y	Y			
Connecticut ³²	Connecticut Community Colleges	Y	Y	Y	7,780	19,851	92%
	Connecticut State University System	Y	Y	Y	6,162		
	University of Connecticut	Y	Y	Y	4,362		
Nevada ³³	Nevada System of Higher Education	Y	Y	Y	10,814	16,097	67%
Utah ³⁴	Utah System of Higher Education	Y	Y	Y	19,252	24,231	97%
	Utah College of Applied Technology (UCAT) System	Y	Y	Y	4,278		
Idaho ³⁵	Boise State University	Y	Y	Y	2,576	8,902	100%
	College of Southern Idaho	Y	Y	Y	1,295		
	Eastern Idaho Technical College	Y	Y	Y	76		
	Idaho State University	Y	Y	Y	1,551		
	Lewis-Clark State College	Y	Y	Y	648		

³² The State provided the number of direct matriculation students in the IHE/IHE system. The data source for the total direct matriculation students in the State is the NCES IPEDS.

³³ The State provided the number of direct matriculation students in the IHE system. The data source for the total direct matriculation students in the State is the NCES IPEDS.

³⁴ The State provided the number of direct matriculation students in the IHE systems and the total direct matriculation students in the State.

³⁵ The data source for the number of direct matriculation students in the IHEs and the total direct matriculation students in the State is the NCES IPEDS.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	North Idaho College	Y	Y	Y	1,047		
	University of Idaho	Y	Y	Y	1,709		
	College of Western Idaho	Y	Y	Y	Opened in 2009		
Maine ³⁶	Maine Maritime Academy	Y	Y	Y	234	6,593	84%
	University of Maine System	Y	Y	Y	2,082		
	Maine Community College System	Y	Y	Y	3,240		
Wisconsin ³⁷	University of Wisconsin System	Y	Y	Y	30,329	46,647	99%
	Wisconsin Technical College System	Y	Y	Y	16,243		
Wisconsin (private)	Wisconsin Association of Independent Colleges and Universities	Y	N	Y	N/A	N/A	N/A
North Carolina ³⁸	University of North Carolina – General Administration	Y	Y	Y	27,347	70,533	85%
	North Carolina Community College System	Y	Y	Y	32,423		
Oregon ³⁹	Oregon University System	Y	Y	Y	12,403	31,158	100%
	Oregon Department of Community Colleges and Workforce Development	Y	Y	Y	18,755		

³⁶ The data source for the number of direct matriculation students in the IHE/IHE system and the total direct matriculation students in the State is the NCES IPEDS.

³⁷ The data source for the number of direct matriculation students in the IHE systems and the total direct matriculation students in the State is the NCES IPEDS.

³⁸ The State provided the number of direct matriculation students in the IHE systems. The data source for the total direct matriculation students in the State is the NCES IPEDS.

³⁹ The State provided the number of direct matriculation students in the IHE and the total direct matriculation students in the State.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
New Mexico ⁴⁰	New Mexico Junior College	Y	Y	Y	395	18,115	2%
Hawaii ⁴¹	University of Hawaii System	Y	Y	Y	4,590	7,963	58%
Vermont ⁴²	–	–	–	–	–	4,470	0%
Kansas ⁴³	Kansas Board of Regents	Y	Y	Y	25,802	25,802	100%
Michigan ⁴⁴	Alpena Community College	Y	Y	Y	1,386	78,985	64%
	Central Michigan University	Y	Y	Y	3,899		
	Eastern Michigan University	Y	Y	Y	2,185		
	Ferris State University	Y	Y	Y	2,690		
	Gogebic Community College	Y	Y	Y	310		
	Kalamazoo Valley Community College	Y	Y	Y	4,218		
	Lansing Community College	Y	Y	Y	3,688		
	Macomb Community College	Y	Y	Y	1,479		
	Michigan State University	Y	Y	Y	7,337		
	Michigan Technological University	Y	Y	Y	1,378		
	Montcalm Community College	Y	Y	Y	646		
North Central Michigan College	Y	Y	Y	289			

⁴⁰ The data source for the number of direct matriculation students in the IHE and the total direct matriculation students in the State is the NCES IPEDS.

⁴¹ The IHE system provided the number of direct matriculation students. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁴² The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁴³ The data source for the number of direct matriculation students in the IHE system and the total direct matriculation students in the State is the NCES IPEDS.

⁴⁴ The State provided the number of direct matriculation students in the IHE. The data source for the total direct matriculation students in the State is the NCES IPEDS.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	Northern Michigan University	Y	Y	Y	2,030		
	Oakland University	Y	Y	Y	2,386		
	Schoolcraft College	Y	Y	Y	1,826		
	Saginaw Valley State University	Y	Y	Y	1,662		
	St. Clair County Community College	Y	Y	Y	992		
	University of Michigan	Y	Y	Y	5,768		
	Wayne State University	Y	Y	Y	2,919		
	Western Michigan University	Y	Y	Y	3,760		
Montana ⁴⁵	Montana State University System	Y	Y	Y	2,067	7,647	57%
	Montana University System ⁴⁶	Y	Y	Y	2,319		
West Virginia ⁴⁷	West Virginia Higher Education Policy Commission	Y	Y	Y	8,993	15,878	57%
Ohio ⁴⁸	University System of Ohio	Y	Y	Y	60,521	77,709	78%
Iowa ⁴⁹	Western Iowa Tech Community College	Y	Y	Y	415	26,979	57%
	Northwest Iowa Community College	Y	Y	Y	225		

⁴⁵ The State provided the number of direct matriculation students in the IHE systems. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁴⁶ Signatures for both Montana State University System and The University of Montana System are on the same signature block.

⁴⁷ The State provided the number of direct matriculation students in the IHE system. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁴⁸ The State provided the number of direct matriculation students in the IHE system. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁴⁹ The data source for the number of direct matriculation students in the IHE and the total direct matriculation students in the State is the NCES IPEDS.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	Iowa Valley Community College District	Y	Y	Y	797		
	Kirkwood Community College	Y	Y	Y	3,420		
	Hawkeye Community College	Y	Y	Y	1,259		
	North Iowa Area Community College	Y	Y	Y	839		
	Eastern Iowa Community College District	Y	Y	Y	1,469		
	Iowa Lakes Community College	Y	Y	Y	736		
	Des Moines Area Community College	Y	Y	Y	3,199		
	Northeast Iowa Community College	Y	Y	Y	597		
	Southeastern Community College	Y	Y	Y	386		
	Southwestern Community College	Y	Y	Y	236		
	Indian Hills Community College	Y	Y	Y	1,012		
	Iowa Western Community College	Y	Y	Y	836		
South Carolina ⁵⁰	College of Charleston	Y	Y	Y	4,126	57,741	86%
	Winthrop University	Y	Y	Y	2,026		
	Lander University	Y	Y	Y	1,011		
	Francis Marion University	Y	Y	Y	1,362		

⁵⁰ The State provided the number of direct matriculation students in the IHE and the total direct matriculation students in the State.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	Clemson University	Y	Y	Y	5,880		
	University of South Carolina System ⁵¹	Y	Y	Y	11,821		
	Spartanburg Community College	Y	Y	Y	1,784		
	Florence-Darlington Technical College	Y	Y	Y	1,515		
	Orangeburg-Calhoun Technical College	Y	Y	Y	855		
	Technical College of the Lowcountry	Y	Y	Y	420		
	York Technical College	Y	Y	Y	1,469		
	Northeastern Technical College	Y	Y	Y	324		
	Williamsburg Technical College	Y	Y	Y	153		
	Aiken Technical College	Y	Y	Y	910		
	Trident Technical College	Y	Y	Y	3,593		
	Denmark Technical College	Y	Y	Y	429		
	Piedmont Technical College	Y	Y	Y	1,357		
	Midlands Technical College	Y	Y	Y	3,578		
	Horry-Georgetown Technical College	Y	Y	Y	1,724		
Greenville Technical College	Y	Y	Y	4,420			
Central Carolina Technical College	Y	Y	Y	816			
South Dakota ⁵²	South Dakota Regental System	Y	Y	Y	5,125	7,081	72%

⁵¹ Individual letters from IHEs that are part of this System (Upstate, Beaufort, and Aiken) also submitted Letters of Intent.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
Colorado ⁵³	Colorado Higher Education System	Y	Y	Y	39,862	39,862	100%
North Dakota ⁵⁴	Bismarck State College	Y	Y	Y	308	7,708	75%
	Dickinson State University	Y	Y	Y	203		
	Lake Region State College	Y	Y	Y	62		
	Mayville State University	Y	Y	Y	86		
	Minot State University	Y	Y	Y	319		
	Dakota College at Bottineau	Y	Y	Y	59		
	North Dakota State College of Science	Y	Y	Y	141		
	North Dakota State University	Y	Y	Y	2,525		
	University of North Dakota	Y	Y	Y	1,858		
	Valley City State University	Y	Y	Y	123		
Williston State College	Y	Y	Y	119			
Delaware ⁵⁵	Delaware Technical & Community College (all campuses)	Y	Y	Y	3,297	7,889	100%
	University of Delaware	Y	Y	Y	3,862		
	Delaware State University	Y	Y	Y	730		
Alabama ⁵⁶	Alabama Southern Community College	Y	Y	Y	362	41,372	76%

⁵² The State provided the number of direct matriculation students in the IHE system. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁵³ The data source for the number of direct matriculation students in the IHE system and the total direct matriculation students in the State is the NCES IPEDS.

⁵⁴ The State provided the number of direct matriculation students in the IHEs. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁵⁵ The data source for the number of direct matriculation students in the IHEs and total direct matriculation students in the State is the NCES IPEDS.

⁵⁶ The data source for the number of direct matriculation students in the IHEs is the Alabama Statewide Student Database. Note: Totals provided do not include 22 students with unknown date of birth. The data source for the total direct matriculation students in the State is the NCES IPEDS.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	Bevill State Community College	Y	Y	Y	682		
	Bishop State Community College	Y	Y	Y	392		
	Calhoun Community College	Y	Y	Y	1,232		
	Central Alabama Community College	Y	Y	Y	396		
	Chattahoochee Valley Community College	Y	Y	Y	247		
	J. F. Drake State Technical College	Y	Y	Y	87		
	Enterprise State Community College	Y	Y	Y	389		
	Gadsden State Community College	Y	Y	Y	1,081		
	Jefferson Davis Community College	Y	Y	Y	187		
	Jefferson State Community College	Y	Y	Y	1,138		
	Lawson State Community College	Y	Y	Y	607		
	Lurleen B. Wallace Community College	Y	Y	Y	252		
	Marion Military Institute	Y	Y	Y	261		
	Northeast Alabama Community College	Y	Y	Y	495		
	Northwest-Shoals Community College	Y	Y	Y	633		
	Reid State Technical College	Y	Y	Y	79		

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	Shelton State Community College	Y	Y	Y	1,060		
	Snead State Community College	Y	Y	Y	492		
	H. Councill Trenholm State Technical College	Y	Y	Y	172		
	Wallace Community College – Dothan	Y	Y	Y	667		
	Wallace State Community College – Hanceville	Y	Y	Y	960		
	Wallace Community College – Selma	Y	Y	Y	332		
	Alabama A & M University	Y	Y	Y	1,018		
	Alabama State University	Y	Y	Y	1,262		
	Auburn University	Y	Y	Y	3,965		
	Auburn University – Montgomery	Y	Y	Y	577		
	Jacksonville State University	Y	Y	Y	1,159		
	Troy University	Y	Y	Y	1,482		
	University of Alabama	Y	Y	Y	5,069		
	University of Alabama – Birmingham	Y	Y	Y	1,266		
	University of Alabama – Huntsville	Y	Y	Y	767		
	University of North Alabama	Y	Y	Y	1,008		
	University of South Alabama	Y	Y	Y	1,470		
University of West Alabama	Y	Y	Y	278			
Alabama (private)	Stillman College	Y	Y	Y	N/A	N/A	N/A
	Talladega College	Y	Y	Y			
	Concordia College – Selma	Y	Y	Y			

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
	Stamford University	Y	Y	Y			
Kentucky ⁵⁷	Kentucky Council on Postsecondary Education	Y	Y	Y	22,744	30,341	75%
New Hampshire ⁵⁸	University System of New Hampshire	Y	Y	Y	5,255	7,897	100%
	Community College System of New Hampshire	Y	Y	Y	2,642		
Pennsylvania ⁵⁹	Pennsylvania State System of Higher Education	Y	Y	Y	21,625	79,284	79%
	Pennsylvania Commission for Community Colleges	Y	Y	Y	40,755		
Oklahoma ⁶⁰	Oklahoma State System of Higher Education	Y	Y	Y	24,249	30,390	80%
New Jersey ⁶¹	Ocean County College	Y	Y	Y	3,915	55,933	16%
	Montclair State University	Y	Y	Y	2,293		
	Camden County College	Y	Y	Y	2,594		
New Jersey (private)	Rider University	Y	Y	Y	N/A	N/A	N/A
	Monmouth University	Y	Y	Y			
	Georgian Court University	Y	Y	Y			
	Bloomfield College	Y	Y	Y			
	Caldwell College	Y	Y	Y			

⁵⁷ The State provided the number of direct matriculation students in the IHE system. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁵⁸ The data source for the number of direct matriculation students in the IHE systems and the total direct matriculation students in the State is the NCES IPEDS.

⁵⁹ The State provided the number of direct matriculation students in the IHE systems. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁶⁰ The State provided the number of direct matriculation students in the IHE system. The data source for the total direct matriculation students in the State is the NCES IPEDS.

⁶¹ The State provided the number of direct matriculation students in the IHEs. The data source for the total direct matriculation students in the State is the NCES IPEDS.

State (List all in Consortium)	Name of Participating IHE or IHE System	IHE Committed to Participating w/Consortium? (Y/N)	IHE Committed to Exempting Students from Remediation? (Y/N)	LOI Signed by IHE Leader(s)? (Y/N)	Number of Direct Matriculation Students in IHE in 2008-2009	Total Direct Matriculation Students in State in 2008-2009	Direct Matriculation Students in Participating IHEs as % of State Total
Georgia ⁶²	University System of Georgia	Y	Y	Y	51,353	69,978	92%
	Technical College System of Georgia	Y	Y	Y	12,921		
TOTAL	N/A	N/A	N/A	N/A	739,377	995,538	74%

⁶² The State provided the number of direct matriculation students in the IHE systems. The data source for the total direct matriculation students in the State is the NCES IPEDS.

References

- Allalouf, A. (2007). An NCME Instructional Module on: Quality control procedures in the scoring, equating and reporting of test scores. *Educational Measurement: Issues and Practice*, 26, 36–43.
- Almond, P., Winter, P., Cameto, R., Russell, M., Sato, E., Clarke, J., . . . Lazarus, S. (2010). *Technology enabled and universally designed assessment: Considering access in measuring the achievement of students with disabilities—A foundation for research*. Dover, NH: Measured Progress; Menlo Park, CA: SRI International.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for psychological testing*. Washington, DC: American Educational Research Association.
- Association of Test Publishers (ATP), & the Council of Chief State School Officers (CCSSO). (in press). *Operational best practices for statewide large scale assessment programs*. Retrieved June 10, 2010, from http://www.ccsso.org/best_practices/index.html
- Baker, E. L. (2007). Model-based assessments to support learning and accountability: The evolution of CRESST's research on multiple-purpose measures. *Educational Assessment*, 12(3&4), 179–194.
- Baker, E. L., & Herman, J. (2010, May). *Assessment planning*. Presented at the SMARTER Balanced Assessment Consortium Advisory Panel Meeting, Stanford, CA.
- Bennett, R. E. (2004). How the Internet will help large-scale assessment reinvent itself. In F. C. Blumberg, H. T. Everson, & M. Rabinowitz (Eds.), *The design of instruction and evaluation: Affordances of using media and technology* (pp. 101–128). Mahwah, NJ: Erlbaum.
- Betebenner, D. A. (2009). Norm- and criterion-referenced student growth. *Educational Measurement: Issues and Practice*, 28(4), 44–51.
- Betebenner, D. A., & Linn, R. L. (2010). *Growth in student achievement: Issues of measurement, longitudinal data analysis and accountability*. Retrieved June 1, 2010, from <http://www.k12center.org/publications.html>
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum.

- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.
- Council of Chief State School Officers (CCSSO), Technical Issues in Large-Scale Assessment (TILSA), & State Collaborative on Assessment and Student Standards (SCASS). (2003a). *Quality Control Checklist for Processing, Scoring, and Reporting*. Retrieved June 1, 2010, from <http://www.ccsso.org/content/pdfs/scorereportQCchk1st.pdf>
- Council of Chief State School Officers (CCSSO), Technical Issues in Large-Scale Assessment (TILSA) & State Collaborative on Assessment and Student Standards (SCASS). (2003b). *Quality Control Checklist for Item Development and Test Form Construction*. Retrieved June 1, 2010, from www.ccsso.org/content/pdfs/ItemandTestDevQCChk1st.pdf
- Darling-Hammond, L., & Pecheone, R. (2010). *Developing an internationally comparable balanced assessment system that supports high-quality learning*. Presented at the National Conference on Next Generation K–12 Assessment Systems, Center for K–12 Assessment & Performance Management with the Education Commission of the States (ECS) and the Council of Great City Schools (CGCS), Washington, DC.
- Dunn, J. L., & Allen, J. (2009). Holding schools accountable for the growth of nonproficient students: Coordinating measurement and accountability. *Educational Measurement: Issues and Practice*, 28(4), 27–41.
- ETS. (2010). *Using natural language processing (NLP) and psychometric methods to develop innovative scoring technologies*. Retrieved June 11, 2010, from <http://www.ets.org/Media/Home/pdf/AutomatedScoring.pdf>
- Gong, B., Perie, M., & Dunn, J. (2006). *Using student longitudinal growth measures for school accountability under No Child Left Behind: An update to inform design decisions*. Center for Assessment. Retrieved June 1, 2010, from www.nciea.org
- Heritage, M. (2007). *Learning progressions: Supporting instruction and formative assessment*. Paper prepared for the Formative Assessment for Students and Teachers (FAST), State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO). Retrieved June 1, 2010, from <http://www.k12.wa.us/assessment/ClassroomAssessmentIntegration/pubdocs/FASTLearningProgressions.pdf>

- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher*, 25(4), 12–21. Retrieved June 12, 2010, from <http://www.jstor.org/stable/1176776>
- International Test Commission. (2005). *International guidelines on computer-based and Internet delivered testing*. Retrieved June 1, 2010, from http://www.intestcom.org/_guidelines/guidelines/index.html
- Joint Committee on Standards for Educational Evaluation (JCSEE). (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- Joint Committee on Standards for Educational Evaluation (JCSEE). (2002). *The student evaluation standards: How to improve evaluations of students*. Thousand Oaks, CA: Corwin Press.
- Joint Committee on Standards for Educational Evaluation (JCSEE). (2008). *The personnel evaluation standards: How to assess systems for evaluating educators*. Thousand Oaks, CA: Corwin Press.
- Kane, M. (2007). Terminology, emphasis, and utility in validation. *Educational Researcher*, 37(2), 76–82.
- Kingsbury, G.C. and Hauser, C. (2004). *Computerized adaptive testing and No Child Left Behind*. Northwest Evaluation Association. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Kingston, N., Kahl, S., Sweeney, K., & Bay, L. (2001). Setting performance standards using the body of work method. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives*. (pp. 219–248). Mahwah, NJ: Erlbaum.
- Kopriva, R. J., Emick, J. E., Hipolito-Delgado, C. P., & Cameron, C. A. (2007). Do proper accommodation assignments make a difference? Examining the impact of improved decision making on scores for English language learners. *Educational Measurement: Issues and Practice*, 26(3), 11–20.
- Kosty, D., McBride, J., Poggio, J., Wise, L., & Way, D. (2006). *What's next in online testing*. Presentation from the 36th Annual National Conference on Large-Scale Assessment, San Francisco, CA. Retrieved November 4, 2008, from <http://www.ccsso.org/content/PDFs/Session38Way.pdf>

- Lilley, M., Barker, T., & Britton, C. (2004). The development and evaluation of a software prototype for computer-adaptive testing. *Computers & Education*, 43(1–2), 109–123.
- Martineau, J. A. (2006). Distorting value added: The use of longitudinal, vertically scaled student achievement data for growth-based, value-added accountability. *Journal of Educational and Behavioral Statistics*, 31(1), 35–62.
- Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York, NY: W. H. Freeman.
- McCall, M., & Hauser, C. (2007). Item response theory and longitudinal modeling: The real world is less complicated than we fear. In R. W. Lissitz (Ed.), *Assessing and modeling cognitive development in schools: Intellectual growth and standard setting*. Maple Grove, MN: JAM Press.
- McLaughlin, M. (2005). Listening and learning from the field: Tales of policy implementation and situated practice. In A. Lieberman (Ed.), *The roots of educational change* (pp. 58–72). New York, NY: Springer.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). *On the structure of educational assessments*. (CSE Report 597). Los Angeles: The National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education & Information Studies, UCLA. <http://www.cse.ucla.edu/products/Reports/TR597.pdf>
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). *A brief introduction to evidence-centered design*. (CSE Report 632). Los Angeles: The National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education & Information Studies, UCLA. <http://www.cse.ucla.edu/products/reports/r632.pdf>
- Mislevy, R. J., & Riconscente, M. M. (2006). Evidence-centered assessment design: Layers, concepts, and terminology. In S. Downing, & T. Haladyna (Eds.), *Handbook of test development* (pp. 61–90). Mahwah, NJ: Erlbaum.
- Mitzel, H., Lewis, D., Patz, R., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249–281). Mahwah, NJ: Erlbaum.
- National Research Council (NRC). (1999). *How people learn: Brain, mind, experience, and school*. J. D. Bransford, A. L. Brown, & R. R. Cocking (Eds.). Committee on Developments in the Science of Learning. Washington, DC: National Academies Press.

- National Research Council (NRC). (2001). *Knowing what students know: The science and design of educational assessment*. J. Pellegrino, N. Chudowsky, & R. Glaser (Eds.). Committee on the Foundations of Assessment. Washington, DC: National Academies Press.
- National Staff Development Council (NSDC). (2009). Ensure great teaching for every student. *NSDC Policy Points, 1(2)*. Retrieved June 11, 2010, from <http://www.nsd.org/news/policypoints/policypoints4-09.pdf>
- Obama, B. (2009, March 10). *Remarks by the President to the Hispanic Chamber of Commerce on a complete and competitive American education*. Retrieved from http://www.whitehouse.gov/the_press_office/Remarks-of-the-President-to-the-United-States-Hispanic-Chamber-of-Commerce
- Oregon Department of Education (2009). 2008–2009 Technical Report: Oregon’s Statewide Assessment System. Annual Report.
- Pellegrino, J. W. (2006). *Rethinking and redesigning curriculum, instruction and assessment: What contemporary research and theory suggests*. Retrieved Jan. 5, 2009, from <http://www.skillscommission.org/commissioned.htm>
- Popham, W. J. (2006, October). *Defining and enhancing formative assessment*. Paper presented at the Annual Large-Scale Assessment Conference, Council of Chief State School Officers, San Francisco, CA.
- Popham, W. J. (2007). The lowdown on learning progressions. *Educational Leadership, 64(7)*, 83–84.
- Quellmalz, E. S., & Moody, M. (2004). *Models for multi-level state science assessment systems*. Report commissioned by the National Research Council Committee on Test Design for K-12 Science Achievement.
- Quellmalz, E. S., Timms, M. J., & Buckley, B. C. (2009, November). *Recommendations for technology and innovation in assessment*. Race to the Top Assessment Program Public and Expert Input Meeting—Technology and Innovation, Boston, MA.
- Rabinowitz, S. N. (2005). *Leading edge U.S. computer-based assessment developments and issues*. Presentation at the TASA Institute: Innovation in Testing Technology for Canadian Schools. Victoria, BC.
- Rabinowitz, S. (2009). *Assumptions about assessments*. San Francisco, CA: WestEd.

- Rabinowitz, S. (2010). Commentary: Next-generation assessment systems. *Education Week*, 29(22), 36.
- Reckase, M. D. (2006). A conceptual framework for a psychometric theory for standard setting with examples of its use for evaluating the functioning of two standard setting methods. *Educational Measurement*, 25(2), 4–18.
- Resnick, L. B. (1987). *Education and learning to think*. Washington, DC: National Academies Press.
- Resnick, L. B., & Berger, L. (2010). *An American examination system*. National Conference on Next Generation K–12 Assessment Systems, Center for K–12 Assessment & Performance Management with the Education Commission of the States (ECS) and the Council of Great City Schools (CGCS), Washington, DC.
- Rivera, C., Acosta, B., & Willner, L. (2008). *Guide for refining state assessment policies for accommodating English language learners*. Arlington, VA: The George Washington University Center for Equity and Excellence in Education. Retrieved from <http://ells.ceee.gwu.edu/>
- Russell, M., Goldberg, A., & O'Connor, K. (2003). Computer-based testing and validity: A look back into the future. *Assessment in Education: Principles, Policy and Practice*, 10, 279–293.
- Sato, E., Rabinowitz, S., Gallagher, C., Huang, Chun-Wei (2010). *Accommodations for English language learner students: the effect of linguistic modification of math test item sets*. The National Center for Education Evaluation and Regional Assistance, Washington, D.C.
- Schmidt, W. H., Houang, R. T., & McKnight, C. C. (2005). Value-added research: Right idea but wrong solution? In R. Lissitz (Ed.), *Value added models in education: Theory and applications* (pp. 145–164). Maple Grove, MN: JAM Press.
- Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of U.S. mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, 37(5), 525–559.
- Shepard, L., Hammerness, K., Darling-Hammond, L., & Rust, F., with Baratz Snowden, J., Gordon, E., Gutierrez, C., & Pacheco, A. (2005). Assessment. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 275–326). San Francisco, CA: Jossey-Bass.

- Southern Regional Education Board (SREB). (2009). *Ready for tomorrow: Six proven ideas to graduate and prepare more students for college and 21st-century careers*. Retrieved June 1, 2010, from http://publications.sreb.org/2009/09V20_Ready_for_Tomorrow.pdf
- Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *PhiDelta Kappan*, 83(10), 758–765.
- Thompson, S., Thurlow, M., & Malouf, D. B. (2004). Creating better tests for everyone through universally designed assessments. *Journal of Applied Testing Technology*. Retrieved June 1, 2010, from http://www.testpublishers.org/Documents/Creating_Better_Tests%20Final%20Revision%205.15.04.pdf
- Way, W., Twing, J., Camara, W., Sweeney, K., Lazer, S., & Mazzeo, J. (2010). Some considerations related to the use of adaptive testing for the Common Core Assessments. Retrieved June 11, 2010, from www.ets.org/s/commonassessments/pdf/AdaptiveTesting.pdf
- Zenisky, A. L., Hambleton, R. K., & Sireci, S. G. (2009). Getting the message out: An evaluation of NAEP score reporting practices with implications for disseminating test results. *Applied Measurement in Education*, 22(4), 359–375.