CTB/McGraw-Hill LLC presents this Proposal to
the Washington Office of Superintendent of Public Instruction
April 12, 2012
Table of Contents

Letter of Submittal
   Exhibit A Certifications and Assurances
   Exceptions

Introduction ................................................................................................................................... 1
Technical Proposal ........................................................................................................................ 1

OBJECTIVE AND SCOPE OF WORK .................................................................................... 1

2.1 Submit a draft schedule of tasks and activities as a part of the bid ................................................1
2.2 Submit a final detailed schedule of tasks and activities .................................................................1
2.3 Provide weekly status reports ......................................................................................................2
2.4 Schedule and facilitate a contractor kick-off meeting upon execution of the contract ..............2
2.5 Project the initial ALDs, when complete, in print ready format as well as digital formats ....2
2.6 Solicit volunteers and select, in collaboration with the Consortium leadership ..................3
2.7 Develop the methodology for designing ALDs ............................................................................5
2.8 Panelists to follow the methodology in designing initial ALDs ............................................. 24
2.9 Edit the draft initial ALDs arising from the development sessions .......................................... 25
2.10 Elicit and collect comprehensive feedback regarding the draft initial ALDs ....................... 25
2.11 Revise the draft initial ALDs based on Consortium states’ feedback ..................................... 27
2.12 Elicit and collect feedback from the Consortium’s TAC and the public ..................................27
2.13 Revise the draft initial ALDs based on TAC feedback and public feedback .........................28
2.14 Present at one TAC meeting before finalizing the initial ALDs and technical report ..........33
2.15 Revise initial ALDs to provide any needed consistency of initial ALDs with PARCC ..........33
2.16 Gather validity evidence to be included in the final technical report .................................... 34
2.17 Draft a final technical report ......................................................................................................36

A. Project Approach/Methodology ........................................................................................... 38

Oversight, Panelist Recruitment, and Methodological Design .......................................................38
Workshop Implementation, Draft Initial ALDs, and Draft Technical Report .................................44
Review Cycles and Validation Studies ..........................................................................................49
Final Technical Report and Final Initial ALDs ............................................................................50

B. Work Plan and Schedule ...................................................................................................... 52

C. Deliverables .......................................................................................................................... 55
EXHIBIT K: Deliverables Matrix ........................................................................................................................ 55

D. Outcomes and Performance Measurement .................................................................................................................. 59

E. Risks ............................................................................................................................................................................... 61

Management Proposal ....................................................................................................................................................... 1

A. Project Management ......................................................................................................................................................... 1

1. Project Team Structure .................................................................................................................................................... 1

2. Project Management Capabilities ................................................................................................................................. 3

3. Project Management Deliverables .................................................................................................................................. 3

4. Staff Qualifications/Experience ..................................................................................................................................... 9

B. Experience of the Vendor .................................................................................................................................................... 20

General Capabilities ............................................................................................................................................................. 20

1. Relevant experience that indicates the qualifications for the performance of the contract. ....................................... 24

2. List of contracts that relate to the ability to perform the services .................................................................................. 26

C. References ........................................................................................................................................................................ 26

D. Contractor Intake Form ....................................................................................................................................................... 30

Cost Proposal ......................................................................................................................................................................... 1

Appendices

Appendix A: Schedule

Appendix B: ALD Framework Discussion

Appendix C: Résumés
Dear Mr. Middleton:

In response to the State of Washington, Office of Superintendent of Public Instruction, Olympia, Washington Request for Proposal (RFP) SBAC-12 please find attached the CTB-College Board (Collaborative) proposal for the "SMARTER Balanced Assessment Consortium Initial Achievement Level Descriptor (ALDs) Development" Request for Proposal (RFP).

CTB and the College Board look forward to the opportunity to work with the SMARTER Balanced Assessment Consortium in preparing the way for rich assessments that promote student learning by supporting the development of initial ALDs. We believe our proposed work with College Board has the potential to create outstanding ALDs for SMARTER Balanced and state membership. In particular, the Collaborative presents a very thoughtful solution that is backed by proven success in the following areas:

- Successful experience working with states in developing ALDs
- Nationally recognized research, publications and in-house experts in this area
- High success rate in meeting all of the federal peer review qualifications regarding ALDs for contracts
- Overall, nationally recognized experts in assessment, design, research, and college readiness

CTB/McGraw-Hill is a Limited Liability Company and a subsidiary of The McGraw-Hill Companies. It was first established in 1926. CTB was purchased by The McGraw-Hill Companies in 1965 and has operated in substantially the same manner since this time. CTB/McGraw-Hill's Federal Tax Identification number is: 52-2358325. Our Washington Uniform Business Identification number is: 602 190 023.

This letter is signed by Sandor Z. Nagy, Chief Operating Officer of CTB/McGraw-Hill, who is authorized to legally bind CTB to a contractual relationship with the State of Washington. CTB is located at 20 Ryan Ranch Road, Monterey, California 93940, which is our principal place of business. The main exchange for this location is 831.393.0700.

Our Collaborative looks forward to the opportunity to work with SMARTER Balanced in developing and providing the services, resources, and products set forth in SBAC-12. If you would like additional information or clarification about the CTB response, please do not hesitate to contact Dr. Janice "Jan" Barth, State Solutions Manager, at 304.941.9061, or by email at Janice_Barth@ctb.com.

Sincerely,

Sandor Z. Nagy

Attachment 1: CTB and proposed subcontractors required information

Enclosures: CTB's response to RFP SBAC-12 is enclosed. Please see the Table of Contents for a list of enclosures to the proposal.
### Attachment 1

<table>
<thead>
<tr>
<th>Name of Legal Entity</th>
<th>CTB/McGraw-Hill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address, Phone, email of individual whom contract would be written</td>
<td>20 Ryan Ranch Rd., Monterey, CA 93940</td>
</tr>
<tr>
<td>Phone:</td>
<td>831.393.7070</td>
</tr>
<tr>
<td>Sandor Nagy, Chief Operating Officer</td>
<td><a href="mailto:sandor_nagy@ctb.com">sandor_nagy@ctb.com</a></td>
</tr>
<tr>
<td>Name of Contact Person</td>
<td>Jan Barth</td>
</tr>
<tr>
<td>Principal Officers</td>
<td>Ellen Haley President 831.393.7757</td>
</tr>
<tr>
<td></td>
<td>Sandor Nagy Chief Operating Officer 831.393.7070</td>
</tr>
<tr>
<td></td>
<td>Mark Limbach Vice President Finance 831.393.6336</td>
</tr>
<tr>
<td></td>
<td>David Seitter Vice President Sales 212.904.6392</td>
</tr>
<tr>
<td>Location of Facility which Vendor would operate</td>
<td>20 Ryan Ranch Rd., Monterey, CA 93940</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Legal Entity</th>
<th>College Entrance Examination Board (dba The College Board)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address, Phone, email of individual whom contract would be written</td>
<td>45 Columbus Avenue, New York, NY 10023</td>
</tr>
<tr>
<td>Phone:</td>
<td>212.373.8791</td>
</tr>
<tr>
<td><a href="mailto:sharris@collegeboard.org">sharris@collegeboard.org</a></td>
<td></td>
</tr>
<tr>
<td>Name of Contact Person</td>
<td>Shelly Harris, Director, State and District Response Administration</td>
</tr>
<tr>
<td>Principal Officers</td>
<td>Gaston Caperton President and Chief Executive Officer</td>
</tr>
<tr>
<td></td>
<td>Herb Elish Chief Operating Officer</td>
</tr>
<tr>
<td></td>
<td>Wayne Camara Vice President Research &amp; Development</td>
</tr>
<tr>
<td></td>
<td>Diane Duggan Senior Vice President and Chief Information Officer</td>
</tr>
<tr>
<td></td>
<td>Hal Higginbotham Senior Vice President</td>
</tr>
<tr>
<td></td>
<td>Thomas Higgins Chief Financial Officer and Senior Vice President, Finance</td>
</tr>
<tr>
<td></td>
<td>Peter Kauffmann Vice President, Communications</td>
</tr>
<tr>
<td></td>
<td>Neil Lane Senior Vice President, Administration and General Counsel</td>
</tr>
<tr>
<td></td>
<td>Kanika Lichthardt Vice President and Chief of Staff</td>
</tr>
<tr>
<td></td>
<td>Andrea Mainelli Senior Vice President, Region and Account Services</td>
</tr>
<tr>
<td></td>
<td>Steve Meyer Vice President, Office of Strategy Management</td>
</tr>
<tr>
<td></td>
<td>Peter Negroni Senior Vice President, Relationship Development</td>
</tr>
<tr>
<td></td>
<td>Trevor Packer Senior Vice President, AP and College Readiness</td>
</tr>
<tr>
<td></td>
<td>Tom Rudin Senior Vice President, Advocacy, Government Relations and Development</td>
</tr>
<tr>
<td></td>
<td>Juliet Weissman Vice President, Organizational Effectiveness</td>
</tr>
<tr>
<td></td>
<td>Mary Carroll Scott Vice President, Membership</td>
</tr>
<tr>
<td></td>
<td>Dorothy Sexton Vice President, Operations</td>
</tr>
<tr>
<td></td>
<td>Ranjit Sidhu Vice President, Governance and Secretary of the Corporation</td>
</tr>
<tr>
<td>Legal Status</td>
<td>Non-profit organization, established in New York, NY in 1957.</td>
</tr>
<tr>
<td>Location of Facility which Vendor would operate</td>
<td>The College Board, 45 Columbus Avenue, New York, NY 10023</td>
</tr>
</tbody>
</table>
EXHIBIT A
CERTIFICATION AND ASSURANCES

1/we make the following certifications and assurances as a required element of the proposal to which it is attached, understanding that the truthfulness of the facts affirmed here and the continuing compliance with these requirements are conditions precedent to the award or continuation of the related contract(s):

1. 1/we declare that all answers and statements made in the proposal are true and correct.
2. The prices and/or cost data have been determined independently, without consultation, communication, or agreement with others for the purpose of restricting competition. However, 1/we may freely join with other persons or organizations for the purpose of presenting a single proposal.
3. The attached proposal is a firm offer for a period of 60 days following receipt, and it may be accepted by OSPI without further negotiation (except where obviously required by lack of certainty in key terms) at any time within the 60-day period.
4. In preparing this proposal, 1/we have not been assisted by any current or former employee of the state of Washington whose duties relate (or did relate) to this proposal or prospective contract, and who was assisting in other than his or her official, public capacity. (Any exceptions to these assurances are described in full detail on a separate page and attached to this document.)
5. 1/we understand that OSPI will not reimburse me/us for any costs incurred in the preparation of this proposal. All proposals become the property of OSPI, and 1/we claim no proprietary right to the ideas, writings, items, or samples, unless so stated in this proposal.
6. Unless otherwise required by law, the prices and/or cost data which have been submitted have not been knowingly disclosed by the Proposer and will not knowingly be disclosed by him/her prior to opening, directly or indirectly to any other Proposer or to any competitor.
7. 1/we agree that submission of the attached proposal constitutes acceptance of the solicitation contents and the attached sample contract and general terms and conditions. If there are any exceptions to these terms, 1/we have described those exceptions in detail on a page attached to this document.
8. No attempt has been made or will be made by the Proposer to induce any other person or firm to submit or not to submit a proposal for the purpose of restricting competition.
9. 1/we grant OSPI the right to contact references and others, who may have pertinent information regarding the Proposer's prior experience and ability to perform the services contemplated in this procurement.

On behalf of the firm submitting this proposal, my name below attests to the accuracy of the above statements.

Signature of Proposer  Sandor Z. Nagy

Chief Operating Officer  April 51 2012
Title  Date
CTB/McGraw-Hill’s Contractual Clarifications and Exceptions
For State of Washington Office of the Superintendent of Public Instruction (OSPI)
RFP No. SBAC-12

SMARTER Balanced Assessment Consortium Request for Proposals (RFP) for SBAC Initial Achievement Level Descriptors.

It is CTB/McGraw-Hill LLC’s (“CTB”) understanding that a mutually agreed upon contract (the "Contract") will be negotiated between the OSPI and the winning bidder ("Contractor"), and the Contract will be substantially in the form of Exhibit B to the State of Washington Office of the Superintendent of Public Instruction’s (the “OSPI” or the “Superintendent”) Request for Proposals No. SBAC-12 (the “RFP”), including Contractor’s proposal (“Proposal”). Pursuant to RFP Section B, Clause II, “Contract and General Terms & Conditions” (page 13), and Clauses 3, 5, and 7 of Exhibit A “Certification and Assurances” (page 27), CTB is submitting the following with the intention of clarifying its intent with respect to the indicated provisions, and wording similar to that set forth below (subject to the changes mutually agreed upon by the parties) will be included in any resulting Contract between the OSPI and CTB as Special Terms and Conditions.

Exhibit B: Contract for Personal Services:
1. Clause I. “Duties of the Contractor” (page 28); Clause III “Duties of The Superintendent” (page 29)
   To clarify these sections, in Clause I, Section B, the first sentence shall be modified as follows: "In order to accomplish the general objectives(s) of this Contract, Contractor shall perform the following specific duties in conformance with the requirements of this Contract.” Clause III, Section B, last paragraph shall include, as a new second sentence: “Approval and acceptance will be granted for services provided in accordance with this Contract.”

2. Clause I. “Duties of the Contractor” (page 28)
   To clarify this section, the following will be added to the end of paragraph C: “Contractor shall not be responsible for any delay or compromise caused by persons or circumstances not under Contractor’s control including without limitation riot, war, declared emergency, statutory or regulatory enactment, flood, fire or act of God, or acts or failures to act on the part of OSPI or the Consortium.”

3. Clause IV. “Renewal (Optional Clause)” (page 30)
   To clarify this section, the second sentence will be modified as follows: “If OSPI provides such notice to the Contractor, the Contractor and the OSPI shall negotiate any revision of additional services, goals, and compensation beyond those encompassed in the previous contract.”

Exhibit C: Service Level Requirements and Remedies (pages 32-34):
4. With respect to these provisions the parties acknowledge and agree as follows:
   (a) Financial remedies shall not be assessed if the failure of the Contractor (or any subcontractors or vendors of Contractor) to complete a Key Service Requirement within the agreed upon timeline is caused by force majeure, as defined in this Contract. The amount of any financial remedies shall be based upon a good faith estimate of actual damages the OSPI would sustain if a Key Deliverable Date is missed or delayed due to causes under the Contractor’s control, and that the total amount of financial remedies for all non-conforming Key Activities during the term of the Contract shall not exceed ten percent (10%) of the total Contract price.
   (b) The OSPI will notify the Contractor in writing if the OSPI believes the Contractor is not in
compliance with any provision of the Contract, so the parties may allow a reasonable time for Contractor to cure such non-compliance before terminating the Contract for default, notwithstanding any term or provision in the Contract to the contrary.

(c) The Contractor will not be held liable or deemed in default under the Contract if and to the extent that Contractor's performance is prevented by events beyond its (or any of its vendors, suppliers or subcontractors') reasonable control, including force majeure events, which are deemed to include without limitation, acts or failure to act on the part of the OSPI, SBAC, any of SBAC's Governing States, or other OSPI or SBAC contractors (collectively, "OSPI"), such as failures to provide necessary materials or information, required approvals, or changes in the scope of work, or any other obligation upon which the Contractor's performance depends. In such cases the Contractor's time to perform shall be extended for the same period as the delay caused by the force majeure event. The parties further agree that in the event a force majeure event is caused by OSPI, the OSPI shall be responsible for providing additional funding to Contractor to cover the costs of any additional work or delays caused by such acts or failures to act on the part of OSPI. Such additional work, changed deadlines, and payments due to Contractor will be negotiated in good faith by the parties and set forth in an amendment to the Contract signed by both parties.

Exhibit D: Attachment A, Contract for Personal Services, “General Terms and Conditions”:

5. Clause 10 “Confidentiality” (page 36)

Add these standard exemptions to the definition of Confidentiality: “The obligations of Clause 10 do not apply to any information that: (a) is already in the public domain or becomes available to the public through no breach of this Contract by the receiving party; (b) was lawfully in the receiving party’s possession prior to receipt from the disclosing party; (c) is received independently from a third party free to lawfully disclose such information to the receiving party; (d) is subsequently independently developed by the receiving party; or (e) is required to be disclosed by operation of law, provided that the disclosing party is promptly notified in writing prior to such disclosure.”

6. Clause 10 “Confidentiality” (page 36), Clause 11 “Copyright Provisions” (page 36); Clause 38 “Termination Procedure” (page 42), and Clause 39 “Treatment of Assets” (page 43)

To clarify these provisions, add the following paragraph: “While the parties agree all deliverables under this Contract shall be OSPI’s sole property, it is also agreed the Contractor (including its subcontractors) will be using pre-existing secure, confidential, trade secret, and proprietary intellectual property in performing its obligations under this Contract, including without limitation, Contractor’s templates, development tools, research, software, documentation, know-how, forms, blueprints, processes, procedures, methodologies, and ancillary and related materials, including their revisions and derivative works (all herein collectively the “Contractor Proprietary Property”). The OSPI agrees that the Contractor Proprietary Property were not developed under the Contract and, notwithstanding anything in the Contract to the contrary, all copyrights, patent rights, trade secret and all other intellectual property rights in the Contractor Proprietary Property shall be and remain Contractor’s sole property and no license is granted to the OSPI relating thereto including, without limitation, the right to modify or remove restrictive markings on any Contractor Proprietary Property. The OSPI agrees to maintain the security and confidentiality of the Contractor Proprietary Property it may have access to and agrees not release the same to any third party without the prior written consent of Contractor. Similarly, in the event that the OSPI desires Contractor to use stimuli, the copyright in which is held by a third party (the "Third Party Materials"), all rights in and to such Third Party Materials are and shall remain the sole property of the third party rights holder. The Contractor will obtain copyright permission to use the Third Party Materials for the purposes of the Contract during the term of the Contract and any renewals or extensions of the Contract.”

7. Clause 18 “Indemnification” (page 37)

To clarify this provision, add: “The parties acknowledge and agree Contractor’s indemnification of
“Claims” refers to third party claims and claims by Contractor’s agents, employees, representatives, or any subcontractor, and not claims by OSPI’s agents, employees or representatives or SBAC Partners. Thus, notwithstanding anything to the contrary in this Contract, the obligations of this Section 18 shall apply only to claims that Contractor (or any of its subcontractors’) performance or failure to perform was negligent, in violation of law or in breach of this Contract, and such will be reduced or eliminated to the extent of the OSPI’s or its agents’ negligence, violation of law or breach of this Contract. When Contractor is obligated to defend, indemnify or hold the OSPI harmless, the OSPI shall provide: (i) prompt notification in writing to Contractor of any such Claim; (ii) the opportunity for CTB to conduct the defense and the exclusive right to settle the claim; (iii) full information and records concerning the claim; and (iv) reasonable cooperation in the defense of the same. The OSPI shall have the right to participate in the defense with attorneys of its choosing and at its own expense. The OSPI further acknowledges and agrees the Contractor shall have no liability for any claims resulting from any test materials previously developed by OSPI and provided to Contractor by OSPI or its agents, employees or representatives for use in the test materials under this Contract, or for their actions, inaction or conduct. In no event shall Contractor be liable for any incidental, consequential, or special damages.”

8. Exhibit D Clause 20 “Insurance” (pages 38-39), and RFP Section B.17. “Insurance Coverage” (page 13)
   The third sentence of Clause 20.a. “Worker’s Compensation Coverage” is modified as follows: “Such worker’s compensation and occupational disease requirements shall include coverage for all employees of the Contractor suffering bodily injury (including death) by accident or disease, which arises out of or in connection with the performance of this Contract. Contractor will be responsible for ensuring that any subcontractors Contractor uses for this project will comply with this requirement separately.”

9. Exhibit D Clause 25 “Payments” (page 40); Exhibit D Clause 38 “Termination Procedure” (page 42)
   To clarify this provision, add the following to the end of Clause 25: “The State will not withhold its acceptance, certification or voucher for services performed in accordance with this Contract.”
   The first sentence of the second paragraph of Clause 38, is modified as follows: "The Superintendent shall pay to the Contractor the agreed upon price, if separately stated, for completed work and services performed in accordance with this Contract, and reasonable compensation agreed upon by the Contractor and the Superintendent for (a) completed work and services for which no separate price is stated, (b) partially completed work and services, (c) other property, and (d) the protection and preservation of the property, unless the termination is for default by Contractor, in which case the Superintendent shall determine the extent of the liability.” The last sentence of the same second paragraph will be replaced by: “The Superintendent may withhold from any amounts due to the Contractor the compensation for the disputed services until such dispute is resolved.”
   Subparagraph 38.d. shall be modified as follows: “Endeavor to settle all outstanding liabilities and all claims arising out of such termination of orders and subcontracts, with the approval or ratification of the Superintendent to the extent the Superintendent may require.”

10. Exhibit D Clause 28 “Records Maintenance” (page 40)
    Add to this provision: “No competitor of CTB will perform any auditing that will require release of any CTB confidential, trade secret or proprietary information.” Additionally, since CTB’s site may contain other CTB customers’ confidential information, any third party auditor acting on the OSPI’s behalf shall sign a non-disclosure agreement.”

11. Exhibit D Clause 36 “Termination for Default” (page 41)
In order to clarify this provision, add to it: “The parties acknowledge that termination herein refers exclusively to termination as a result of Contractor’s breach of a material provision of this Contract, and Contractor is granted a reasonable period of time to remedy the claimed breach. In the event Contractor’s failure to perform or delay in performing is caused by a force majeure event, the OSPI agrees that Contractor will not be assessed penalties or damages, and no payments will be withheld. A force majeure event means a circumstance or event beyond Contractor’s reasonable control or Contractor’s delay or inability to perform is due to acts or failures to act on the part of the OSPI or another OSPI contractor, agent, employee or representative. In no event shall Contractor be liable for any special, consequential or incidental damages.”

12. Exhibit D Clause 37 “Termination Due to Funding Limitations” (page 41-42)

Add to this provision: "The OSPI shall notify the Contractor in writing when such termination due to funding limitations is likely to occur and Contractor shall be paid for all services rendered and materials furnished by Contractor (including work commenced by Contractor or its subcontractors) up to the effective date of termination."

13. RFP Section A.4 “Objective and Scope of Work” (page 7); and Section A.5 “State’s Role” (page 7-8)

CTB understands there will be significant collaboration among Contractor, OSPI, SBAC, SBAC leadership, and other OSPI or SBAC contractors, and that OSPI will be responsible for managing all contracting activities with the Contractor.
Introduction

The Smarter Balanced Assessment Consortium (SBAC) recognizes the importance of clearly delineating performance continua as a foundation for its summative assessments. Strong Achievement Level Descriptors (ALDs) are fundamental to quality item and task development, test development, and standard setting—all essential for producing valid and reliable assessments.

SBAC’s Request for Proposals (RFP) # 12: Initial Achievement Level Descriptor Development calls for a contractor to develop general and grade-specific Achievement Level Descriptors in English/Language Arts (ELA) and Mathematics in grades 3–8 and grade 11.

CTB/McGraw-Hill (CTB) and College Board are pleased to collaborate in response to this RFP, offering the Smarter Balanced a state-of-the-art solution and a highly respected team to implement the well-researched methodology. CTB and College Board bring to the Consortium more than 185 years combined experience in assessment and a proven record of success in work similar to the scope of this RFP.

Both CTB and College Board have extensive experience in the development of ALDs, with a clear understanding that well-defined ALDs provide a sound foundation for a standards and assessment system, as well as serving as a valuable tool in the development of test blueprints, test items, and the reporting of results.

Both organizations have in-depth knowledge of Common Core State Standards (CCSS) and work daily with these real-world skill definitions in their assessment development with districts and states nationwide. In addition, College Board has data from more than seven million students, which could be used to assure linkage of the initial ALDs to the CCSS.

The research team dedicated to this program includes leaders in evidence-centered design, and will incorporate strategies to ensure that the ALDs will serviceably inform test design and item development and will support effective standard setting and intended test score uses.

The program team brings to this project valuable content expertise in ELA and Mathematics, based on years of developing curriculum plans, test blueprints, and items and performance tasks for customized state assessments and standardized assessments that serve thousands of districts nationwide. The team also offers practical experience in shaping ALDs with the needs of all stakeholders in mind.

The team is well versed in higher education readiness, with specific experience in the knowledge and skills required for students to be prepared for entry-level college courses. College Board has an explicit commitment to college readiness initiatives, including promoting curricula, assessment tools, and guidance resources that help students prepare for the academic rigors of higher education.

CTB is nationally recognized for innovations in psychometric research, including all aspects of the standard setting process. CTB’s work in the development of valid ALDs for standard setting has led many states to Full Approval without Recommendation Status with the Federal Peer Review Process as required by United States Department of Education.

CTB and College Board understand that the ALDs required under this RFP are initial ALDs, created as a foundation for final ALDs, which will be developed under the guidance of the Smarter Balanced standard setting contractor.
We offer the experience needed to deliver ALDs to meet Consortium’s needs and to ensure that the program is efficiently and effectively managed to achieve all established goals.

In writing this proposal, CTB assumed that the review cycles must take place consecutively and 10 days were needed prior to each cycle for panelists to review the ALDs. Working backwards from the stated end-contract date through all review cycles (including the TAC review), CTB came up with the date for the workshop. The answers to the Questions and Answers released on April 9, 2012 show that this assumption may be relaxed.

With relaxed assumptions, our timelines will change. The dates of the workshop will likely change to an August date. This change gives us approximately two months to recruit panelists. CTB has a long tradition of recruiting panelists for workshops, including our own standard setting workshops for our internal products. These panelists are selected from across the United States. We will bring this experience to our work in recruiting panelists for the Smarter Balanced ALD-development workshop. Given the change to assumptions, we recommend that each governing state nominate panelists who will later be selected by the Consortium and CTB working together to identify the strongest potential panelists. To select panelists from higher education, governing states can again nominate panelists, or we could review the relevant literature (e.g., achievement gains within mathematics) to identify potential members.

**Well-Researched Methodology**

CTB and College Board are partnering on this program to provide a high quality solution for the deliverables in SBAC-12. The proposed solution includes the following:

- The creation of initial ALDs through the ALD Framework proposed by Egan, Schneider, and Ferrara (2012). This framework creates an interrelated set of ALDs whose breadth and depth vary depending on their use. For SBAC-12, the most important deliverables will be the policy ALDs (to guide ALD development), range ALDs (to guide item writing), and target ALDs (to guide standard setting).

- The development of methodology for defining policy ALDs. CTB and College Board propose to work collaboratively with the Smarter Balanced-appointed group to reach agreement on these ALDs. This step is critical to the success of the project, as these ALDs must be in place before the rest of the work can be implemented.

- The determination of the methodology for developing range and target ALDs. This phase will include the following activities:
  - Develop panelist selection strategy with SBAC approval
  - Annotate the common core content standards; annotate the SBAC content specifications which contain their claims and evidence
  - Ensure procedural validity of the method
  - Survey the panelists

- The creation of Validity Studies that will allow for linking high school ALDs to SAT or AP data, studying placement of items in relationship to the ALDs, surveying item writers and teachers about the usefulness of the ALDs.

- The development of an initial and final Technical Report, which will document the processes used and products developed to support the validity of the initial ALDs.

**Proven Leadership**

The expertise and experience of CTB and the College Board are well documented in the assessment community. Both organizations have delivered valid and reliable assessments for large-scale statewide programs as well as for thousands of district programs, providing useful information to improve student performance and inform instruction. A proven record such as this is important; however, overall program management is also integral to the success of the project.
The Smarter Balanced Assessment Consortium recognizes the need for a contractor with the capability to respond to the Consortium’s needs in a fast-paced environment with multiple tasks to be completed concurrently. The contractor must support recruitment of committee members, facilitate multiple meetings simultaneously, and work collaboratively with a variety of groups and additional contractors as needed. CTB will provide this high level of program management and work closely with the College Board staff and Smarter Balanced staff, committees, and contractors to ensure that multiple tasks are managed efficiently and deliverables are provided within the agreed upon timeline.

Ms. Lindy Wienand will be SBAC-12 program manager. Ms. Wienand has eight years of program management experience and has successfully managed a number of statewide assessment programs, including Department of Defense Education, Ohio, Alaska, Arizona, and Hawaii.

As a Program Manager, Ms. Wienand is responsible for the management and oversight of all program activities including schedule creation, management of cross-functional staff, risk identification and mitigation, decision documentation and record keeping, monitoring project implementation progress, establishing priorities, acting as a team leader for representatives of all departments involved, and interfacing with and ensuring the satisfaction of the Smarter Balanced Assessment Consortium.

CTB’s experience in program management, open communication, and integrated planning/documentation is nationally recognized as one of the most customer focused and cost efficient approaches offered today.

**Collaboration for Success**

As assessment and learning systems evolve to meet the changing needs of education, one thing remains constant—the need for quality assessment solutions that present reliable and accurate date about student achievement. Such solutions begin with careful attention to the foundation, especially Achievement Level Descriptors that distinguish knowledge and skills needed and the level of performance required of students. CTB and the College Board stand ready to dedicate an experienced and expert team to complete this project and meet the highest quality standards. We consider each program a collaborative initiative to address specific needs in the most efficient and effective way possible. The Smarter Balanced leadership, member states, education stakeholders, and other contractors working with the Consortium are all part of this dynamic process.

We succeed only if our products and services help teachers and students advance in every stage of their educational processes. The deliverables for this phase of SBAC’s program provide the foundation for a rigorous instructional and assessment system that will assist teachers in improving instructional delivery for all students. The successful implementation of this program will also support the Consortium’s vision of the next generation of assessments and the related work currently in progress in Smarter Balanced member states—all ensuring sustainable success for the future.
OBJECTIVE AND SCOPE OF WORK

2.1 Submit a draft schedule of tasks and activities as a part of the bid. Our management team will develop and maintain comprehensive project documentation and communications plans that assure the transparency and oversight supports required for the delivery of all program deliverables. The primary forms of program documentation/communication include (each of these have been explained in greater detail in the Management Plan of this proposal):

- Program Work Plan - Top-level program planning document providing specifics relative to program scope and defined roles/responsibility.
- Program Master Schedule - Organized in a Work Breakdown Structure, the master schedule will provide the identification, organization, and sequencing of all program tasks, deliverables, and milestones.
- Deliverable Matrix - Definition of all major deliverables and assigned resources and delivery dates (including all dates for review cycles and approval).
- Status Reports - weekly management report on project status, key risks, upcoming events, critical required actions.
- Record of Decisions - Catalogued record for key decisions made during program activities.

CTB will comply with all existing documentation formats as provided by the Smarter Balanced Assessment Consortium in order to maintain consistency across Smarter Balanced projects and vendors.

CTB has prepared a schedule of the tasks and activities required to successfully deliver the SBAC-12 project scope of work. A draft of the Master Project Schedule is provided in Appendix A.

2.2 Submit a final detailed schedule of tasks and activities within one (1) week of contract award as described above.

CTB will maintain full responsibility for complete program oversight, including the maintenance of timelines and critical program deliverables, complete program communication across the College Board and Smarter Balanced leadership, and consolidation of all work completed on the contract. As detailed in our Management Proposal, we have assigned a senior-level management team with the skills and experience necessary to maintain program oversight and effectively coordinate all program activities.

The complexity and accelerated timeline of the program will require the team to build consensus quickly and move decisively in order to maintain the measured speed required to develop, review, revise and deliver the initial achievement level descriptors within the required timelines. It will require a “one-department” mindset to develop quickly across Smarter Balanced and our alliance organization, the College Board, in order to deliver all that is required.

The team, under the leadership of Program Manager Lindy Wienand will maintain a full program schedule and complete program documentation. A Master Project Schedule will be created to assure
identification, organization, and sequencing of all project tasks, deliverables, and milestones. The project schedule will take the key elements of the project and translate them into a time-based plan. The complete schedule will include a work breakdown structure, all tasks and activities associated with the project, and the interdependencies of the tasks to be performed. The Project Schedule will be created using CTB’s project management scheduling software, and will be continuously monitored, updated and analyzed by Kristal Taylor, the Program Schedule Analyst.

The project schedule will be reviewed and negotiated at the kick-off meeting with Consortium leadership within one week of the contract award. The schedule will be revised and finalized immediately after this meeting. Then the final schedule will be submitted to the Consortium and will become one of the primary tools used to manage the project to successful completion.

The team will provide comprehensive program documentation using a number of management tools, including a Program Work Plan. This top-level program-planning document provides the details of the program scope and defined roles and responsibilities of all participants to assure everyone is able to be productive and effective right from the start. The Program Work Plan provides the answers to the who, what, when, and how, questions related to key activities, milestones, deliverables, timeline, resources, risks, program controls, and quality controls. This document will become the central control document for all teams working on the program, and the basis for change management. The work plans for each program part align to the overarching Program Work Plan, and will define at a greater level of detail the tasks, services, and activities to accomplish the scope of each part.

2.3 Provide weekly status reports to the CONSORTIUM’s Project Management Partner regarding the progress of the project.

Weekly Status Reports – CTB will provide a detailed weekly management report to the Consortium. The weekly management report will address how the project is progressing, in comparison with the schedule. The weekly status report will also address issues and risks that have been identified, and how they are being managed. The weekly status reports are intended to provide the necessary communication to all Smarter Balanced stakeholders groups about the progression of the project work.

2.4 Schedule and facilitate a contractor kick-off meeting upon execution of the contract.

Program Management will maintain fiscal and organizational responsibility for the kick-off meeting. Logistics information will be available through established collaboration channels. Key fiscal and meeting summary information will be provided through weekly management reports.

CTB will schedule a kick-off meeting immediately upon execution of the contract. Because tasks associated with the initial achievement level descriptors project need to commence as soon as possible, we propose that the kick-off meeting be conducted via WebEx webinar. If necessary, we will follow up with a formal face-to-face meeting in conjunction with the Achievement Level Descriptors Development Committee Meeting.

2.5 Project the initial ALDs, when complete, in print ready format as well as digital formats that can be incorporated into a variety of databases.

CTB will provide final copies of the ALD documents in both print and digital formats. As we are drafting and reviewing the initial ALDs, we will work in a format that is readily accessible by a variety of stakeholders and one in which comments are easily captured. We will work with SBAC to determine the appropriate formats. One recommended format would be Adobe PDF files with comments captured in an online form or survey. Final document delivery will be in print-ready PDF format along with the source Word document files. A proposed Table of Contents for the final deliverable is shown in Table 1.
Table 1. Proposed organization of final ALDs

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Introduction to ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>I.a</td>
<td>Grade 3 ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>I.b</td>
<td>Grade 4 ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>I.c</td>
<td>Grade 5 ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>I.d</td>
<td>Grade 6 ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>I.e</td>
<td>Grade 7 ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>I.f</td>
<td>Grade 8 ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>I.g</td>
<td>High School ELA Achievement Level Descriptors</td>
</tr>
<tr>
<td>II</td>
<td>Introduction to Mathematics Achievement Level Descriptors</td>
</tr>
<tr>
<td>II.a</td>
<td>Grade 3 Mathematics Achievement Level Descriptors</td>
</tr>
<tr>
<td>II.b</td>
<td>Grade 4 Mathematics Achievement Level Descriptors</td>
</tr>
<tr>
<td>II.c</td>
<td>Grade 5 Mathematics Achievement Level Descriptors</td>
</tr>
<tr>
<td>II.d</td>
<td>Grade 6 Mathematics Achievement Level Descriptors</td>
</tr>
<tr>
<td>II.e</td>
<td>Grade 7 Mathematics Achievement Level Descriptors</td>
</tr>
<tr>
<td>II.f</td>
<td>Grade 8 Mathematics Achievement Level Descriptors</td>
</tr>
<tr>
<td>II.g</td>
<td>High School Mathematics Achievement Level Descriptors</td>
</tr>
</tbody>
</table>

2.6 Solicit volunteers and select, in collaboration with and final approval of CONSORTIUM leadership, content and/or curriculum specialists in ELA and mathematics to participate in the development of initial ALDs, as described more fully above.

Recruiting panelists for the purpose of developing ALDs is a serious undertaking. It is important that the diversity of stakeholder interest is represented when ALDs are written. In this section, CTB presents a methodology systematically recruiting content and/or curriculum specialists that takes into account the varied stakeholder groups that are part of the Smarter Balanced Assessment Consortium.

In developing this methodology, we are particularly concerned with meeting Section 2.6 of the Peer Review guidance, which requires that the sponsoring agency be able to describe the selection process and to involve a diverse group of stakeholders. To meet this requirement, it is important that the methodology for panelist selection be well planned so that a wide variety of stakeholders will be represented at the workshop. In this section, we will present:

- a methodology for panelist selection, and
- an overview of how the panelists will be placed into groups.
Methodology

Within the RFP, Smarter Balanced has provided a basic sampling design showing from where panelists should be recruited. CTB suggests that CTB and Consortium representatives work together to refine this design to include other relevant characteristics, like demographic information or teaching experience. In the interest of time, CTB suggests that the pool of nominees come from prior recruiting work that the Smarter Balanced Assessment Consortium has done. CTB can quickly examine the nominee information from previous pools to assure that all governing states are included. If available, we can also check the qualifications of the nominees. Qualified, but unselected, nominees from prior recruiting work can be contacted to ascertain their interest in participating in the ALD development. If a governing state does not have nominees in the pool, then CTB will work with Consortium representatives to identify five potential nominees from their governing states. If the Consortium prefers, CTB can work with Consortium representatives to create a new pool of nominees for the purposes of creating initial ALDs; however, this may add more time to the schedule.

Once a pool of potential nominees has been selected, CTB will create a matrix that shows the grade-level experience, content-area experience, and Common Core State Standards (CCSS) experience of each nominee. CTB will work with Smarter Balanced to include other desired information in the matrix, which might include demographic information or types of teaching experience. We will use this information to find the most qualified nominee from each governing state. A list of nominees will be created by CTB and will be approved by the Consortium. Once approved by Smarter Balanced, the nominee will be contacted to participate in the ALD-development workshop (see Section 2.7).

In writing this proposal, CTB assumed that the review cycles must take place consecutively and 10 days were needed prior to each cycle for panelists to review the ALDs. Working backwards from the stated end-contract date through all review cycles (including the TAC review), CTB came up with the date for the workshop. The answers to the Questions and Answers released on April 9, 2012 show that this assumption may be relaxed.

With relaxed assumptions, our timelines will change. The dates of the workshop will likely change to an August date. This change gives us approximately two months to recruit panelists. CTB has a long tradition of recruiting panelists for workshops, including our own standard setting workshops for our internal products. These panelists are selected from across the United States. We will bring this experience to our work in recruiting panelists for the Smarter Balanced ALD-development workshop. Given the change to assumptions, we recommend that each governing state nominate panelists who will later be selected by the Consortium and CTB working together to identify the strongest potential panelists. To select panelists from higher education, governing states can again nominate panelists, or we could review the relevant literature (e.g., achievement gains within mathematics) to identify potential members.

Groupings

The panelists will work in small groups that span two grades, except for grade 11. The proposed grade groupings will be:

- Grades 3-4
- Grades 5-6
- Grades 7-8
- Grade 11

Within each content area, each grade span will have four panelists, except for Grade 7-8 which will have five panelists.

Below, Table 2 shows the disaggregation of panelists into each grade and grade-span group. As shown in
Table 2, one grade-level expert from a governing state will be in his/her appropriate grade; one content expert will in each grade span group; and one expert in CCSS will be in each grade-span group. A representative of higher education will participate in the grade 7-8 group and the grade 11 group.

Table 2. Panelists by Stakeholder Group & Committee Role

<table>
<thead>
<tr>
<th>Stakeholder group/committee role</th>
<th>Grade level/span</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Grade-level experts from governing states</td>
<td>1</td>
</tr>
<tr>
<td>Content experts from governing states</td>
<td>1</td>
</tr>
<tr>
<td>CCSS content experts</td>
<td>1</td>
</tr>
<tr>
<td>IHE rep</td>
<td>1</td>
</tr>
<tr>
<td>Total grade band</td>
<td>4</td>
</tr>
<tr>
<td>Total governing</td>
<td>11</td>
</tr>
<tr>
<td>Total other</td>
<td>6</td>
</tr>
<tr>
<td>Total panelist per Content area</td>
<td>17</td>
</tr>
</tbody>
</table>

2.7 Develop the methodology for designing ALDs, which will create descriptors in accordance with the requirements listed previously.

Achievement level descriptors have historically been used to support score interpretation and in relationship to standard setting, either to guide the standard setting or to summarize the KSAs expected by the standard setting panelists once cut scores are set. With the advent of ECD, ALDs have become central to test design and to item writing. As such, modern ALDs are more comprehensive and structured than are their predecessors. In a comprehensive review of the existing ALD literature, CTB Research Scientists uncovered four distinct yet interrelated uses for ALDs, including guidance for policy, test and item development, standard setting, and score interpretation (Egan, Schneider, & Ferrara, 2012). No single ALD can do all of these things well. A single multi-use ALD tends to conflate uses and confuse users. In response to this, CTB Researchers created a framework that comprises four linked, interrelated ALDs that are directly related to the uses explicated above.

Smarter Balanced has indicated that its ALDs will be used for test and item development, standard setting guidance, and score interpretation. Thus, there is a good fit between the way in which the Consortium wishes to use its ALDs and the framework that CTB would use to develop those ALDs. This framework will allow Smarter Balanced to create ALDs that are clear, concise, and appropriate for its intended use. In this section, we first discuss a framework for developing ALDs that are appropriate for the multiple uses Smarter Balanced desires of its ALDs. Next, we propose a workshop where Consortium stakeholders can take advantage of the framework to create initial ALDs.
A Framework for Developing Achievement Level Descriptors

In current psychometric practice, ALDs have four interrelated uses, including policy guidance, standard setting guidance, test and item development, and score interpretation (Egan, Schneider, & Ferrara, 2012). While interrelated, each of these uses intersects with ALDs in different ways. For example, item writers will need ALDs that have knowledge, skills, and abilities (KSAs) disaggregated by cluster and strand of the content standards in order to guide their task while a parent may want a summative description of the KSAs possessed by his/her child. CTB Research Scientists recognized that a single, multi-use ALD could not fit all of these needs. In response, we combined our operational experience and theoretical knowledge to create an ALD development framework that comprises four linked interrelated ALDs that are directly related to the uses explicated above.

The framework includes:

1. Policy ALDs that are general descriptors that name the performance levels that will be used on the assessment, and articulate the goals and rigor for the final performance standards. These descriptors set the tone for and are embedded within subsequent descriptors.

2. Range ALDs that are grade- and content-specific ALDs that may be used by test developers to identify which aspects of items align to a particular performance level in regard to the cognitive and content rigor that has been defined.

3. Target ALDs that are created in conjunction with or following the Range ALDs and are used to guide the standard setting. They use only the information from the Range ALD that defines the minimum performance required to be considered as meeting the performance level expectation.

4. Reporting ALDs are the final ALDs that are developed following the standard setting once cut scores are finalized, and they define the appropriate intended interpretations regarding what a test score means, because they are developed to be consistent with the scores and reporting information to be produced by summative computer adaptive and performance tasks that SBAC will implement.

This system of four linked interrelated ALDs explicates the construct being measured and supports the intended test score interpretations in regard to that construct. In other words, the framework supports an evidence-centered design approach. Although the framework ALDs are discussed as four interrelated ALDs, they also can be conceptualized as the evolution that an ALD must undergo from its inception to completion because the ALD development process is iterative (Plake, Huff, & Reshetar, 2010). These ALDs define the construct that is being measured and describe what students should know and be able to do in relation to the construct student achievement relative to the common core standards.

---

1 For an in-depth discussion of the ALD framework, please see Egan, Schneider, & Ferrara (2012) attached in Appendix B.
Figure 1 shows how the four ALDs are linked and inform the development of the next ALD type. As shown, the general ALDs are overarching statements that are policy claims that encompass all grade content areas. They are the educational goal regarding what students within the performance level are expected to do. From these policy claims grow the range ALDs that incorporate the grade- and content-specific information about the knowledge, skills, and abilities (KSAs) along the proficiency continuum that students are expected to demonstrate. Range ALDS describe the types of evidence items within the performance level should elicit to support the policy claims, and it is for these reasons that they support item writing. Range ALDs are built using the common core standards, Smarter Balanced content specifications, and the expected learning trajectory the standard implies within a specific grade and content area. The target ALDs are the preliminary conceptualization of the minimum evidence a student should demonstrate from the range ALDs to be considered just proficient. The reporting ALDs are the final indication of the target ALDs based upon the final approved cut score.

CTB understands that final ALDs that are consistent with score interpretation, information reported from the adaptive summative tests and performance tasks, and learning gains achievable through reasonable effective instruction, will be developed following the standard setting workshop. We discuss this intended ALD use, along with the supporting ALDs, to provide an orientation to the ALD framework. The scope of the current proposal is primarily associated with the policy, range, and target ALDs within the framework.

**Initial ALDs and the ALD Framework**
The Consortium has stated the three purposes for its initial ALDs, including (a) item and test development, (b) standard setting guidance, and (c) score interpretation. As such, SBAC’s goal and CTB’s ALD framework are well aligned. In addition, SBAC has indicated that the initial ALDs comprise general descriptors and grade-specific ALDs for each content area. Within the context of the ALD framework, general ALDs are equivalent to policy ALDs, and grade-specific ALDs encompass range and target ALDs.

**Utility of the Framework**
Some may be concerned that the use of a system of interconnected ALDs creates the “potential for misuse or confusion” (Camara & Quenemoen, 2012); however, it is the single multi-use ALD that has directly led to confusion in the past. For example, much of the debate surrounding ALDs focused on the use of “should” statements versus “can” statements. This debate is resolved when we recognize that “should” statements provide guidance, and standard setters to operationalize ALDs can use this type of guidance. The “can” statements describe KSAs held by students, and such statements are generally useful to parents or teachers so they can understand what knowledge the student demonstrated based on the assessment. By disaggregating the types of ALDs by their uses, we decrease the potential for misuse of the ALDs and are able to focus ALD development that is appropriate for its use. The use of the different ALD types allows us to use the right tools for the right job.

---

2 Target ALDs should reflect the standard setting methodology to be used. For the purposes of this discussion, we have assumed that the borderline student is appropriate for the standard setting methodology. If SBAC prefers, Target ALDs may be created for typical students (or students at the middle of the range).
Figure 1. Framework for ALD Development

Step 1. Develop General Policy Achievement Level Descriptor and Define College Readiness

The first step in creating the interconnected system of ALDs will be to construct the general policy ALDs. These ALDs will provide guidance for the remaining work. SBAC has indicated that it desires that general ALDs be developed for each content area. CTB suggests that the work for developing the general ALDs begins with the overall claims for the summative assessments. The overall claims are a sort of policy guidance that the expectations that Smarter Balanced has for its students. Using the overall claims along with the claims associated with particular aspects of the common core state standards, the panelists can discuss the expectations of the students in each of the four achievement levels in terms of these claims.

Table 3 shows overall grade 11 claim for the summative Mathematics assessment. Much like the policy ALDs from the framework, these claims show expectations that Smarter Balanced has set forth for students. The panelists can begin with these claims and show the progress of these claims for students in Levels 1, 2, 3, and 4. Using just the overall claim as an example, the panelists may begin by stating that grade 11, Level 3, students demonstrate CCR in mathematics. From here, they will move up the proficiency continuum to discuss the general claim for the Level 4 students. They will then complete their work by discussing their expectations for the Level 1 and Level 2 student.
**Table 3. Example of how SBAC Claims may be Scaffolded into Achievement Levels**

<table>
<thead>
<tr>
<th>Overall Claim for Grade 11</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Claim for Grade 11</td>
<td>Level 1 students require intensive support to develop basic math skills. They require substantial remediation in skills considered prerequisite to college mathematics.</td>
<td>Level 2 students require remedial coursework in mathematics and are currently unprepared for entry-level college mathematics.</td>
<td>Level 3 students demonstrate college and career readiness in mathematics. These students are prepared for entry-level college mathematics without the need for remediation.</td>
<td>Level 4 students demonstrate advanced CCR in mathematics. These students are prepared for more advanced mathematics coursework that might be expected of students after their first year of college mathematics.</td>
</tr>
</tbody>
</table>

**Step 2. Develop Grade- and Content-Specific Achievement Level Descriptors – Range and Target ALDs**

In the context of the ALD framework, content-specific ALDs for each grade are the range and target ALDs. CTB has successfully used the method described to create Target PLDs to guide standard setting, and it is easily understood and utilized by panelists. This method has been extended so that both Range and Target ALDs can be developed.

The process of writing the first draft of the range and target ALDs can be overwhelming for panelists. Through our experience with standard setting, we have found that annotating the content standards provides the panelists with a doable starting point. In doing this, panelists can see how the work can be parsed and accomplished so that KSAs can be compiled into achievement levels. In the proposed process, panelists will study the general ALDs, parse CCSS and content specifications, and compile the KSAs into achievement levels.

**Study the General Policy Descriptors.** Writing range and target ALDs is most efficiently accomplished in a workshop (that will be described next) with small groups of panelists. To develop range and target ALDs, panelists should begin studying the general policy ALDs to understand the claims that will be made using the assessments. In writing the range and target ALDs the panelists operationalize intent expressed in the general policy ALDs.

**Parse the CCSS and Content Specifications.** Next, the panelists will study and begin to annotate the content standards and the content specifications. The panelists will first analyze the cognitive expectations, content information, and skills found in the CCSS. Figure 2 shows an example of this process using a cluster from the Ratios and Proportional Relationship domain of the seventh-grade mathematics CCSS. For the purposes of this description, we use the achievement level labels “Proficient” and “Basic.”
Figure 2. Parsed CCSS to Create Range and Target ALDs

KEY:

P-: Skills of the student just entering proficiency
P: Skills of the average Proficient student
P+: Skills of the high-performing Proficient student
B-: Just entering Basic
B: Average Basic
B+: High-performing Basic

Ratios and Proportional Relationships 7.RP

Analyze proportional relationships and use them to solve real-world and mathematical problems.

1. Compute unit rates associated with ratios of fractions, including ratios of lengths, areas and other quantities measured in like or different units. For example, if a person walks $\frac{1}{2}$ mile in each $\frac{1}{4}$ hour, compute the unit rate as the complex fraction $\frac{1}{2}/\frac{1}{4}$ miles per hour, equivalently 2 miles per hour.

2. Recognize and represent proportional relationships between quantities.

   a. Decide whether two quantities are in a proportional relationship, e.g., by testing for equivalent ratios in a table or graphing on a coordinate plane and observing whether the graph is a straight line through the origin.

Skills represented by P- are compiled for the Target Proficient PLD. The Target Proficient student is assumed to possess the KSPs assigned to the Basic category.

After identifying the range of cognitive complexity within the content standard and strands, the panelists should annotate the content standards, indicating the performance level to which each skill aligns crossed with the cognitive level. The general policy ALDs provides broad guidance for analyzing the content and skills in the content standards. Within each performance level, the skill can be further delineated into just Proficient (P-), average Proficient (P), or highly Proficient (P+) using the annotation process shown in Figure 2. A similar delineation should be made for the other performance levels. For example, Figure 1 shows panelists have indicated that recognizing proportional relationships is a skill expected of the just Basic examinee (B-), whereas representing proportional relationships is a skill expected of the average Basic examinee (B).

Compile the KSAs into ALDs. As this analysis is performed, the content and skills can be compiled into a database showing their alignment to each performance level. Then, a matrix can be created showing the delineation of KSAs into range ALDs, which include all of the content and skills assigned to a performance level. The range ALDs should eventually resemble the matrix presented in Table 4. A
A separate matrix can be compiled for those KSAs classified in the just Proficient (P-) category. These KSAs constitute the target ALD for Proficient. The KSAs of both the target and range ALDs will be examined for fluency across the grade levels within a content area. The KSAs will be compiled into a single set of ALDs for each grade level and content area.

Because Range ALDs will guide test development, they should deconstruct what evidence of achievement for each standard found on the assessment looks like at each performance level. Table 4 shows an example of a portion of a Range ALD for a Grade 8 Mathematics test. As demonstrated in Table 4, this Range ALD scaffolds KSAs across the ability range of each performance level and in some cases connotes when straight computation is expected versus an increase in cognitive complexity due to context. For example, a Level 1 student can determine perimeter of a regular figure without context being present. The Level 2 student is expected to show how to apply this skill in a real world situation. Notice that the descriptor for Level 3 does not distinguish the lowest Level 3 student from the mid-level and highest Level 3 student. In some cases, portions of the matrix may be incomplete if panelists believe that students in a level cannot do any portion of the work or if the panelists do not have enough information to predict what students in the next level should be able to do.

**Table 4. Example of Portion of Range ALD for Grade 8 Mathematics**

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determine perimeter of regular figure when no context is present</td>
<td>Determine perimeter of regular figure when context is present</td>
<td>Given perimeter, determine the missing measurement of a regular figure when context is present</td>
<td>Given perimeter, determine the missing measurement of an irregular figure when context is present</td>
</tr>
<tr>
<td>Find area of regular figure</td>
<td>Find area of irregular figure</td>
<td>Find missing measurement of regular figure given area when context is present</td>
<td>Find missing measurement of irregular figure given area when context is present</td>
</tr>
<tr>
<td>Find volume of regular figure</td>
<td>Find volume of irregular figure</td>
<td>Find missing measurement of regular figure given volume when context is present</td>
<td>Find missing measurement of irregular figure given volume when context is present</td>
</tr>
<tr>
<td>Use data in a circle graph to determine if statements are true</td>
<td>Use data in a single bar graph to determine if statements are true or false</td>
<td>Use data in a line graph to determine if statements are true</td>
<td></td>
</tr>
<tr>
<td>Determine the mode of a set of data</td>
<td>Determine the range of a set of data</td>
<td></td>
<td>Determine the mean and median of a set of five 2-digit numbers presented in a table of data</td>
</tr>
</tbody>
</table>

*From Egan, Schneider, & Ferrara, 2012

**Challenges to Creating Initial ALDs**

In developing a methodology to create initial ALDs, it is necessary to address challenges that may arise as we go through the process. Here, we address five challenges.
Creating General Policy ALDs
General policy ALDs are typically developed and approved by policy makers ahead of any committee work. The timelines of this work necessitate that the general policy ALDs are created in tandem with the grade-specific range and target ALDs. This means that the general policy ALDs will be reviewed at the same time as the grade-specific ALDs. If changes are made to the general policy ALDs, then it will likely affect the grade-specific ALDs. If SBAC can allow more time, then we recommend that policy ALDs be created and approved prior to writing grade-specific ALDs.

College and Career Readiness
The general claim of “college and career readiness” may present a challenge to panelists as they begin to parse out the claims for each achievement level. The meaning of CCR as defined by Smarter Balanced Assessment Consortium policymakers will provide important guidance to panelists as they create the grade-specific, content-based ALDS. CTB recommends that SBAC define its policy meaning of CCR prior to convening the content-panels. For purposes of developing the College and Career Ready ALDs for the SBAC assessments, we are assuming that the KSAs for college and the KSAs for career readiness are largely the same as they apply to the common core state standards, and that separate ALDs will not be developed for college and career readiness. However, we recognize that as more is learned about differences between college and career readiness, refinements may need to be made to better reflect these differences. In a similar fashion, one could argue that college readiness is not a single monolithic construct. For example, being college ready for a STEM field likely requires a different set of capabilities than being college ready for a fine arts field.

If Smarter Balanced desires that the panelists define CCR during the ALD-development process, then the College Board is uniquely positioned to assist the Consortium in defining CCR with data prior to the ALD development process. For example, College Board may provide external data and to conduct empirical analyses to aid in establishing ALDs. The College Board has the capacity to match state K-12 student records to College Board records, thereby merging students’ state assessment scores with any available scores on national tests such as the SAT, PSAT/NMSQT, and AP. The College Board currently works with several states and districts in linking College Board data with local data. SAT data trends and reports are routinely published by the College Board in efforts to inform education. These types of data and information can be presented to panelists to help inform and guide the development of achievement level descriptors that are appropriate for the population of students they are meant to support. These matched data can be used to provide evidence supporting the initial ALDs in two ways.

First, because the SAT and PSAT/NMSQT have college-ready benchmarks, estimates can be made of the percent of students in a state that are college ready. These estimates are derived through a multistep process, where, through the use of common students, the College Board benchmarks are mapped onto a given state’s High School test. The percent of students exceeding that benchmark can then be calculated. If this percentage is consistent with a priori expectations of the number of students anticipated to be college ready, it may be useful to use the definition of SAT college-ready benchmarks as starting points for the level 3 achievement level descriptor.

Additionally the College Board has undertaken a scale anchoring procedure, where descriptions of what students know and can do are associated with score ranges on the SAT and PSAT/NMSQT. To the extent that the scale anchoring statements are consistent with the KSAs identified with each ALD, they can be used to provide confirming evidence of the utility of the ALDs. Alternately, the scale anchoring information could be used to help with the initial development of the ALDs.

To support the use of College Board data in this capacity, evidence suggests that assessments such as the SAT may be used as an indicator of national trends for college and career readiness. For example, the National Assessment Governing Board (NAGB) recently contracted with the College Board to obtain SAT scores for public school students who were twelfth graders in 2009 and who had taken the SAT by June of 2009. The matching process was developed to protect students’ identity and
confidentiality, which was assured through the use of pseudo-IDs for both groups of students. SAT Critical Reading scores were provided for 16,200 students and SAT Math scores were provided for 15,300 students. These data were then used by NAGB to conduct statistical linking studies. Results established a very strong relationship between the cutscore for proficient on NAEP math and the college readiness benchmark of 500 on the SAT Math test and a moderate relationship between the cutscore for proficient on NAEP reading and the college readiness benchmark of 500 on the SAT Critical Reading test. NAGB is completing additional studies on college readiness before releasing any reports publically, but these linking studies and additional alignment studies begin to establish evidence that performance on SAT may be used as an indicator of national trends for college and career readiness.

In addition to SAT and PSAT/NMSQT data, College Board can also bring to bear student performance, and the associated descriptions of that performance, on Advanced Placement exams to inform initial development of the ALDs, or to provide confirmatory evidence of the ALDs for describing college level work. A unique contribution of the Advanced Placement exams is that they are, by definition college-level work, and those students who are successful on them are generally considered college-ready.

Assessable Standards
The CCSS encompass more standards than will be included in the summative assessment. Since the RFP states that the “initial ALDs are germane to judgments made about performance levels on the summative assessments…” (p. 15), we are recommending that panelist discussion focus on the parts of the CCSS to those that have been deemed to be a target of the summative assessment in the content specifications. If SBAC would prefer that the achievement levels be based on all cluster and standards of the CCSS, then the proposed methodology can easily accommodate this.

Initial ALDs and External Benchmarks
One aspect that is currently unknown and bears reflection is the relationship between the initial ALDs that will be developed as part of the work required by this RFP and the external benchmarks that may be used during the standard setting process. If external benchmarks are to drive the standard setting methodology, then it is worth considering how these same external benchmarks may be used to inform the initial range and target ALDS.

The first step is to identify the external benchmarks that are relevant for the purposes of SBAC. For example, SBAC may choose to base its own performance standards on NAEP, the SAT, PISA, and/or TIMMS. Since these initial benchmarks will be written prior to availability of item data, it is not feasible to relate performance on SBAC items to the external benchmarks (as it will be right before standard setting). Even so, information can be gleaned from the external benchmarks that can guide the work being done to create the initial ALDs.

As an example, let us assume that Smarter Balanced wishes to use SAT performance standards of college readiness as one of its external benchmarks. For the purpose of writing initial ALDs, College Board could map its items onto its proficiency distribution, and the content experts could create a matrix of the knowledge, skills, and abilities students in each achievement level are likely to possess using the SAT cut scores of college readiness. This matrix along with sample items to exemplify the matrix would provide panelists with a set of expectations for student performance that SBAC has deemed important, and could inform Smarter Balanced panelists of how the achievement level skills found in the common core standards might best align in accordance with the SAT college readiness construct.

If no external benchmarks are used for the development of the initial ALDs yet they are nonetheless used for the purposes of standard setting, then there may be a disconnect between the initial ALDs and the standard setting process. The utility of the initial ALDs will be limited and perhaps counterproductive if they are not informed by the same external information as the standard setting process.
Learning Gains
It is desired that the ALDs that results from the ALD-development process demonstrate “consistency with learning gains achievable through reasonable effective instruction during a school year for a typical student.” This is not a simple task, and though we will direct panelists to consider learning gains during the workshop, we believe that it will be necessary to conduct a review of the literature on learning gains for ELA and mathematics. CTB proposes that the panelists representing higher education also be experts in learning gains. If this is not possible, then CTB will negotiate with SBAC to identify an external consultant with expertise in learning gains to create a short document (no more than five pages) for each content area that explains the current research on achievable learning gains. Panelists will be asked to study this document prior to coming to the ALD-Development workshop.

A Note on Accessibility
CTB is committed to providing appropriate accommodations on assessments and to increasing accessibility to assessments for all students. Accessibility is not addressed as part of the ALD development process. Appropriate accommodations increase accessibility to the test. As long as students receive appropriate accommodations, then the ALDs should be equally applicable to all students (Egan, Schneider, & Ferrara, 2011).

The ALD-Development Workshop
CTB proposes a five-day workshop in which the general policy ALDs and grade- and content-specific range and target ALDs can be developed. We first propose a workshop that will meet the time considerations laid out by the Consortium.

Table 5 overviews the activities of the five-day workshop. At a high level, a meta-committee will develop the general policy ALDs. Next, the grade 11 committees will develop the range and target ALDs for mathematics and ELA. Meta-committee members will participate in the creating the range and target ALDs for grade 11 by providing policy guidance and clarification of intent as necessary to the content-based committee. Finally, the grades 3 through 8 committees will develop range and target ALDs for mathematics and ELA for each grade level working backwards from the Grade 11 range and target ALDs.

This will be a five-day workshop for meta-committee members. For all other panelists, this will be a two-day activity. The grade 11 range and target ALDs will be completed before the grades 3-8 ALDs are written so that panelists may use the grade 11 ALDs to work backward to the KSAs that are necessary in their own grades to meet the grade 11 expectations.

In this section, we describe the committees who will meet and the proposed day-by-day activities in which they will engage. All activities are subject to the agreement and approval of SBAC and its TAC. CTB understands and expects that SBAC will engage in a discussion of the proposed activities, and we will refine the workshop based on the discussion.

Table 5. Overview of five-day workshop

<table>
<thead>
<tr>
<th>Day(s)</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Meta-Committee creates General Policy ALDs</td>
</tr>
<tr>
<td>2 - 3</td>
<td>Grade 11 creates range and target ALDs</td>
</tr>
<tr>
<td>4 - 5</td>
<td>Grades 3-8 create range and target ALDs</td>
</tr>
</tbody>
</table>
Committees
Within each content area, two types of committees will engage in the ALD-development workshops: a meta-committee and grade-span committees. Section 2.6 above details the types of panelists that CTB suggests recruiting for both types of committees.

Meta-Committee. The meta-committee will comprise one representative from each grade-span group, for a total of four members within each content area. The members of this committee will also be part of the grade-span committees within their assigned content areas. The meta-committee members will assist with their own grade-span as well as helping with work of other grade-spans when appropriate. The members of this committee will create the general policy ALDs, and they will assist with the cross-grade articulation of descriptors.

Grade-Span Committees. Within each content area, four committees will meet to create ALDs for each grade level. Table 6 shows the grade spans and the number of committee members. These committees will be responsible for creating the grade- and content-specific range and target ALDs.

Table 6. Number of Members by Committee Type for a Single Content Area

<table>
<thead>
<tr>
<th>Grade-Span Committees</th>
<th>Meta-Committee</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 – 4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>5 – 6</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>7 – 8</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Pre-Work Materials
Prior to the workshop, a packet of materials will be created that will be sent to panelists prior to the workshop. The contents of the final packet will be decided upon with input from SBAC. As an example, this packet may include:

- a workshop agenda
- samples of general policy ALDs (for meta-committee members)
- samples of range and target ALDs (for all members)
- Common Core State Standards for the appropriate content area
- SBAC content specifications for the appropriate content area

Prior to the workshop, all committee members will be asked to engage in five to seven hours of self-study of the CCSS, the SBAC content specifications, and exemplar general and grade-specific ALDs. The panelists will be sent a packet of materials that they will be asked to study prior to coming to the workshop.

Day 1 Activities
The first day of activities will consist of opening session, training, and writing policy ALDs. Table 7 shows the proposed Day 1 schedule. The final schedule would be decided upon with SBAC following the award of the contract.
### Table 7. Proposed Schedule of Day 1 Activities

<table>
<thead>
<tr>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00 – 9:00 am</td>
<td><strong>Opening Session</strong></td>
</tr>
<tr>
<td></td>
<td>• Training on general policy ALDs</td>
</tr>
<tr>
<td></td>
<td>• Orientation to materials</td>
</tr>
<tr>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td>9:00 – 10:30 am</td>
<td><strong>Round 1. Overall SBAC Claims</strong></td>
</tr>
<tr>
<td></td>
<td>• Study and discuss overall SBAC claims</td>
</tr>
<tr>
<td></td>
<td>• Study examples of general policy ALDs</td>
</tr>
<tr>
<td></td>
<td>• Discuss college and career readiness</td>
</tr>
<tr>
<td></td>
<td>• Scaffold overall claims into achievement levels</td>
</tr>
<tr>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td>10:30 – 10:45 am</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td>10:45 am – 12:00 pm</td>
<td><strong>Round 2. Overall SBAC Claims</strong></td>
</tr>
<tr>
<td></td>
<td>• Study and discuss Round 1 general policy ALDs</td>
</tr>
<tr>
<td></td>
<td>• Revise as needed</td>
</tr>
<tr>
<td></td>
<td>• Discuss defining phrases that differentiate achievement levels</td>
</tr>
<tr>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td>12:00 – 1:00 pm</td>
<td><strong>Lunch</strong></td>
</tr>
<tr>
<td>1:00 – 2:45 pm</td>
<td><strong>Round 3. Content-based SBAC Claims</strong></td>
</tr>
<tr>
<td></td>
<td>• Study and discuss content-specific SBAC claims</td>
</tr>
<tr>
<td></td>
<td>• Scaffold content-specific claims into achievement levels</td>
</tr>
<tr>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td>2:45 – 3:00 pm</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td>3:00 – 4:45 pm</td>
<td><strong>Round 4. Content-based SBAC Claims</strong></td>
</tr>
<tr>
<td></td>
<td>• Study and discuss Round 3 content-specific general policy ALDs</td>
</tr>
<tr>
<td></td>
<td>• Cross-content discussion of content-specific general policy ALDs</td>
</tr>
<tr>
<td></td>
<td>• Revise as needed</td>
</tr>
<tr>
<td>4:45 – 5:00 pm</td>
<td><strong>Evaluation</strong></td>
</tr>
<tr>
<td>5:00 pm</td>
<td><strong>Dismissal</strong></td>
</tr>
</tbody>
</table>

**Opening Session.** The day will begin with all meta-committee members meeting for an opening session where panelists will receive an overview of the purposes and uses of ALDs, their specific activities for
the day, and training on the materials that they will use. During training, CTB will provide guidance to panelists on using plain language that can be understood by a lay audience.

**Round 1 Activities.** Meta-committee panelists from both content areas will jointly participate in the first round of activities. This round will begin by studying the overall SBAC claims for both content areas. CTB and College Board will provide exemplars of general policy ALDs that have been used by other organizations. Table 8 shows general policy ALDs that were used by NAEP for its ALD and standard setting activities. Prior to the meeting, CTB will work with SBAC to identify particular organizations or states with general policy ALDs that fit with the guidance SBAC wants to provide panelists.

*Table 8. NAEP General Policy PLDs*

<table>
<thead>
<tr>
<th>NAEP*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced. Superior performance.</td>
</tr>
<tr>
<td>Proficient. Solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.</td>
</tr>
<tr>
<td>Basic. Partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.</td>
</tr>
</tbody>
</table>


Following the discussion, panelists will discuss the progression of how students in each achievement level demonstrate readiness for college and career in grade 11, followed by a similar discussion for grades 3 through 8. One of the keys to this activity will be for panelists to create defining phrases for each achievement level. The defining phrase is the “intended or take-away message in the definitions” of each achievement level (Egan, Schneider, & Ferrara, 2012, p. 86).

To facilitate this discussion, CTB and College Board will create a list of defining phrases that others have used to distinguish among the achievement levels. For example, “satisfactory” or “solid performance” are defining phrases that have been used to define the general performance in the proficient achievement level in some states (Egan, Schneider, & Ferrara, 2012). This list will provide the panelists with a starting point as they attempt to parse out how students in Levels 1, 2, 3, and 4 are progressing against the overall claims.

Once panelists have created a general policy descriptor for the overall claims, they will complete a short evaluation. The evaluations will be explained at the end of this section.

**Round 2 Activities.** Following the break, all meta-committee panelists will study, discuss, and revise the Round 1 general policy ALDs. They will also discuss the defining phrases that differentiate achievement levels. The panelists will create a list of the important defining phrases, and these will be used to guide the Round 3 and 4 activities. Once panelists indicate that their work is complete, they will be asked to complete an evaluation.

**Round 3 Activities.** Unlike the Rounds 1 and 2 activities where meta-committee panelists worked as a single group, panelists will be divided by content area for the Rounds 3 and 4 activities. In Round 3, panelists will study and discuss the content-specific SBAC claims. Using the defining phrases from Round 2, they will scaffold the content-specific claims into achievement levels. Following this work, they will complete an evaluation of the Round 3 activities.

**Round 4 Activities.** In Round 4, panelists will share their Round 3 work with the meta-committee for the other content area. This will be done to assure consistency between the content areas. Following this discussion, panelists can revise their Round 3 general policy descriptors. At the end of this round,
panelists will compile the overall general policy ALDs and the content-specific general policy ALDs into a single general policy ALD that will be used to guide the development of the grade-specific ALDs. An evaluation will be completed at the end of the day.

Day 2 and 3 Activities
The Grade 11 committee members for both content areas will convene on Days 2 and 3 of the workshop. The meta-committee members from the other grade-spans will participate in the Grade 11 activities. Table 9 shows the proposed schedule of Day 2 and 3 activities.

Opening Session. The day will begin with all Grade 11 committee members meeting for an opening session where panelists will receive an overview of the purpose of ALDs, their specific activities for the day, and training on the materials that they will use. During training, CTB will provide guidance to panelists on using plain language that can be understood by a lay audience. Upon completion of the opening session, panelists will split into their respective content areas.

Round 1. Panelists will begin Round 1 by studying and discussing the general policy ALDs that were created by the meta-committee. The majority of the day will be spent discussing and parsing the CCSS in the achievement levels as described above or parsing the summative assessment targets from each claim of the content specifications into achievement level. Once panelists have discussed and parsed the CCSS, the will compile the KSAs into achievement levels.

Round 2. Panelists will study and discuss the CAT and cluster emphases in Grade 11 to assure that there is continuity between the Grade 11 Range and Target ALDs and intended emphases of the clusters. Following this discussion, panelists will edit their Round 1 descriptors. The panelists will spend time making the ALDs usable for the grades 3 through 8 committees that will meet on Days 4 and 5 of the workshop.

Table 9. Proposed Schedule of Day 2 and 3 Activities

<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8:00 – 9:00 am</td>
<td>Opening Session</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Training on general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Orientation to materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Evaluation</td>
</tr>
<tr>
<td></td>
<td>9:00 am – 12:00 pm</td>
<td>Round 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Study and discuss general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Study example matrices of knowledge, skills, and abilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Discuss and Parse Common Core Content Standards</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Discuss and Parse content specifications</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td>Break</td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td>Lunch</td>
</tr>
<tr>
<td></td>
<td>1:00 – 2:45 pm</td>
<td>Continue Round 1</td>
</tr>
<tr>
<td>Day</td>
<td>Time</td>
<td>Activities</td>
</tr>
<tr>
<td>-----</td>
<td>----------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>2</td>
<td>2:45 – 3:00 pm</td>
<td>Break</td>
</tr>
<tr>
<td>3</td>
<td>3:00 – 5:00 pm</td>
<td>Complete Round 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Compile into ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td>5</td>
<td>5:00 pm</td>
<td>Dismissal</td>
</tr>
<tr>
<td>3</td>
<td>8:00 am – 12:00 pm</td>
<td>Round 2.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study cluster emphases and CAT emphases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise descriptors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td>Break</td>
</tr>
<tr>
<td>12:00 – 1:00 pm</td>
<td>Lunch and Dismiss</td>
<td></td>
</tr>
</tbody>
</table>

**Day 4 and 5 Activities**

The grade 3 through 8 committee members for both content areas will convene on Days 4 and 5 of the workshop. The meta-committee members from the grade 11 group will participate in these activities. Table 10 shows the proposed schedule of Day 4 and 5 activities.

As mentioned above, grade 11 ALDs will be written prior to the grades 3 through 8 ALDs so that the grade 3-8 panelists can work backward from grade 11. This will help assure that the scope and sequence of KSAs makes sense as students move from grade 3 to grade 11.

The activities of the grades 3 through 8 committee will look much like those of the grade 11 committee, except that panelists will study the grade 11 range and target ALDs.

**Round 1.** Panelists will begin Round 1 by studying and discussing the general policy ALDs that were created by the meta-committee. In addition, they will study the grade 11 range and target ALDs. The meta-committee members from grade 11 will float between the grade-span groups to answer questions about the grade 11 range and target ALDs, and they will help the other grade-span groups as they attempt to parse the CCSS. As with the grade 11 group, the majority of the day will be spent discussing and parsing the CCSS in the achievement levels as described above and parsing the summative assessment targets from each claim of the content specifications into achievement level. Once panelists have discussed and parsed the CCSS and summative assessment targets from each claim of the content specifications, the will compile the KSAs into achievement levels.

**Round 2.** Panelists will begin by engaging in a cross-grade discussion. To facilitate this, two members from the grade 4–5 span will work with the grade 3–4 span and the other two members will work with the grade 7–8 span. The meta-committee members from each grade span will float among groups to assist with the cross-grade articulation of standards. Following the cross-grade discussions, panelists will revise their Round 1 range and target ALDs.

**Round 3.** Panelists will study and discuss the CAT and cluster emphases in their respective grade levels to assure that there is continuity between the range and target ALDs and intended emphases of the clusters. Following this discussion, panelists will edit their Round 2 descriptors. Again, the meta-committee members will float among the groups to assist with the cross-grade articulation.

**Meta-Committee.** At the end of the workshop, the meta-committee for each content area will examine the range and target ALDs to assure that there is a logical progression of KSAs across grade levels.
levels. They will pay particular attention to the scope and sequence of the ALDs to assure that the learning gains are achievable through reasonable instruction during the school year.

<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>8:00 – 9:00 am</td>
<td><strong>Opening Session</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Training on general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Orientation to materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>9:00 am – 12:00 pm</td>
<td><strong>Round 1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study and discuss general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study example matrices of knowledge, skills, and abilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss and Parse Common Core Content Standards</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss and Parse content specifications</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td><strong>Lunch</strong></td>
</tr>
<tr>
<td></td>
<td>1:00 – 2:45 pm</td>
<td><strong>Continue Round 1</strong></td>
</tr>
<tr>
<td></td>
<td>2:45 – 3:00 pm</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>3:00 – 5:00 pm</td>
<td><strong>Complete Round 1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Compile into ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>5:00 pm</td>
<td><strong>Dismissal</strong></td>
</tr>
<tr>
<td>5</td>
<td>8:00 am – 12:00 pm</td>
<td><strong>Round 2</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Cross-grade discussions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise as needed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td><strong>Lunch</strong></td>
</tr>
<tr>
<td></td>
<td>1:00 – 2:45 pm</td>
<td><strong>Round 3</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study cluster emphases and CAT emphases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise descriptors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td>Day</td>
<td>Time</td>
<td>Activities</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>3:00 – 5:00 pm</td>
<td><strong>Meta-Committee</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise descriptors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>5:00 pm</td>
<td><strong>Dismissal</strong></td>
</tr>
</tbody>
</table>

**After the Workshop**

Our experience with ALD development is that panelists are very successful in creating a rough draft; however, panelists find the editing and style task burdensome when working in a small group. Thus, we recommend that the content experts at CTB and College Board assist with editing the general policy, range, and target ALDs so that they are concise yet detailed enough to support item and task development, test design, and standard setting. CTB and College Board experts will also be to edit the rough drafts so that they are in plain language that is easily understood by non-technical stakeholders, such as parents and policy makers.

If desired by SBAC, CTB could ask the meta-committee to evaluate the final product before it goes to the next group for revisions.

**Evaluations**

CTB routinely conducts evaluations of the process and results as part of the workshops that we conduct, including workshops for standard setting and for writing ALDs. Through our extensive experience conducting workshop, we have learned that it is valuable to collect evaluations not only at the end of the workshop but also at key phases of the workshop process to gauge participant understanding and reaction to elements of the ALD-writing process. CTB recognizes the importance of the panelist evaluation as evidence of the validity of the ALDs. As such, the panelist evaluation is one more piece of the workshop design that must be carefully considered before the workshop begins. In this section, CTB highlights the administration of the panelist evaluations and then summarizes the construction and content of the panelist evaluations.

**Administration**

CTB will administer panelist evaluations after all key activities. Evaluations will be administered

1. after the opening session;
2. immediately after Round 1, 2, and 3 judgments;
3. after the presentation of final recommendations.

The administration of the evaluations is called out in the agenda presented in Tables 8, 9, and 10.

**Construction**

To construct panelist evaluation for the ALD workshop, we can look for guidance from those evaluations constructed for standard setting workshops. Hambleton and Pitoniak (2006) offer some general guidance indicating panelists should be asked about their understanding of key elements of the standard setting process, their feedback on the timing of various aspects of the standard setting and the quality of training, and their confidence in the performance standards. This guidance can easily be translated to questions for the ALD workshop.

Panelist evaluations may be used as evidence of procedural validity because they will be used to measure panelist understanding, confidence, and comfort with the ALD process, the ALD recommendation, the workshop materials, etc. Panelist understanding is especially crucial to the evidence of procedural
validity: if panelists do not understand the ALD task (e.g., how to parse the CCSS), then this can undermine confidence in the final recommended ALDs. In constructing the panelist evaluation, CTB pays attention to length, rating scale, wording, and layout. Procedural validity is discussed in more detail in Section 2.13.

**Length.** The panelist evaluation must be of a reasonable length so that it does not interfere with the timing of the ALD workshop nor does it fatigue panelists. This is often difficult to do in practice because there are so many questions that we would like to ask panelists. CTB has allotted 10 minutes at various intervals in the agenda for the completion of the evaluation. CTB anticipates that panelists spend approximately 30 seconds on typical evaluation like-type questions and 1 minute on items that measure understanding. Using these estimations, survey tools will be constructed. Because timing is a consideration, CTB recommends that surveys being administered at points where there are obvious breaks (e.g., at the end of each round).

**Rating Scale.** CTB recommends using a rating scale that asks panelists for level of agreement from strongly disagree to strongly agree. The use of a neutral category is preferential, and CTB will look to SBAC for guidance on its use in the rating scale.

**Wording.** CTB pays special attention to the way in which evaluation items are constructed. The wording of these items must be accessible to all panelists; thus, it is important to assure any esoteric words and phrases have been well-defined for panelists. In addition, CTB is careful to construct items that neither leading nor double-barreled (Nachmias and Nachmias, 1992). CTB also recommends wording survey items so that negative responses are sometimes desired and positive response are sometimes desired. This practice helps detect response sets, where panelists answer all items in the same direction, regardless of the content of the item (Nachmias and Nachmias, 1992). Finally, it is especially important that the item is worded in such a way that panelists understand the question we are asking. Ideally, think alouds could be conducted for all evaluation items to ensure panelists and facilitators are on the same page; however, this type of study is usually logistically impossible.

**Layout.** The evaluation should be laid out in a visually attractive manner that is easy for panelists to use. CTB will create evaluations that are easy to understand, simple to complete, and efficient to analyze.

**Summary**

This section began with a request from the Consortium that the proposed methodology results in grade-specific ALDs that demonstrate numerous characteristics. The table below shows each of these characteristics along with the manner in which the proposed CTB and College Board methodology will result in ALDs with these characteristics.

<table>
<thead>
<tr>
<th>Table 11. Aspects of Methodology that Support SBAC’s Desired Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristic</strong></td>
</tr>
<tr>
<td>Consistency with the Common Core State Standards.</td>
</tr>
<tr>
<td>Consistency with the CONSORTIUM’s content specifications</td>
</tr>
<tr>
<td>Consistency with skills and knowledge students will be able to demonstrate on the CONSORTIUM’s computer adaptive summative assessments and performance tasks, as described by the CONSORTIUM’s content specifications.</td>
</tr>
<tr>
<td>Consistency with the scores and reporting information to be produced by summative computer adaptive tests and performance tasks, as described by the CONSORTIUM’s content specifications.</td>
</tr>
<tr>
<td>Consistency across grade levels, with logical developmental progression from one grade level to the next.</td>
</tr>
</tbody>
</table>
### Characteristic
- Consistency with expectations for performance at each grade that will be necessary for students to be college and career ready by the end of high school (i.e., working backward from college and career readiness at the end of high school, students’ knowledge and skills should build sequentially to ensure steady annual progress).
- Consistency with learning gains achievable through reasonable effective instruction during a school year for a typical student demonstrating achievement of the ALDs.
- A balance of conciseness and level of detail that will support item and task development, test design, and standard setting.
- Plain language that is easily understood by non-technical stakeholders such as parents and policy makers.

### Methodology
- Cross-grade discussions for grades 3-8
- Meta-committee members work with other grade groups
- Meta-committee reviews ALDs
- Final edit by CTB and CB content experts
- Participation of higher education member with expertise in achievement gains
- Training for panelists with examples of final products
- Final edit by CTB and CB content experts

### Facilitation of all steps in the ALD development process with expertise in ALD development

Staff from CTB and College Board will facilitate all steps in the ALD development process. Dr. Karla Egan will serve as the lead psychometric facilitator from CTB. Dr. Egan has led over 50 standard setting workshops, where target and reporting ALDs were developed as part of the workshop. In addition, she has conducted extensive research on the ALD development process which led to the ALD framework proposed here (Egan, Schneider, & Ferrara, 2012; Egan, Schneider, & Ferrara, 2011; Schneider, Egan, Huff, & Tully, 2011; Egan, Ferrara, Schneider, & Barton, 2009; Schneider, Egan, Siskind, Brailsford, & Jones, 2009; Mercado & Egan, 2005).

Gretchen Schultz will serve as the lead content facilitator from CTB. Ms. Schultz has had a professional career as an educator, state assessment specialist, and developer of high-stakes tests. She has a Master of Liberal Arts, with emphasis on History of Ideas and English studies and a Bachelor of Arts in Humanities. She has been pivotal in developing materials to guide and support item and test development as a teacher, as a Maryland State Department of Education assessment specialist, and as Principal Assessment Editor for CTB. In her role at CTB, Ms. Schultz has led the study of the CCSS and the application of evidence-based development processes to assessing the CCSS. In addition to assisting in the development of the initial ALDs, Ms. Schultz will oversee the revision of the ALDs during the five review rounds.

Dr. Andrew Wiley and Dr. Pamela Kaliski will serve as lead psychometric facilitators from the College Board for grade 11 mathematics and ELA, respectively. Dr. Wiley has extensive experience working within the educational environment of college and career readiness, including the management of a team of content specialist who were instrumental in the development of the college and career readiness standards in collaboration with the CCSSO and Achieve. His experience includes the identification of benchmarks for college readiness on the SAT, and working in the development of validation of assessments designed to assess college readiness. Dr. Kaliski has led standard settings for several programs at the College Board. She has co-facilitated the development of ALDs for several Advanced Placement exams, and worked on developing ALDs used for standard setting purposes for ACCUPLACER. Dr. Kaliski is also a co-facilitator of a full day evidence-centered assessment design training workshop at the annual meeting of the National Council of Measurement in Education, where she will be facilitating the component about ECD-based ALDs. Additionally, Dr. Kaliski conducts research related to ALD development, validation, and item alignment with ALDs (Kaliski, Huff, & Barry, 2011; Reshetar, Chajewski, Lionberger, Plake, & Kaliski, 2012).
2.8 If the methodology proposed requires the use of panel judgments, convene panelists to follow the methodology in designing initial ALDs, including all requirements listed previously.

CTB will be responsible for the logistics associated with Initial Achievement Level Descriptors workshop, including:

- locating and contracting for suitable meeting rooms that are well-lighted, that have plenty of work space, and that are comfortable for working groups,
- arranging for catering services, as needed, for the workshop,
- arranging for meeting equipment as necessary for the work of the committees,
- establishing and maintaining procedures for item security,
- handling travel and lodging arrangements and reimbursements for participants, and
- preparing all training and meeting materials.

As explained in Section 2.7, CTB and College Board staff who are experienced in developing ALDs will guide the committee members through the approved methodology so that they can draft initial ALDs. Dr. Karla Egan will serve as the overall facilitator for the workshop. Heather Farina will facilitate committee members in grades 3-8 Mathematics and Gretchen Schultz will lead panelists in grades 3-8 ELA. Dr. Andrew Wiley will facilitate the grade 11 Mathematics group, and Dr. Pamela Kaliski will facilitate the grade 11 ELA group.
2.9 Edit the draft initial ALDs arising from the development sessions to meet all criteria listed above, including a two-week time period for review other contractors and CONSORTIUM workgroups as specified previously.

Content experts from CTB and College Board will edit the draft initial ALDs that are created during the ALD development workshop. Our content experts have extensive experience in editing draft ALDs so that they remain true to the original intent of the workshop panelists. As our content experts edit the draft initial ALDs, they will pay particular attention to:

1. Consistency with the Common Core State Standards.
2. Consistency with the CONSORTIUM’s content specifications in ELA and mathematics (see Exhibits I and J).
3. Consistency with skills and knowledge students will be able to demonstrate on the CONSORTIUM’s computer adaptive summative assessments and performance tasks, as described by the CONSORTIUM’s content specifications.
4. Consistency with the scores and reporting information to be produced by summative computer adaptive tests and performance tasks, as described by the CONSORTIUM’s content specifications.
5. Consistency across grade levels, with logical developmental progression from one grade level to the next.
6. Consistency with expectations for performance at each grade that will be necessary for students to be college and career ready by the end of high school (i.e., working backward from college and career readiness at the end of high school, students' knowledge and skills should build sequentially to ensure steady annual progress.
7. Consistency with learning gains achievable through reasonable effective instruction during a school year for a typical student demonstrating achievement of the ALDs.
8. A balance of conciseness and level of detail that will support item and task development, test design, and standard setting.
9. Plain language that is easily understood by non-technical stakeholders such as parents and policy makers.

CTB and College Board content experts have deep knowledge of the CCSS and we bring a substantial body of expertise to this work. CTB's assessment experts have developed item specifications to facilitate the evidence-based development of items for catalog products, state assessments, and other projects. Our ELA experts have studied the ELA standards, including the standards for Literacy in History/Social Studies, Science, and Technical Subjects. Our mathematics experts are familiar with the content and organization of the mathematics standards as well as the intent and application of the Mathematical Practice Standards. We are familiar as well with the SBAC Content Specifications for ELA and Mathematics from Smarter Balanced and the similar content frameworks from PARCC. We have developed CCSS-aligned assessments in both ELA and Mathematics for large-scale assessment projects for states as well as for secure catalog products. CTB is currently developing CCSS-aligned items for the Smarter Balanced RFP 14, Pilot Test Item/Task Development.

2.10 Elicit and collect comprehensive feedback regarding the draft initial ALDs from CONSORTIUM K-12 and higher education state membership. Particular emphasis on comprehensive feedback from higher education is required for the high school ALDs.

Once the draft initial ALDs are completed, CTB will manage the process of gathering stakeholder feedback. We will work closely with SBAC staff and work group leaders to identify appropriate candidates for the review panels. We will begin recruiting these reviewers immediately upon award so
that the various review panels will be in place as soon as the initial draft of the ALDs is completed.

The College Board is a membership institution, comprised of more than 5,900 individuals, institutions, and agencies. Higher education delegates represent both public and private institutions, as well as higher education systems from all states and territories. The College Board has extensive experience in soliciting professionals from higher education to participate in standard setting and assessment development activities. Using these resources, the College Board will work with Smarter Balanced to determine a meaningful sample of representatives from various states and institution types to serve on the ALD development panel.

We propose to use a series of webinars and questionnaires to gather information from the various review panels. Webinar dates will be set in early June and reviewers will be recruited who will be able to participate in the reviews during the established timeframe. Due to the short timeframe for the completion of the panel reviews, we will schedule the webinars and begin recruiting reviewers immediately upon contract award.

We anticipate the reviewer groups will consist of volunteers who are willing to participate in a live webinar session followed by the completion of an online review questionnaire. Each one-hour live webinar will present the following information:

- Introduction to the development of the ALDs
- Brief description of the methodology used by the development committee
- Overview of the organization of the ALD document(s)
- Guidelines for review and instructions for completing the online survey.

Review questions will reflect the criteria described in Section 2.10.

**Version 2**

The first review panel will consist of representatives from the contractors who have worked on the previous SBAC projects: 04, 05, 06, 09, 14, and 15. To date, that would include representatives from Measured Progress, Educational Testing Service (ETS), and CTB/McGraw-Hill. We will also include representatives from the associated work groups: Item Development, Test Design, Performance Tasks, Accessibility and Accommodations, and the Validation and Psychometric work groups. Because of the webinar and questionnaire format that we are proposing, we could accommodate 1-2 members from each of these stakeholder groups and anticipate between 10 and 15 participants in this round of review. CTB content experts will review the comments from this stakeholder group, and appropriate revisions will be made to produce Version 2. The specific process for incorporating comments is described in Sec. 2.11.

**Version 3**

The second review panel will consist of the Smarter Balanced Executive Committee and the Smarter Balanced higher education state membership. Because of the summer timeframe, it is possible that not all Higher Education Leads would be available to participate, but would be able to name a qualified delegate. We would anticipate that this review would consist of 20-25 participants. College Board will facilitate the webinar and review activities for this review. CTB content experts will work with the College Board to apply the feedback from these reviews to produce Version 3.

**Version 4**

The third review panel will solicit feedback for the Smarter Balanced TAC and the K-12 state membership. Given the timing of these reviews, it is likely that this review will occur after the regular TAC meeting in July. We will work closely with the SBAC Executive Committee to identify the appropriate state representatives to participate in this review. In order to be inclusive of all SBC member states, we anticipate that this review may accommodate up to 30 people. CTB will coordinate this review and the preparation of Version 4.
**Version 5**
Throughout the feedback and review cycles described above, we will work with the SBAC leadership to identify other appropriate stakeholder groups from which to elicit feedback. We will identify PARCC representatives to invite perhaps PARCC TAC members or others working on similar ALD development for PARCC. We will also reach out to other policy or educational groups from which SBAC might wish to gain feedback on this draft of the ALD documents. The participants in this review will be identified early in the project. CTB will facilitate this review and recommend revisions to SBAC to create Version 5.

**Version 6**
Version 6 will represent the final ALD deliverables from this contract.

**2.11 Revise the draft initial ALDs based on CONSORTIUM states’ feedback as described previously.**
As described in Section 2.10, CTB will gather feedback and revise the draft ALD documents in each version. Our content experts will be responsible for editing all drafts as the ALDs are reviewed and revised. Our content experts have experience in crafting performance level descriptor language for large-scale assessments and will receive training in the methodology for the development of the ALDs that will be used by the Development Committee.

Content editors will be responsible for incorporating revisions to the ALDs at each content and grade level after each review round. The Lead Senior Research Scientist and the Content Director will oversee this work. As review comments are received from each round of review, content experts will organize and combine the comments from various stakeholders. Potential edits will be reviewed with the Senior Research Director and Content Director for internal approval. In addition to tracked changes or similar notes within the document, documentation will be provided to SBAC in the form of a change record to summarize the revisions made at each step. Before each new review round, each set of ALDs will go through a thorough proofing review to ensure that drafts reviewed by stakeholders are free from errors.

The proposed process for editing the draft ALDs will include the following steps:

- Receive multiple survey responses from individual stakeholder reviews.
- Combine review comments into single document.
- Review stakeholder comments with Lead Research Scientist and Content Lead to identify edits/comments to be incorporated.
- Apply approved edits to document using tracked changes or similar notations. Prepare accompanying tracking document with notations explaining each applied edit.
- Finalize document as new version and submit for next round of review or final approval.

**2.12 Elicit and collect comprehensive feedback regarding the draft initial ALDs from the CONSORTIUM’s TAC and the public as described previously.**
As described in Sec. 2.10, CTB will manage the process of gathering stakeholder feedback once the draft initial ALDs are completed. We will work closely with SBAC staff and work group leaders to identify appropriate candidates for the review panels. We will begin recruiting these reviewers immediately upon award so that the various review panels will be in place as soon as the initial ALD-Development Committee completes the initial draft of the ALDs.

We propose to use a series of webinars and questionnaires to gather information from the various review panels. Webinar dates will be set in early June and reviewers will be recruited who will be able to participate in the reviews during the established timeframe. Due to the short timeframe for the completion of the panel reviews, we will schedule the webinars and begin recruiting reviewers immediately upon contract award.
2.13 Revise the draft initial ALDs based on TAC feedback and public feedback as described previously.

Once the draft initial ALDs are completed, CTB will manage the process of gathering stakeholder feedback as described in Sec. 2.10. We will work closely with Smarter Balanced staff and work group leaders to identify appropriate candidates for the review panels. We will begin recruiting these reviewers immediately upon award so that the various review panels will be in place as soon as the initial draft of the ALDs is completed.

We propose to use a series of webinars and questionnaires to gather information from the various review panels. Webinar dates will be set in early June and reviewers will be recruited who will be able to participate in the reviews during the established timeframe. Due to the short timeframe for the completion of the panel reviews, we will schedule the webinars and begin recruiting reviewers immediately upon contract award.

We anticipate the reviewer groups will consist of volunteers who are willing to participate in a live webinar session followed by the completion of an online review questionnaire. Each one-hour live webinar will present the following information:

- Introduction to the development of the ALDs
- Brief description of the methodology used by the development committee
- Overview of the organization of the ALD document(s)
- Guidelines for review and instructions for completing the online survey
- Review questions will reflect the criteria describe in Sec. 2.10.

The documentation of the development of the initial ALDs is an important step for SBAC in establishing validly interpretable and meaningful performance standards. We also understand that establishing ALDs requires ensuring the defensibility of the procedures we implement and providing evidence of the validity of interpretations about student achievement based on the ALDs. In this section we describe the how we will document procedures and panelists’ understandings of their tasks in aligning expectations with particular achievement levels. We also describe our plans for documenting evidence to support procedural validity and developing validity arguments to support interpretations about student achievement based on the Achievement Levels. The draft technical report will describe the methodological, conceptual, and logistical elements of the five-day ALD development process. The final technical report will incorporate the process and feedback activities that are a component of the RFP after the workshop.

We begin this section with an overview of the proposed structure of the draft technical report, and we then provide an in-depth discussion of procedural validity, which will be important as we create the initial ALDs.

**Proposed Structure of the Draft Technical Report**

We understand that the draft technical report is an important means for SBAC to present information regarding the initial ALDs to various groups of SBAC stakeholders. The technical report also presents argumentation and related evidence for external reviewers and assessment professionals to judge the extent to which the intended assessment uses of initial ALDs are justified.

Prior to the ALD workshop, CTB will document the planned ALD development activities through a rigorous planning document that describes each activity. This meticulous planning will provide the foundation for the comprehensive and coherent technical document that SBAC will use as a central source of validity evidence for its initial ALDs.

We have successfully supported many other states as they successfully completed the peer review process. One of the key components of our customers’ overwhelming success in peer review has come
directly from our rigorous technical documentation. Our technical reports will include specific links to appropriate evidence within the Standards for Educational and Psychological Testing (AERA, APA, and NCME) and the Peer Review Guidance for ESEA.

One important structural change in the proposed updates for the Standards is the categorization of standards into three sections: Foundations (e.g., reliability, validity, test fairness), Operations (e.g., test design and development, administration, scoring, and reporting), and Applications (e.g., tests as educational assessment). This categorization happens to align with three critical elements that are present in argument-based approaches such as the evidenced-centered assessment design. These elements are qualities of score-based inferences, assessment activities to produce and distribute scores, and uses of the scores. Accordingly, we propose to organize the draft technical report into three major sections: Section I will provide information about the applications of the initial ALDs; Section II describes key operations, including the workshop design; and Section III presents evaluative information, including validity evidence.

Table 12 shows the proposed outline for the draft technical report. CTB is happy to adjust the design to incorporate feedback or requirements from the Consortium or its TAC.

**Table 12. Outline of the Technical Report**

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Section Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXECUTIVE SUMMARY</strong></td>
<td>Goal of the Executive Summary: To present key information and highlights of the draft technical report.</td>
</tr>
<tr>
<td><strong>SECTION I: APPLICATIONS</strong></td>
<td>Goals of Section I (Chapter 1): To introduce the initial ALDs and their intended uses.</td>
</tr>
<tr>
<td>Chapter 1. Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1 Intended Uses of the Initial ALDs</td>
<td></td>
</tr>
<tr>
<td>1.2 Target Population for Using Initial ALDs</td>
<td></td>
</tr>
<tr>
<td><strong>SECTION II: OPERATIONS</strong></td>
<td>Goals of Section II (Chapters 2-6): To describe major activities in operations.</td>
</tr>
<tr>
<td>Chapter 2. Panelists</td>
<td></td>
</tr>
<tr>
<td>2.1 Panelist Recruitment</td>
<td></td>
</tr>
<tr>
<td>2.2 Panelist Demographics</td>
<td></td>
</tr>
<tr>
<td>Chapter 3. Workshop Design</td>
<td></td>
</tr>
<tr>
<td>3.1 Description of Workshop Goal</td>
<td></td>
</tr>
<tr>
<td>3.2 Description of General Policy ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.3 Description of Grade- and Content-Specific Range and Target ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.4 Quality Control Evidence</td>
<td></td>
</tr>
<tr>
<td>Chapter 4. General Policy Descriptors</td>
<td></td>
</tr>
<tr>
<td>Chapter 5. Range and Target Achievement Level Descriptors</td>
<td></td>
</tr>
<tr>
<td><strong>SECTION III: EVALUATION</strong></td>
<td>Goals of Section III (Chapter 6): To present critical info. for evaluating qualities of the initial ALDs.</td>
</tr>
<tr>
<td>Chapter 6. Validity</td>
<td></td>
</tr>
<tr>
<td>6.1 Types of Validity Evidence</td>
<td></td>
</tr>
<tr>
<td>6.1.a Procedural Validity</td>
<td></td>
</tr>
<tr>
<td>6.1.b Panelist Feedback</td>
<td></td>
</tr>
<tr>
<td>6.2 Summary of Validity Evidence for the Intended Test Uses</td>
<td></td>
</tr>
<tr>
<td>Appendix A. Training Materials</td>
<td></td>
</tr>
<tr>
<td>Appendix B. Results of Evaluations</td>
<td></td>
</tr>
<tr>
<td>Appendix C. Planned Review Cycles</td>
<td></td>
</tr>
</tbody>
</table>
Evidence of Procedural Validity

Throughout the ALD development workshop, we will collect panelist evaluations that will enable us to clarify and reinforce concepts and procedures as needed and to document procedural validity. Documentation of the myriad elements of the ALD process is intrinsically important to the evidence of procedural validity as well as content validity. This documentation should detail all of the implemented steps of the ALD development process, information used to inform the process, and the panelists understanding of concepts and information (Plake, 2008, Kane, 2003; Hambleton and Pitoniak, 2006) associated with the construct of being college ready. The elements associated with the ALD process can be conceptualized in three categories: methodological, conceptual, and logistical. These elements are introduced throughout the ALD workshop, and thus may be broken out by round. At a high level, the proposed plan for the ALD workshops calls for an opening session, multiple rounds of revisions, consensus-building feedback, and workshop finalization. In addition, briefing materials will be sent in advance of the ALD development workshop. Table 13 summarizes the methodological, conceptual, and logistical elements associated with each phase of the process. In this section, we explicate the ways in which these various elements will be captured as evidence of procedural validity.

Methodological Elements

The methodological elements consist of the tasks and materials specifically associated with the ALD development process. For example, the common core standards and content specifications are examples of a methodological element. It is important to gauge panelist understanding and reactions to the methodological elements. Not only must panelists understand the materials that they are using, they should also think that these materials and the associated tasks with those materials enhance the ALD development process. It is very possible that panelists can understand a task (e.g., parsing the common core standards) while, at the same time, think it completely unnecessary to making decisions about which performance level to place a particular knowledge, skills, or abilities.

To capture understanding of and reaction to the methodological elements, CTB will administer surveys following each phase of the ALD development process. The questions concerning methodological elements will ask panelists to gauge the usefulness, effectiveness, and understanding of particular tasks or materials. Where possible, specific test-like questions will be administered that further measure panelist understanding of certain tasks and materials. After all, panelists may believe that they understand a task when, in reality, they do not.

For the methodological elements, their implementation will be explicated in detail in the draft technical report that CTB provides to SBAC following the ALD development workshop. For example, CTB will provide the training materials and overheads used for opening session training in the Technical Report. CTB will also explicate how the training was conducted and the explanation given to panelists for each phase of the workshop.

Table 13 summarizes each of the methodological elements for which panelists will answer survey and/or test questions. As an example, four different methodological elements are introduced in Round 2, and panelist will answer survey and/or test questions for each element immediately following Round 2. The findings associated with the round-by-round survey and test questions will be evaluated and reported in the draft technical report.

Conceptual Elements

The conceptual elements refer to the ideas introduced at various times during the ALD development process. It is important to evaluate how well panelists understood the myriad of concepts that are introduced during the ALD development process. For most panelists, the ideas associated with defining college and career readiness will be completely new.

Panelist understanding of conceptual elements will again be measured by evaluations administered after key phases of the ALD-development process. The questions pertaining to conceptual elements may ask panelists to rate their confidence and comfort with particular elements and understanding of particular
elements. Where possible, panelists will again be asked test like questions to further gauge their understanding of the conceptual elements. Table 13 summarizes concepts that are central to panelist understanding. If panelists do not grasp the bulk of these concepts, then it can be argued that they did not understand the judgment tasks necessary to develop the ALDs.

After final versions of the ALDs are shown, panelists will be asked survey questions related to their perceived validity of the ALDs in terms of usefulness for standard setting, score interpretation, and item writing. We will also investigate the usefulness of the ALDs from a classroom assessment perspective. Do the ALDs adequately describe, for the classroom teacher, the degree of proficiency a student needs to reach in a single year to be, for example, on target for being college ready? If panelists respond negatively to any of these questions, the procedural validity of the ALD-development process may be undermined (Kane, 2003). It will difficult for SBAC to have faith in the target ALDs for standard setting, if panelists, for example, self-report they are not comfortable with descriptors' ability to describe the content-based expectations of a student in a particular grade.

**Logistical Elements**

Finally, the logistical element refers to the planning, coordination, and implementation of various process elements, as an example the time allocated for opening session. Panelist reaction to the timing associated with the different pieces of the process is important to document. If panelists indicate too little time was spent, for example, studying adjacent grade-level ALDs, then this may correlate with their confidence that ALDs are consistent across grade levels: too little time for defining college readiness may equal low confidence that ALDs adequately describe the skills in the CCSS students need to master to be college ready. CTB recognizes the importance of time in the ALD development process: panelists should not be rushed nor should they have large gaps with no activity. Panelists must believe that their time is well-utilized during the process. CTB has extensive experience in designing and implementing such activities. Our experience is brought to bear in the agenda proposed in Section 2.7. The agenda will again be presented in the draft technical report as evidence of procedural validity.

Of particular import among these elements is the selection of panelists. Unlike most elements, panelist selection will be documented in two ways. First, CTB will provide a detailed plan for the selecting the panelists. The details of this plan will be documented in the draft technical report. In addition to documenting the diversity and representativeness of the ALD development panel, CTB will ask panelists to rate their own perceptions of panel diversity and representativeness.

Finally, the logistical elements also cover the lodging and conference facilities as well as the food and snacks panelists received. These elements should be seamless to panelists and should not interfere with their work. CTB has learned that a 3:00 pm snack break with cookies and soda goes far in keeping panelists satisfied. Table 13 summarizes the logistical elements for the phases of the proposed ALD-development workshop.

**Summary**

During the ALD-development workshop, panelists will be continually monitored regarding the methodological, conceptual, and logistical elements that comprise the ALD-development workshop. CTB will examine the results of the various evaluations to assure that favorable panelist attitudes and panelist understanding is maximized throughout the process. For example, survey results following the training session may indicate panelists do not understand the purpose of general policy ALDS. CTB will use this information to re-train panelists on the concept of general policy ALDs and how to use them.

All information related to the methodological, conceptual, and logistical elements will also be fully explicated in the draft technical report that summarizes the ALD-development process. Any training materials and/or overheads pertaining to these elements will be contained in the draft technical report. In this report, CTB will explicate how these elements are supporting evidence for the procedural validity of the ALD-development workshop.
**Table 13. Methodological, Conceptual, and Logistical Elements Associated with each phase of the Grade 11 Range and Target ALD-Development Workshop.**

<table>
<thead>
<tr>
<th>ROUND</th>
<th>Methodological</th>
<th>Conceptual</th>
<th>Logistical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-workshop</td>
<td>• Briefing materials</td>
<td></td>
<td>• Panelist selection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Briefing materials received in a timely manner</td>
</tr>
<tr>
<td>Opening</td>
<td>• Orientation</td>
<td>• Purpose of general policy ALDs</td>
<td>• Timing of presentations</td>
</tr>
<tr>
<td>Session</td>
<td>• Process Training</td>
<td>• Purpose of range and target ALDs</td>
<td>• Timing of session</td>
</tr>
<tr>
<td></td>
<td>• Frameworks</td>
<td>• Purpose of SBAC assessment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Purpose of Defining College Readiness</td>
<td></td>
</tr>
<tr>
<td>Round 1</td>
<td>• CCSS Content Specifications</td>
<td>• Relationship between general policy ALDs and range/target ALDs</td>
<td>• Timing of study of CCSS</td>
</tr>
<tr>
<td>Elements</td>
<td>• Exemplar ALDs</td>
<td>• Relationship between range and target ALDs</td>
<td>• Timing of study of content specifications</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Consistency of ALDs between grade levels</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Consistency with achievable learning gains</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Influence of other panelists</td>
<td></td>
</tr>
<tr>
<td>Round 2</td>
<td>• Training</td>
<td>• Consistency with KSA to be demonstrated on CAT</td>
<td>• Timing of training</td>
</tr>
<tr>
<td>Elements</td>
<td>• CAT Emphases</td>
<td>• Influence of other panelists</td>
<td>• Timing of study of CAT and Cluster Emphases</td>
</tr>
<tr>
<td></td>
<td>• Cluster Emphases</td>
<td>• Influence of meta-committee members</td>
<td>• Timing of Round 2</td>
</tr>
<tr>
<td></td>
<td>• Group discussions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>• Final Feedback</td>
<td>• Comfort with ALDS</td>
<td>• Timing of final presentation</td>
</tr>
<tr>
<td></td>
<td>• Method</td>
<td>• Perceived validity of ALDS</td>
<td>• Overall process time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Perceived validity of method</td>
<td>• Lodging</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Confidence in process</td>
<td>• Food</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Usefulness for purposes</td>
<td>• Diversity of panel</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Representativeness of panel</td>
</tr>
</tbody>
</table>


2.14 Present at one (1) TAC meeting regarding preliminary products before finalizing the initial ALDs and technical report.

CTB is committed to supporting SBAC. Lead Research Scientist, Dr. Karla Egan and Content Lead, Gretchen Shultz will be available to present the results of and facilitate discussion about the preliminary products from the panelists' workshops. Prior to this meeting, we will prepare the preliminary products from the panelists' workshops as well as a draft version of the technical report for review by the TAC. During the TAC meeting, we will present both sets of materials and facilitate discussion. CTB Research Scientist, Dr. Karla Egan, has conducted multiple such presentations for numerous state TACs, and she is skilled at presentation and facilitation with technical and lay audiences.

2.15 Revise initial ALDs as directed by the CONSORTIUM to provide any needed consistency of initial ALDs with PARCC.

CTB will revise the initial ALDs, including the general policy ALDs and the content- and grade-specific range and target ALDs, as directed by SBAC to provide consistency with the PARCC ALDs. The goal of consistency between SBAC and PARCC performance standards requires coordination between the two. First, we note that consistency does not imply equivalence. Different test designs and unique assessment artifacts (e.g., choice of scaling models, scoring methods, etc.) will result in unique constructs and scales. That is, even if the same students took both SBAC and PARCC assessments, and were scored on the respective SBAC standardized scales, different distributions of SBAC student scores would likely result from the different tests, and students’ percentile ranks would likely be different, sometimes substantially, on the two distributions. Thus, the goal of consistency, not equivalence is appropriate. It would seem desirable that students considered “on track” or “four year college-ready” on one consortium’s assessment would have similar levels on the other consortium’s assessment.

Fostering Consistency through Cooperation

Consistency can be fostered (or induced) in many ways, and it requires the cooperation of the two consortia. The degree of cooperation and agreement will be reflected in the consistency of the results. The following cooperative strategies could foster consistency between SBAC and PARCC performance standards:

Common General Policy Achievement Level Descriptors. Consistency between the SBAC and PARCC achievement standards will be supported by each SBAC’s adopting common a priori policy ALDs. Policy ALDs describe the policy makers’ vision of the goals and expected rigor for the final performance standards (Egan, Schneider, & Ferrara, 2012). These ALDs are free of specific content expectations. For instance, the label Proficient may be defined in a policy ALD as “demonstrates mastery of the current grade’s standards and is well prepared for the more challenging content of the next grade.”

Commonality between Range and Target Achievement Level Descriptors. Range descriptors that reflect the expectations of students in each achievement level and should guide the development of the specific content of each consortium’s assessment. The target ALDs reflect the specific knowledge, skills, and abilities that should be possessed by a student who is right at the cut score. Like the range ALDs, the target ALDs will reflect the specific content of each assessment. For this reason, the goal of having the same Range and Target ALDs for both consortia may not be possible. However, to the degree possible, cooperative development of the descriptors may be economically efficient and will support the
desire for as much commonality as possible, given the differences between the two assessment systems.

**Common CCR benchmarks.** Consistency between the two consortiums’ performance standards will be supported by providing panelists with common, or comparable, CCR benchmarks. Thus, if SAT college readiness benchmarks were provided to SBAC, it would be desirable to provide the analogous SAT college readiness benchmarks linked to the PARCC assessments. Thus, communication between the two consortia with respect to the types of CCR benchmarks or international benchmarks being considered would foster consistency.

**Creating Consistency**

If consistency cannot be fostered through cooperation (perhaps the two consortia have very different timelines for ALD development), then consistency between the ALDs can be induced once the PARCC ALDs are available. In this case, CTB, at the request of SBAC, can compare and contrast the KSAs in the PARCC and SBAC ALDs for all grades and content areas. We can parse out the KSAs in each set of ALDs and crosswalk each KSA to where it appears (or is implied) in the other set of ALDs. In this way, we can create a list of major modifications and/or minor modifications that would be needed to the SBAC ALDs to foster consistency with the PARCC ALDs.

**2.16 Gather validity evidence for the usefulness of the initial ALDs to be included in the final technical report.**

SBAC requires that the winning vendor gather validity evidence for the usefulness of the initial ALDs and that this validity information be included in the final technical report. ALDs are meant to define the intended inferences and measured constructs about students, based on test scores. As such they are one of the ultimate products of an assessment program (Ferrara & DeMauro, 2006). SBAC desires ALDs for three main uses, including (a) item and test development, (b) standard setting guidance, and (c) score interpretation.

CTB’s ALD framework was established because the three uses of ALDs require different levels of specificity targeted to different users. The framework also incorporates common ALD development practices such as increasing the rigor of the cognitive level language in the descriptor as performance levels increase. In the next section we discuss issues that can affect the accuracy and therefore the usefulness of ALDs. This discussion is based upon the most up-to-date information in the field. We then discuss ways to collect a priori usefulness information per the RFP requirements. We solicit feedback evidence from stakeholders most directly targeted by the purpose; however, should SBAC desire all stakeholders evaluate the usefulness of the different ALDs types, CTB can also meet this request. We note such validity studies are a priori validity evidence that will be used to establish procedural validity evidence. Until items are developed and student data are collected, ALD accuracy and therefore their usefulness for the three different purposes cannot be empirically substantiated.

**Item and Test Development**

Range ALDs specifically denote the expected observable knowledge and skill development across performance levels. Clearly articulating the development of specific skills across achievement levels is what “sets the stage” for measuring student growth and makes ALDs consistent with reasonable learning gains. When finite skill development is not explicated in such a way, item writers may not write items targeting each stage of the skill progression, which diminishes the assessment’s instructional sensitivity. Bejar, Braun & Tennenbaum (2007) describe this process as developing a matrix in which the differences in expectation for a specific learning target are defined for each achievement level. Because SBAC plans to implement a computer adaptive test, it will be able to make full use of such an approach.

Targeting items to specific ALDs, while consistent with an evidence-centered design approach, is still in its infancy. Schneider, Huff, Tully, and Egan (2011) investigated which type of ALD (range or target) was the most useful for item development, and they found empirical support for the usefulness of range ALDs. CTB is excited to about the prospect of developing range ALDs with SBAC that can be used to
guide item writing. This work will provide how the field writes items and constructs tests.

**Usefulness Investigations**

One method to provide confirming evidence of the utility of the Grade 11 ALDs is to examine the extent to which the KSAs of the Grade 11 ALDs are consistent with the descriptions of what students know and can do are associated with score ranges on the SAT and PSAT/NMSQT. College Board has already undertaken a scale anchoring procedure, so this information is readily available.

In addition, we can ask multiple sources to evaluate the usefulness of the range ALDs for item writing. This evaluation can be conducted during the evaluation cycle that includes feedback from the following sources:

1. Contractors for Item/Task Specifications (#04),
2. Item/Task Writing/Review – Pilot (#14),
3. Test Design and CAT Specifications (#09),
4. Accessibility and Accommodations Policy Guidelines (#06),
5. Psychometric Services (#05), and Report Development (#15).

The evaluation will be administered using an online survey tool. The evaluation will ask several questions where panelists can respond on an agree/disagree continuum regarding the utility of the range ALDs for item writing. If possible, the evaluation could present test items and ask respondents to place those items into the appropriate range ALD. High degrees of agreement between the respondents regarding the range ALD placement of the item would be an indication of the utility of the range ALDs.

**Standard Setting Guidance**

Because target ALDs will specify SBAC’s expectations for students at the threshold of the achievement level, they define both the state’s policy and content-based expectations (that is, what it means to be Proficient). Bourque (2000) asserted that the “most important function” of the ALD “is to provide a mental framework or structure for standard setting panelists (p. 8)”.

In short, target ALDs are the heart of the standard setting process. Researchers (Hurtz & Auerbach, 2003; Giraud, Impara, & Plake, 2005; Impara, Giraud, & Plake, 2000) found when panelists used ALDs to guide their judgments, the variability of panelists ratings decreased. In a meta-analysis of 113 groups of judges, Hurtz and Auerbach (2005) concluded that “…judges tend to reach a relatively higher degree of consensus, and they tend to reach this consensus at a higher cutoff score” (p. 595) when they focus on a common ALD.

**Usefulness Investigations**

In order for target ALDs to be useful for standard setting participants they need to define expectations for students located at the cut score, and participants who develop them should be able to clearly articulate this purpose. In order to be useful, target ALDs also have to represent the expectations from the SBAC’s Consortium State Members. To that end, CTB will collect evaluations of the usefulness of the target ALDs for standard setting by collecting feedback from the some or all of the following sources:

1. Draft 1 Committee Members
2. CONSORTIUM Executive Committee.
3. CONSORTIUM higher education state membership.
4. The CONSORTIUM’s TAC.
5. CONSORTIUM K-12 state membership.
6. CONSORTIUM Executive Committee.
7. Lead Psychometrician and EC Liaison.
The evaluation will be administered using an online survey tool. The evaluation will ask several questions where panelists can respond on an agree/disagree continuum regarding the utility of the target ALDs for standard setting. If possible, the evaluation could present vignettes of students and ask respondents to place those students according to the target ALD. High degrees of agreement between the respondents regarding the target ALD placement of the vignettes would be an indication of the utility of the target ALDs.

The Consortium may want to consider asking the standard setting panelists about the usefulness of the target ALDs while at the standard setting workshop. These panelists will work closely with the target ALDs and will be best informed to address their usefulness. CTB recognizes that this type of study of standard setting panelists would be done by the vendor who conducts the standard setting workshop and is not part of the current RFP.

Score Interpretation

Once final cut scores have been recommended, it is necessary to review the target ALDs based on the KSAs students demonstrated on the test given the final cut scores. Target ALDs reflect expectations for what students in each achievement level should be able to do, and once cut scores are set, the original target student descriptors may no longer represent what students actually can do based upon evidence from the test items. Given that SBAC is interested in benchmarking its standards to other national and international assessments, this area requires particular attention. Given that the final ALDs are noted by the RFP as being handled in a separate standard setting contract, CTB at this time is planning not to collect usefulness evaluations for this intended ALD purpose. Should SBAC desire CTB to do so a priori, we will address the evaluation for this context while collecting data on the other two uses.

Convergent (Divergent) Evidence using Other Descriptors of College & Career Readiness

As discussed in Section 2.7, the College Board is well positioned to assist SBAC with defining and/or validating definitions of CCR. The College Board has undertaken a scale anchoring procedure, where descriptions of what students know and can do are associated with score ranges on the SAT and PSAT/NMSQT. To the extent that the scale anchoring statements are consistent with the KSAs identified with each ALD, they can be used to provide confirming evidence of the utility of the ALDs.

2.17 Draft a final technical report describing the process used for developing initial ALDs, the deliverables produced, and the validity evidence they provide in support of the CONSORTIUM’s goals.

The final technical report will be based on the draft technical report that is described in Section 2.13. Like the draft technical report, the final technical report will describe the methodological, conceptual, and logistical elements of the five-day ALD development process (please see Section 2.13 for a thorough discussion of procedural validity). In addition, we will capture the iterative feedback that occurs following the workshop and through the reporting of ALD usefulness data. The final technical report will incorporate the process and feedback activities that are a component of the RFP after the workshop. In this section, we propose the structure of the final technical report.

Proposed Structure of the Final Technical Report

Like the draft technical report, the final technical report is an important means for SBAC to present information regarding the initial ALDs to various groups of SBAC stakeholders. In addition, the final technical report will show various review cycles undertaken prior to finalizing the initial ALDs. Like the draft technical report, the final technical report also presents argumentation and related evidence for external reviewers and assessment professionals to judge the extent to which the intended assessment uses of initial ALDs are justified.

Prior to the ALD workshop, CTB will document the planned ALD development activities through a rigorous planning document that describes each activity. This meticulous planning will provide the foundation for the comprehensive and coherent technical document that SBAC will use as a central
source of validity evidence for its initial ALDs.

We have successfully supported many other states as they successfully completed the peer review process. One of the key components of our customers' overwhelming success in peer review has come directly from our rigorous technical documentation. Our technical reports will include specific links to appropriate evidence within the Standards for Educational and Psychological Testing (AERA, APA, and NCME) and the Peer Review Guidance for ESEA.

One important structural change in the proposed updates for the Standards is the categorization of standards into three sections: Foundations (e.g., reliability, validity, test fairness), Operations (e.g., test design and development, administration, scoring and reporting), and Applications (e.g., tests as educational assessment). This categorization happens to align with three critical elements that are present in argument-based approaches such as the evidenced-centered assessment design. These elements are qualities of score-based inferences, assessment activities to produce and distribute scores, and uses of the scores. Accordingly, we propose to organize the final technical report into three major sections: Section I will provide information about the applications of the initial ALDs; Section II describes key operations, including the workshop design; and Section III presents evaluative information, including validity evidence.

Table 14 shows the proposed outline for the final technical report. CTB is happy to adjust the design to incorporate feedback or requirements from the Smarter Balanced or its TAC.

**Table 14. Outline of the Final Technical Report**

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Section Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>Goal of the Executive Summary: To present key information and highlights of the draft technical report.</td>
</tr>
<tr>
<td><strong>SECTION I: APPLICATIONS</strong></td>
<td>Goals of Section I (Chapter 1): To introduce the initial ALDs and their intended uses.</td>
</tr>
<tr>
<td>Chapter 1. Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1 Intended Uses of the Initial ALDs</td>
<td></td>
</tr>
<tr>
<td>1.2 Target Population for Using Initial ALDs</td>
<td></td>
</tr>
<tr>
<td><strong>SECTION II: OPERATIONS</strong></td>
<td>Goals of Section II (Chapters 2-6): To describe major activities in operations.</td>
</tr>
<tr>
<td>Chapter 2. Panelists</td>
<td></td>
</tr>
<tr>
<td>2.1 Panelist Recruitment</td>
<td></td>
</tr>
<tr>
<td>2.2 Panelist Demographics</td>
<td></td>
</tr>
<tr>
<td>Chapter 3. Workshop Design</td>
<td></td>
</tr>
<tr>
<td>3.1 Description of Workshop Goal</td>
<td></td>
</tr>
<tr>
<td>3.2 Description of General Policy ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.3 Description of Grade- and Content-Specific Range and Target ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.4 Quality Control Evidence</td>
<td></td>
</tr>
<tr>
<td>Chapter 4. Review Cycles</td>
<td></td>
</tr>
<tr>
<td>4.1 Contractor and Work Groups Review</td>
<td></td>
</tr>
<tr>
<td>4.2 Consortium Executive Committee and Consortium Higher Education State Membership Review</td>
<td></td>
</tr>
<tr>
<td>4.3 Consortium TAC and Consortium K-12 State Membership Review</td>
<td></td>
</tr>
<tr>
<td>4.4 Consistency with PARCC</td>
<td></td>
</tr>
<tr>
<td>4.5 Quality Control Evidence</td>
<td></td>
</tr>
<tr>
<td>Chapter 5. General Policy Descriptors</td>
<td></td>
</tr>
<tr>
<td>Chapter 6. Range and Target Achievement Level Descriptors</td>
<td></td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

SECTION III: EVALUATION

Chapter 7. Validity
7.1 Types of Validity Evidence
7.1.a Procedural Validity
7.1.b Panelist Feedback
7.2 Usefulness of Initial ALDs
7.3 Summary of Validity Evidence for the Intended Test Uses

Appendix A. Training Materials
Appendix B. Results of Evaluations
Appendix C. Diagram of Review Cycles

A. Project Approach/Methodology

Part 1: Oversight, Panelist Recruitment, and Methodological Design

Oversight

CTB will provide coordination of all schedules and activities under the resulting contract(s) as well as ensuring communication related to the tasks within each part of the work and with other Contractors and Smarter Balanced leadership to accomplish all of the work.

CTB will maintain full responsibility for complete program oversight, including the maintenance of timelines and critical program deliverables, complete program communication across other contractors and SBAC leadership, and consolidation of all work completed on the contract. As detailed in this Proposal’s Management Plan, we have assigned a senior-level management team with the skills and experience necessary to maintain program oversight and effectively coordinate all program activities. The complexity and accelerated timeline of the program will require the team to build consensus quickly and move decisively in order to maintain the ongoing movement required to recruit panelists, design and implement the workshop, implement the review committees, and deliver initial ALDs and Technical Reports within the required timelines. It will require a “one-department” mindset to develop quickly across the SBAC and our alliance organizations in order to deliver all that is required.

The team, under the leadership of Program Manager Lindy Wienand, will maintain a full program schedule and complete program documentation. A Master Program Schedule will be created to assure identification, organization, and sequencing of all project tasks, deliverables, and milestones. The project schedule will take the key elements of the project and translate them into a time-based plan. The complete schedule will include a work breakdown structure, all tasks and activities associated with the project, and the interdependencies of the tasks to be performed. The Project Schedule will be created using CTB’s project management scheduling software, and will be continuously monitored, updated, and analyzed by the Program Schedule Analyst, Kristal Gill.

The team will provide comprehensive program documentation using a number of management tools, including a Program Work Plan. This top-level program-planning document provides the details of the program scope and defined roles and responsibilities of all participants to assure everyone is able to be productive and effective right from the start. The Program Work Plan provides the answers to the who, what, when, and how, questions related to key activities, milestones, deliverables, timeline, resources, risks, program controls, and quality controls. This document will become the central control document for all teams working on the program, and the basis for change management. The work plans for each program part align to the overarching Program Work Plan, and will define at a greater level of detail the...
tasks, services and activities to accomplish the scope of each part.

**Panelist Recruitment**

In this section, we present our plan for panelist recruitment. Due to the number of desired reviews, the panelists who will develop the first draft of the initial ALDs must be recruited very quickly. To accomplish this, Dr. Karla Egan will work closely with Consortium representatives to create the sampling plan and to determine what information is relevant to Smarter Balanced in addition to the information in this report (see Table 15 for the known stakeholder groups). CTB suggests that CTB and Consortium representatives work together to refine this design to include other relevant characteristics, like demographic information or teaching experience.

In the interest of time, CTB suggests that the pool of nominees come from prior recruiting work that the Smarter Balanced Assessment Consortium has done, such as under the work of SBAC 08 and SBAC 14. CTB can quickly examine the nominee information from previous pools to assure that all governing states are included. If available, we can also check the qualifications of the nominees. Qualified but unselected nominees from prior recruiting work can be contacted to ascertain their interest in participating in the ALD development. If a governing state does not have nominees in the pool, then CTB will work with Consortium representatives to identify five potential nominees from their governing states. If SBAC prefers, then CTB can work with Consortium representatives to create a new pool of nominees for the purposes of creating initial ALDs; however, this may add additional time to the schedule.

Once a pool of potential nominees has been selected, CTB will create a matrix that shows the grade-level experience, content-area experience, and CCSS experience of each nominee. CTB will work with SBAC to include other desired information in the matrix, which might include demographic information or types of teaching experience. We will use this information to find the most qualified nominee from each governing state. A list of nominees will be created by CTB and will be approved by the Consortium. Once approved by the Consortium, the nominee will be contacted to participate in the ALD-development workshop (see Section 2.7).

**Groupings**

The panelists will work in small groups that span two grades, except for grade 11. The proposed grade groupings will be:

- Grades 3-4
- Grades 5-6
- Grades 7-8
- Grade 11.

Within each content area, each grade span will have four panelists, except for Grade 7-8 which will have five panelists.

Table 15 shows the disaggregation of panelists into each grade and grade-span group. As shown in Table 15, one grade-level expert from a governing state will be in his/her appropriate grade; one content expert will in each grade span group; and one expert in CCSS will be in each grade-span group. A representative of higher education will participate in the grade 7-8 group and the grade 11 group.
Table 15. Panelists by Stakeholder Group & Committee Role

<table>
<thead>
<tr>
<th>Stakeholder group/committee role</th>
<th>Grade level/span</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade-level experts from governing states</td>
<td>3  4  5  6  7  8 11</td>
</tr>
<tr>
<td>Content experts from governing states</td>
<td>1  1  1  1  1  1</td>
</tr>
<tr>
<td>CCSS content experts</td>
<td>1  1  1  1</td>
</tr>
<tr>
<td>IHE rep</td>
<td>4  4  5  4</td>
</tr>
<tr>
<td>Total grade band</td>
<td>11</td>
</tr>
<tr>
<td>Total governing</td>
<td>6</td>
</tr>
<tr>
<td>Total other</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 16 shows the ALD types within CTB’s framework along with their uses and how they align with SBAC’s request for initial descriptors. Within the context of the ALD framework, general ALDs are equivalent to policy ALDs and grade-specific ALDs encompass range and target ALDs.

Table 16. ALD types within CTB’s framework and how the ALDs align with the Consortiums request

<table>
<thead>
<tr>
<th>Smarter Balanced Initial ALD</th>
<th>Framework ALD</th>
<th>Purpose/Use of ALD</th>
</tr>
</thead>
<tbody>
<tr>
<td>General ALD</td>
<td>Policy ALD</td>
<td>Articulate the goals and rigor for the final performance standards, and set the tone for and are embedded within subsequent descriptors.</td>
</tr>
<tr>
<td>Grade-Specific ALDs</td>
<td>Range ALDs</td>
<td>Guide test development and item writing by identifying which aspects of items align with a particular performance level in regard to the cognitive and content rigor that has been defined.</td>
</tr>
<tr>
<td></td>
<td>Target ALDs</td>
<td>Guide the standard setting by defining the minimum performance required to be considered as meeting the achievement level expectations.</td>
</tr>
</tbody>
</table>

This system of four linked interrelated ALDs explicates the construct being measured and supports the intended test score interpretations in regard to that construct. In other words, the framework supports an evidence-centered design approach. Although the framework ALDs are discussed as four interrelated ALDs, they also can be conceptualized as the evolution that an ALD must undergo from its inception to completion because the ALD development process is iterative (Plake, Huff, & Reshetar, 2010). These ALDs define the construct that is being measured and describe what students should know.
and be able to do in relation to the construct–student achievement relative to the common core standards.

**General Policy ALDs.** The first step of this methodology is to create the general policy ALDs that will set the tone for the grade-specific range and target ALDs. CTB proposes that the work for developing the general ALDs begins with the overall claims for the summative assessments. The overall claims are a sort of policy guidance that shows the expectations that SBAC has for its students. Using the overall claims along with the claims associated with particular aspects of the common core state standards, the panelists can discuss the expectations of the students in each of the four achievement levels in terms of these claims.

Table 17 shows overall grade 11 Claim for the summative Mathematics assessment. Much like the policy ALDs from the framework, these claims show expectations that SBAC has set forth for students. The panelists can begin with these claims and show the progress of these claims for students in Levels 1, 2, 3, and 4. Using just the overall claim as an example, the panelists may begin by stating that grade 11, Level 3, students demonstrate CCR in mathematics. From here, they will move up the proficiency continuum to discuss the general claim for the Level 4 students. They will then complete their work by discussing their expectations for the Level 1 and Level 2 student.

**Table 17. Example of how SBAC Claims may be scaffolded into Achievement Levels**

<table>
<thead>
<tr>
<th>Overall Claim for Grade 11</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 1 students require intensive support to develop basic math skills. They require substantial remediation in skills considered prerequisite to college mathematics.</td>
<td>Level 2 students require remedial coursework in mathematics and are currently unprepared for entry level college mathematics.</td>
<td>Level 3 students demonstrate college and career readiness in mathematics. These students are prepared for entry-level college mathematics without the need for remediation.</td>
<td>Level 4 students demonstrate advanced CCR in mathematics. These students are prepared for more advanced mathematics coursework that might be expected of students after their first year of college mathematics.</td>
</tr>
</tbody>
</table>

**Grade-Specific Range and Target ALDs.** The second step in this methodology is to create the grade-specific range and target ALDs. To do this, we propose to have panelists study the general policy ALDs and then go through a procedure where the annotate the CCSS and the content specifications. The panelists will first analyze the cognitive expectations, content information, and skills found in the CCSS. Figure 3 shows an example of this process using a cluster from the Ratios and Proportional Relationship domain of the seventh-grade mathematics CCSS. For the purposes of this description, we use the achievement level labels “Proficient” and “Basic.” Using this methodology to parse the CCSS, the skills can be delineated into just Proficient (P-), average Proficient (P), or highly Proficient (P+), using the annotation process shown in Figure 3. A similar delineation should be made for the other performance levels. For example, Figure X shows panelists have indicated that recognizing proportional relationships is a skill expected of the just Basic examinee (B-), whereas representing proportional relationships is a skill expected of the average Basic examinee (B).

Once panelists have annotated the CCSS and content specifications, they will compile the KSAs into range and target ALDs. The range ALDs encompass all skills indicated within an achievement level (e.g., P-, P, P+) while the target ALDs are represented by those skills necessary to just enter an achievement level (e.g., B-, P-, etc.). The end results will be a matrix that shows the delineation of KSAs into range ALDs, which include all of the content and skills assigned to a performance level. The range ALDs
should eventually resemble the matrix presented in Table 18. (A separate matrix can be compiled for those KSAs classified in the just Proficient (P-) category. These KSAs constitute the target ALD for Proficient.) The KSAs of both the target and range ALDs will be examined for fluency across the grade levels within a content area. The KSAs will be compiled into a single set of ALDs for each grade level and content area.

**Figure 3. Parsed CCSS to Create Range and Target ALDs**

**Ratios and Proportional Relationships 7.RP**

Analyze proportional relationships and use them to solve real-world and mathematical problems.

3. Compute unit rates associated with ratios of fractions, including ratios of lengths, areas and other quantities measured in like or different units. *For example, if a person walks 1/2 mile in each 1/4 hour, compute the unit rate as the complex fraction 1/2/1/4 miles per hour, equivalently 2 miles per hour.*

4. Recognize and represent proportional relationships between quantities.

   *b.* Decide whether two quantities are in a proportional relationship, e.g., by testing for equivalent ratios in a table or graphing on a coordinate plane and observing whether the graph is a straight line through the origin.

Skills represented by **P**- are compiled for the Target Proficient PLD. *The Target Proficient student is assumed to possess the KSPs assigned to the Basic category.*

**Table 18. Example of Portion of Range ALD for Grade 8 Mathematics**

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determine perimeter of regular figure when no context is present</td>
<td>Determine perimeter of regular figure when context is present</td>
<td>Given perimeter, determine the missing measurement of a regular figure when context is present</td>
<td></td>
</tr>
</tbody>
</table>

© 2012 CTB/McGraw-Hill LLC (Unpublished)
### Level 1
- Determine perimeter of irregular figure when no context is present
- Find area of regular figure
- Find volume of regular figure
- Use data in a circle graph to determine if statements are true
- Determine the mode of a set of data

### Level 2
- Determine perimeter of irregular figure when context is present
- Find area of irregular figure
- Find volume of irregular figure
- Use data in a single bar graph to determine if statements are true or false
- Determine the range of a set of data

### Level 3
- Given perimeter, determine the missing measurement of an irregular figure when context is present
- Find missing measurement of regular figure given area when context is present
- Find missing measurement of irregular figure given volume when context is present
- Use data in a line graph to determine if statements are true
- Determine the mean and median of a set of five 2-digit numbers presented in a table of data

### Level 4
- Find missing measurement of irregular figure given area when context is present
- Find missing measurement of irregular figure given volume when context is present
- Use data in a line graph to determine if statements are true
- Determine the mean and median of a set of five 2-digit numbers presented in a table of data

---

*From Egan, Schneider, & Ferrara, 2012*

### Challenges to Creating Initial ALDs

No matter what the methodology utilized, SBAC will encounter several challenges while creating these ALDs. First, general policy ALDs are typically developed and approved by policy makers before any committee work. The timelines of this work necessitate that the general policy ALDs are created in tandem with the grade-specific range and target ALDs. This means that the general policy ALDs will be reviewed at the same time as the grade-specific ALDs. If changes are made to the general policy ALDs, then it will likely affect the grade-specific ALDs. If SBAC can allow more time, then we recommend that policy ALDs be created and approved prior to writing grade-specific ALDs.

Next, the general claim of “college and career readiness” may present a challenge to panelists as they begin to parse out the claims for each achievement level. The meaning of CCR as defined by SBAC policymakers will provide important guidance to panelists as they create the grade-specific content based ALDS. CTB recommends that SBAC define its policy meaning of CCR prior to convening the content panels. If panelists are to define CCR, then College Board is uniquely positioned to assist SBAC in defining CCR with data prior to the ALD development process. For example, College Board may provide external data and to conduct empirical analyses to aid in establishing ALDs.

Third, the CCSS encompasses more standards than will be included in the summative assessment. Since the RFP states that the “initial ALDs are germane to judgments made about performance levels on the summative assessments…” (p. 15), we are recommending that panelist discussion focus on the parts of the CCSS to those that have been deemed to be a target of the summative assessment in the content specifications. If SBAC would prefer that the achievement levels be based on all cluster and standards of the CCSS, then the proposed methodology can easily accommodate this.
Fourth, the relationship between the initial ALDs and the external benchmarks that may be used at standard setting is unknown. If external benchmarks are to drive the standard setting methodology, then it is worth considering how these same external benchmarks may be used to inform the initial range and target ALDS.

Finally, it is desired that the ALDs that results from the ALD-development process demonstrate “consistency with learning gains achievable through reasonable effective instruction during a school year for a typical student.” This is not a simple task, and though we will direct panelists to consider learning gains during the workshop, we believe that it will be necessary to conduct a review of the literature on learning gains for ELA and mathematics. CTB proposes that the panelists representing higher education also be experts in learning gains.

**Part II: Workshop Implementation, Draft Initial ALDs, and Draft Technical Report**

**Workshop Implementation**

In this section, we overview the workshop, the committee types, and the workshop agenda. Table 19 overviews the activities of the five-day workshop. At a high-level, a meta-committee will develop the general policy ALDs. Next, the grade 11 committees will develop the range and target ALDs for mathematics and ELA. Meta-committee members will participate in the creating the range and target ALDs for grade 11 by providing policy guidance and clarification of intent as necessary to the content-based committee. Finally, the grades 3 through 8 committees will develop range and target ALDs for mathematics and ELA for each grade level working backwards from the Grade 11 range and target ALDs.

This will be a five-day workshop for meta-committee members. For all other panelists, this will be a two-day activity. The grade 11 range and target ALDs will be completed before the grades 3-8 ALDs are written so that panelists may use the grade 11 ALDs to work backward to the KSAs that are necessary in their own grades to meet the grade 11 expectations.

**Table 19. Overview of five-day workshop**

<table>
<thead>
<tr>
<th>Day(s)</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Meta-Committee creates General Policy ALDs</td>
</tr>
<tr>
<td>2 - 3</td>
<td>Grade 11 creates range and target ALDs</td>
</tr>
<tr>
<td>4 - 5</td>
<td>Grades 3 – 8 create range and target ALDs</td>
</tr>
</tbody>
</table>

**Committees**

Within each content area, two types of committees will engage in the ALD-development workshops: a meta-committee and grade-span committees. The meta-committee will comprise one representative from each grade-span group, for a total of four members within each content area. The members of this committee will also be part of the grade-span committees within their assigned content areas. The meta-committee members will assist with their own grade-span as well as helping with work of other grade-spans when appropriate. The members of this committee will create the general policy ALDs, and they will assist with the cross-grade articulation of descriptors.

Within each content area, four grade-span committees will meet to create ALDs for each grade level. Table 20 shows the grade spans and the number of committee members. These committees will be responsible for creating the grade- and content-specific range and target ALDs.
Table 20. Number of Members by Committee Type for a Single Content Area

<table>
<thead>
<tr>
<th>Grade-Span Committees</th>
<th>Meta-Committee</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 – 4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5 – 6</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7 – 8</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Workshop Agenda

Table 21 shows the proposed schedule of the Day 1-5 activities that CTB proposes for the workshop. This agenda provides multiple rounds for the creation of each ALD type. It provides for panelist discussion. The Grade 11 grade-specific range and target ALDs will be created prior to the Grades 3-8 ALDs so that Grades 3-8 panelists can work backward from Grade 11 to enhance alignment between the expectations of CCR in Grade 11 and progress towards CCR in Grades 3-8.

Table 21. Proposed Schedule of Day 1-5 Activities

<table>
<thead>
<tr>
<th>Day &amp; Committee</th>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>8:00 – 9:00 am</td>
<td>Opening Session</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Training on general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Orientation to materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>9:00 – 10:30 am</td>
<td>Round 1. Overall SBAC Claims</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study and discuss overall SBAC claims</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study examples of general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss college and career readiness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Scaffold overall claims into achievement levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td>Break</td>
</tr>
<tr>
<td></td>
<td>10:45 am – 12:00 pm</td>
<td>Round 2. Overall SBAC Claims</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study and discuss Round 1 general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise as needed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss defining phrases that differentiate achievement levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td>Lunch</td>
</tr>
<tr>
<td></td>
<td>1:00 – 2:45 pm</td>
<td>Round 3. Content-based SBAC Claims</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study and discuss content-specific SBAC claims</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Scaffold content-specific claims into achievement levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>2:45 – 3:00 pm</td>
<td>Break</td>
</tr>
<tr>
<td>Day &amp; Committee</td>
<td>Time</td>
<td>Activities</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>3:00 – 4:45 pm</td>
<td><strong>Round 4. Content-based SBAC Claims</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study and discuss Round 3 content-specific general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Cross-content discussion of content-specific general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise as needed</td>
</tr>
<tr>
<td></td>
<td>4:45 – 5:00 pm</td>
<td><strong>Evaluation</strong></td>
</tr>
<tr>
<td></td>
<td>5:00 pm</td>
<td><strong>Dismissal</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Day 2</th>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8:00 – 9:00 am</td>
<td><strong>Opening Session</strong></td>
</tr>
<tr>
<td>Meta-committee &amp; Grade 11 committee</td>
<td></td>
<td>• Training on general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Orientation to materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>9:00 am – 12:00 pm</td>
<td><strong>Round 1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study and discuss general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study example matrices of knowledge, skills, and abilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss and Parse Common Core Content Standards</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss and Parse content specifications</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td><strong>Lunch</strong></td>
</tr>
<tr>
<td></td>
<td>1:00 – 2:45 pm</td>
<td><strong>Continue Round 1</strong></td>
</tr>
<tr>
<td></td>
<td>2:45 – 3:00 pm</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>3:00 – 5:00 pm</td>
<td><strong>Complete Round 1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Compile into ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>5:00 pm</td>
<td><strong>Dismissal</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Day 3</th>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8:00 am – 12:00 pm</td>
<td><strong>Round 2.</strong></td>
</tr>
<tr>
<td>Meta-committee &amp; Grade 11 committee</td>
<td></td>
<td>• Study cluster emphases and CAT emphases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise descriptors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td><strong>Lunch and Dismiss</strong></td>
</tr>
<tr>
<td>Day &amp; Committee</td>
<td>Time</td>
<td>Activities</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Day 4</td>
<td>8:00 – 9:00 am</td>
<td><strong>Opening Session</strong></td>
</tr>
<tr>
<td>Meta-committee &amp; Grades 3 - 8 committee</td>
<td></td>
<td>• Training on general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Orientation to materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>9:00 am – 12:00 pm</td>
<td><strong>Round 1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study and discuss general policy ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study example matrices of knowledge, skills, and abilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss and Parse Common Core Content Standards</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discuss and Parse content specifications</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td><strong>Lunch</strong></td>
</tr>
<tr>
<td></td>
<td>1:00 – 2:45 pm</td>
<td><strong>Continue Round 1</strong></td>
</tr>
<tr>
<td></td>
<td>2:45 – 3:00 pm</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>3:00 – 5:00 pm</td>
<td><strong>Complete Round 1</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Compile into ALDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>5:00 pm</td>
<td><strong>Dismissal</strong></td>
</tr>
<tr>
<td>Day 5</td>
<td>8:00 am – 12:00 pm</td>
<td><strong>Round 2</strong></td>
</tr>
<tr>
<td>Meta-committee &amp; Grades 3 - 8 committee</td>
<td></td>
<td>• Cross-grade discussions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise as needed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>10:30 – 10:45 am</td>
<td><strong>Break</strong></td>
</tr>
<tr>
<td></td>
<td>12:00 – 1:00 pm</td>
<td><strong>Lunch</strong></td>
</tr>
<tr>
<td></td>
<td>1:00 – 2:45 pm</td>
<td><strong>Round 3</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Study cluster emphases and CAT emphases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revise descriptors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td></td>
<td>2:45 – 3:00 pm</td>
<td><strong>Break</strong></td>
</tr>
</tbody>
</table>
### Day & Committee

<table>
<thead>
<tr>
<th>Time</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>3:00 – 5:00 pm</td>
<td>Meta-Committee</td>
</tr>
<tr>
<td></td>
<td>- Revise descriptors</td>
</tr>
<tr>
<td></td>
<td>- Evaluation</td>
</tr>
<tr>
<td>5:00 pm</td>
<td>Dismissal</td>
</tr>
</tbody>
</table>

**Evaluations.** CTB recognizes that panelist evaluations are an important piece of evidence of procedural validity, and we plan to survey panelists after each round of discussion and on important aspects of the workshop. The survey will ask panelists for their feedback on the utility of the ALDs, the cross-grade articulation of KSAs within each content area’s ALDs; the validity of the method; as well as many other logistical, conceptual, and methodology elements of the workshop.

### Draft Initial ALDs

Panelists will create the first version of the draft initial ALDs. Content Experts from CTB and College Board with deep knowledge of CCSS will edit those ALDs for consistency and style prior to the first review cycle. Our content experts have extensive experience in editing draft ALDs so that they remain true to the original intent of the workshop panelists. As our content experts edit the draft initial ALDs, they will pay particular attention to:

1. Consistency with the Common Core State Standards.
2. Consistency with the Consortium’s content specifications in ELA and mathematics (see Exhibits I and J).
3. Consistency with skills and knowledge students will be able to demonstrate on the Consortium’s computer adaptive summative assessments and performance tasks, as described by the Consortium’s content specifications.
4. Consistency with the scores and reporting information to be produced by summative computer adaptive tests and performance tasks, as described by the Consortium’s content specifications.
5. Consistency across grade levels, with logical developmental progression from one grade level to the next.
6. Consistency with expectations for performance at each grade that will be necessary for students to be college and career ready by the end of high school (i.e., working backward from college and career readiness at the end of high school, students’ knowledge and skills should build sequentially to ensure steady annual progress).
7. Consistency with learning gains achievable through reasonable effective instruction during a school year for a typical student demonstrating achievement of the ALDs.
8. A balance of conciseness and level of detail that will support item and task development, test design, and standard setting.
9. Plain language that is easily understood by non-technical stakeholders such as parents and policy makers.

### Draft Technical Report

Following the implementation of the workshop, CTB will create the draft technical report which will summarize the activities of panelist recruitment and ALD development. This draft technical report will describe the methodological, conceptual, and logistical elements of the five-day ALD development process. The final technical report will incorporate the process and feedback activities that are a component of the RFP after the workshop.
The draft technical report is an important means for SBAC to present information regarding the initial ALDs to various groups of SBAC stakeholders. The technical report also presents argumentation and related evidence for external reviewers and assessment professionals to judge the extent to which the intended assessment uses of initial ALDs are justified. The technical report is an important component of establishing the procedural validity of the ALD development.

Table 22 shows the proposed outline for the draft technical report. CTB is happy to adjust the design to incorporate feedback or requirements from the Consortium or its TAC.

**Table 22. Outline of the Technical Report**

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Section Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>Goal of the Executive Summary: To present key information and highlights of the draft technical report.</td>
</tr>
<tr>
<td>SECTION I: APPLICATIONS</td>
<td>Goals of Section I (Chapter 1): To introduce the initial ALDs and their intended uses.</td>
</tr>
<tr>
<td>Chapter 1. Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1 Intended Uses of the Initial ALDs</td>
<td></td>
</tr>
<tr>
<td>1.2 Target Population for Using Initial ALDs</td>
<td></td>
</tr>
<tr>
<td>SECTION II: OPERATIONS</td>
<td>Goals of Section II (Chapters 2-6): To describe major activities in operations.</td>
</tr>
<tr>
<td>Chapter 2. Panelists</td>
<td></td>
</tr>
<tr>
<td>2.1 Panelist Recruitment</td>
<td></td>
</tr>
<tr>
<td>2.2 Panelist Demographics</td>
<td></td>
</tr>
<tr>
<td>Chapter 3. Workshop Design</td>
<td></td>
</tr>
<tr>
<td>3.1 Description of Workshop Goal</td>
<td></td>
</tr>
<tr>
<td>3.2 Description of General Policy ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.3 Description of Grade- and Content-Specific Range and Target ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.4 Quality Control Evidence</td>
<td></td>
</tr>
<tr>
<td>Chapter 4. General Policy Descriptors</td>
<td></td>
</tr>
<tr>
<td>Chapter 5. Range and Target Achievement Level Descriptors</td>
<td></td>
</tr>
<tr>
<td>SECTION III: EVALUATION</td>
<td>Goals of Section III (Chapter 6): To present critical info. for evaluating qualities of the initial ALDs.</td>
</tr>
<tr>
<td>Chapter 6. Validity</td>
<td></td>
</tr>
<tr>
<td>6.1 Types of Validity Evidence</td>
<td></td>
</tr>
<tr>
<td>6.1.a Procedural Validity</td>
<td></td>
</tr>
<tr>
<td>6.1.b Panelist Feedback</td>
<td></td>
</tr>
<tr>
<td>6.2 Summary of Validity Evidence for the Intended Test Uses</td>
<td></td>
</tr>
<tr>
<td>Appendix A. Training Materials</td>
<td></td>
</tr>
<tr>
<td>Appendix B. Results of Evaluations</td>
<td></td>
</tr>
<tr>
<td>Appendix C. Planned Review Cycles</td>
<td></td>
</tr>
</tbody>
</table>

**Part III. Review Cycles and Validation Studies**

**Review Cycles**

CTB will manage multiple review cycles of the ALDs as described in Sec. 2.10. Review panels will be recruited to represent all stakeholder groups and each panel will engage in an online review of Versions 1-5 of the ALD documents. CTB will collect, evaluate and apply revision comments as described in Sec.
2.11 resulting in a final version of the ALD documents.

**Validation Studies**

CTB proposes three types of validity studies that investigate the usefulness of the ALDs. First, CTB will survey SBAC stakeholders to ascertain the utility of the ALDs for item writers. In the interest of time, CTB will ask about the usefulness of the ALDs during the review cycles. The following groups will be surveyed regarding the utility of the ALDs for item writing.

| 1. Contractors for Item/Task Specifications (#04), |
| 2. Item/Task Writing/Review – Pilot (#14), |
| 3. Test Design and CAT Specifications (#09), |
| 4. Accessibility and Accommodations Policy Guidelines (#06), |
| 5. Psychometric Services (#05), and Report Development (#15), |

Next, CTB will survey SBAC stakeholders to ascertain the utility of the ALDs for standard setters. The following groups will be surveyed regarding the utility of the ALDs for item writing.

| 1. Draft 1 Committee Members |
| 2. Consortium Executive Committee. |
| 3. Consortium higher education state membership. |
| 4. The Consortium’s TAC. |
| 5. Consortium K-12 state membership. |
| 7. Lead Psychometrician and EC Liaison |

All evaluations will be administered using an online survey tool. The evaluation will ask several questions where panelists can respond on an agree/disagree continuum regarding the utility of the target ALDs for standard setting.

Third, College Board will investigate the alignment of the KSAs within the Grade 11 range and target ALDs with the KSAs in its own CCR descriptors. This investigation will be completed assuming that College Board descriptors of CCR are not used to guide the development of the initial ALDs.

**Part IV. Final Technical Report and Final Initial ALDs**

**Final Technical Report**

Following the review cycles, validity studies, and creation of final initial ALDs, CTB will create the final technical report which will summarize the all the major activities of the contract. This final technical report will describe the methodological, conceptual, and logistical elements of the five-day ALD development process. The final technical report will incorporate the process and feedback activities that are a component of the RFP after the workshop. In addition, we will capture the iterative feedback that occurs following the workshop and through the reporting of ALD usefulness data. The final technical report will incorporate the process and feedback activities that are a component of the RFP after the workshop. In this section, we propose the structure of the final technical report. It will also incorporate any changes that the TAC indicated following its review of the draft technical report.

The final technical report is an important means for SBAC to present information regarding the initial ALDs to various groups of SBAC stakeholders. The technical report also presents argumentation and related evidence for external reviewers and assessment professionals to judge the extent to which the intended assessment uses of initial ALDs are justified. The technical report is an important component
of establishing the procedural validity of the ALD development.

Table 23 shows the proposed outline for the final technical report. CTB is happy to adjust the design to incorporate feedback or requirements from the SBAC or its TAC.

Table 23. Outline for final technical report

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Section Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>Goal of the Executive Summary: To present key information and highlights of the draft technical report.</td>
</tr>
<tr>
<td>SECTION I: APPLICATIONS</td>
<td>Goals of Section I (Chapter 1): To introduce the initial ALDs and their intended uses.</td>
</tr>
<tr>
<td>Chapter 1. Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1 Intended Uses of the Initial ALDs</td>
<td></td>
</tr>
<tr>
<td>1.2 Target Population for Using Initial ALDs</td>
<td></td>
</tr>
<tr>
<td>SECTION II: OPERATIONS</td>
<td>Goals of Section II (Chapters 2-6): To describe major activities in operations.</td>
</tr>
<tr>
<td>Chapter 2. Panelists</td>
<td></td>
</tr>
<tr>
<td>2.1 Panelist Recruitment</td>
<td></td>
</tr>
<tr>
<td>2.2 Panelist Demographics</td>
<td></td>
</tr>
<tr>
<td>Chapter 3. Workshop Design</td>
<td></td>
</tr>
<tr>
<td>3.1 Description of Workshop Goal</td>
<td></td>
</tr>
<tr>
<td>3.2 Description of General Policy ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.3 Description of Grade- and Content-Specific Range and Target ALD Workshop</td>
<td></td>
</tr>
<tr>
<td>3.4 Quality Control Evidence</td>
<td></td>
</tr>
<tr>
<td>Chapter 4. Review Cycles</td>
<td></td>
</tr>
<tr>
<td>4.1 Contractor and Work Groups Review</td>
<td></td>
</tr>
<tr>
<td>4.2 Consortium Executive Committee and Consortium Higher Education State Membership Review</td>
<td></td>
</tr>
<tr>
<td>4.3 Consortium TAC and Consortium K-12 State Membership Review</td>
<td></td>
</tr>
<tr>
<td>4.4 Consistency with PARCC</td>
<td></td>
</tr>
<tr>
<td>4.5 Quality Control Evidence</td>
<td></td>
</tr>
<tr>
<td>Chapter 5. General Policy Descriptors</td>
<td></td>
</tr>
<tr>
<td>Chapter 6. Range and Target Achievement Level Descriptors</td>
<td></td>
</tr>
<tr>
<td>SECTION III: EVALUATION</td>
<td>Goals of Section III (Chapter 7): To present critical info. for evaluating qualities of the initial ALDs.</td>
</tr>
<tr>
<td>Chapter 7. Validity</td>
<td></td>
</tr>
<tr>
<td>7.1 Types of Validity Evidence</td>
<td></td>
</tr>
<tr>
<td>7.1.a Procedural Validity</td>
<td></td>
</tr>
<tr>
<td>7.1.b Panelist Feedback</td>
<td></td>
</tr>
<tr>
<td>7.2 Usefulness of Initial ALDs</td>
<td></td>
</tr>
<tr>
<td>7.3 Summary of Validity Evidence for the Intended Test Uses</td>
<td></td>
</tr>
</tbody>
</table>

Appendix A. Training Materials
Appendix B. Results of Evaluations
Appendix C. Diagram of Review Cycles

Final Initial ALDs
After the completion of the Review cycles, Content Experts from CTB and College Board will complete the revision of the draft initial ALDs. They will review all revision comments and changes to ensure that
the edited versions of the ALDs remain true to the original intent of the workshop panelists. Final ALD drafts will be reviewed for:

- Consistency with original intent of the committee
- Consistency with the CCSS and CONSORTIUM’s content specifications
- Consistency across grade levels and college and career readiness expectations
- Appropriate level of detail and plain language

Final initial ALDs will undergo a final thorough editorial style and proofing review.

**B. Work Plan and Schedule**

We have created an integrated work plan that describes the major activities involved in each task and highlights the interdependencies and relationships among the separate tasks of the proposal. For clarity, we have divided the work plan by task and will note interdependencies with other tasks and other deliverables. The details of our processes and technical methodology have been discussed in the above sections. This section provides an outline of the planned work tasks with a brief summary of each activity. For additional details of the approach for each task, please refer to the above sections.

Our Management Proposal provides details of how we will organize and manage the work plan as well as our communication plan.

The proposed work plan will be reviewed and negotiated with SBAC during the proposed Program Start-Up meeting to assure understanding and agreement before work commences. Work plans will be revised and finalized immediately after this meeting, and will become primary tools in ongoing program management, providing both the framework and the transparency required for successful program execution. We have organized the work plan to reflect the logical progression of tasks, activities, and deliverables that will assure the delivery of the Initial Achievement Level Descriptors as outlined in your RFP.

In writing this proposal, CTB assumed that the review cycles must take place consecutively and 10 days were needed prior to each cycle for panelists to review the ALDs. Working backwards from the stated end-contract date through all review cycles (including the TAC review), CTB came up with the date for the workshop. The answers to the Questions and Answers released on April 9, 2012 show that this assumption may be relaxed.

With relaxed assumptions, our timelines will change. The dates of the workshop will likely change to an August date. This change gives us approximately two months to recruit panelists. CTB has a long tradition of recruiting panelists for workshops, including our own standard setting workshops for our internal products. These panelists are selected from across the United States. We will bring this experience to our work in recruiting panelists for the Smarter Balanced ALD-development workshop. Given the change to assumptions, we recommend that each governing state nominate panelists who will later be selected by the Consortium and CTB working together to identify the strongest potential panelists. To select panelists from higher education, governing states can again nominate panelists, or we could review the relevant literature (e.g., achievement gains within mathematics) to identify potential members.

**Program Schedule**

An overall program schedule has been developed to facilitate the successful delivery of work products and deliverables. Our program team will use the schedule to carefully manage and monitor the program to ensure all customer deliverables and contract requirements are met within the agreed-upon timeline. The Project Manager will be responsible for updating information, to ensure SBAC has access to reporting on all scheduled task. Each assigned program team member has specific roles and
responsibilities related to the schedule to make certain that successful implementation of these tasks occur. The Program Manager and program team will work closely with the SBAC team to continuously review key tasks and dates, provide status updates, and make any adjustments to the timelines as required and agreed upon.

The Program Manager will also outline the critical path with the program team and customer, to highlight all critical dates and ensure key milestones are met. Throughout all phases of the program, the Program Manager continuously monitors and analyzes the schedule, looking downstream to assure all dates on the critical path are on target leading to key milestones. Working closely with the contributing team members, the Program Manager works to assure that task-level scheduled remain in alignment with the program schedule. If any impacts to the schedule are identified, the Program Manager will work with the team to bring the schedule back into alignment with the final program deliverable requirements.

We propose the following activities as part of the overall work plan for the developing achievement level descriptors.

**Activity 1: Approval of overall methodology by Smarter Balanced**

The initial activity of the contract will be to discuss the methodology described in Part 1: Oversight, Panelist Recruitment, and Methodological Design of this proposal with the Consortium Executive Committee to determine any changes or modifications that may be appropriate and to obtain approval for the recommended plan. This discussion will be facilitated by Dr. Egan and should be completed within the first week after contract award.

**Activity 2: Recruitment of ALD Development and Review Committee members**

Concurrently with the discussion and approval of our proposed methodology, CTB and the College Board will begin identifying and recruiting potential ALD Development Committee members. Recruiting panelists for the purpose of developing ALDs is a serious undertaking. It is important that the diversity of stakeholder interest is represented when ALDs are written. We have described our plan for the timely recruitment of panelists in Sec. 2.6. If CTB appears to have been named the successful contractor, we will immediately make preliminary logistical arrangements to facilitate the timely convening of the face-to-face committee. It will be essential for the recruiting of panelists to begin during the first week of the contract in June.

At the same time, we will begin recruiting panelists to review the draft ALDs that are developed by the initial committee. We will incorporate this recruiting into the processes described above in order to have the review panels fully identified prior to the beginning of each review cycle.

**Activity 3: Facilitation of ALD Development Committee**

As described in Sec. 2.7 and 2.8, we will facilitate the development of the initial ALDs during the five-day committee meeting. Dr. Karla Egan will oversee this activity.

Within each content area, two types of committees will engage in the ALD-development workshops: a meta-committee and grade-span committees. The meta-committee will comprise one representative from each grade-span group, for a total of four members within each content area. The members of this committee will also be part of the grade-span committees within their assigned content areas. The meta-committee members will assist with their own grade-span as well as helping with work of other grade-spans when appropriate. The members of this committee will create the general policy ALDs, and they will assist with the cross-grade articulation of descriptors.

Within each content area, four committees will meet to create ALDs for each grade level. Table 20 shows the grade spans and the number of committee members. These committees will be responsible for creating the grade- and content-specific range and target ALDs.

The deliverables from the committee meetings will be a complete set of policy, range, and target ALDs.
Activity 4: Preparation of Version 1 of ALD documents
After the committees have created a rough draft of the ALDs for each grade and content area, content experts at CTB and College Board will edit the general policy, range, and target ALDs so that they are concise yet detailed enough to support item and task development, test design, and standard setting. The deliverable from this activity will be a set of initial ALD documents (Version 1) ready for the review panels.

Activity 5: Review of ALDs Version 1 by Smarter Balanced Contractors and Work Group representatives
Once the initial set of ALDs is developed by the committee, the draft documents will be reviewed by four separate review panels. The first review panel will consist of Smarter Balanced contractors from previous contracts (04, 14, 09, 06, 05, 15) and representatives from the Smarter Balanced Work groups. Members of this review panel will be recruited in advance, and the reviews will be conducted online within two-week window beginning shortly after the committee meeting concludes. Comments from this review will be collected, evaluated, and applied to the initial drafts to produce Version 2 of the ALD documents. Dr. Egan and the CTB Content Lead will oversee this panel evaluation.

Activity 6: Review of ALD draft by the SBAC TAC.
As described above, CTB will present the current draft of the ALDs at the Smarter Balanced TAC meeting in July. CTB will be available to present the results of and facilitate discussion about the preliminary products from the panelists' workshops. Prior to this meeting, we will prepare the preliminary products from the panelists' workshops as well as a draft version of the technical report for review by the TAC. During the TAC meeting, we will present both sets of materials and facilitate discussion.

Activity 7: Review of ALDs Version 2
The second review panel will consist of the Smarter Balanced Executive Committee and higher education membership. This review will also be conducted online within a two week window. Recommendations from this review will be applied to the ALD documents to create Version 3.

Activity 8: Review of ALDs Version 3
The third review panel will consist of the Smarter Balance TAC and K-12 state membership. Since this review is scheduled to take place after the TAC meeting in July, TAC members may see a slightly different version of the ALD documents. At this time the will have an opportunity to provide specific comments. In addition to TAC member, this version of the ALD will also be reviewed by identified members of the Consortium K-12 membership. We will make changes based on review comments using the process described in Sec. 2.11.

Activity 9: Review of ALDs Version 4
The final round of reviews will solicit comments from PARCC and other key stakeholder groups identified by Smarter Balanced. The review process and incorporation of review comments will be the same as described previously. It is anticipated that this activity will not result in significant edits to the document. Any review comments that would result in changes to the documents will be discussed with the Consortium leadership before major edits are made.

Activity 10: Review of ALDs Version 5
This activity will provide final review and approval by the Consortium Executive Committee, Lead Psychometrician, and Validation and Psychometric Work Group. CTB will apply all requested edits to complete the final version of the documents.

The final activity of the contact will be to prepare the final set of documents for delivery to Smarter Balanced. This will include a final editorial style and proofing review and preparation of the documents in the required formats.
C. Deliverables

For each RFP task, we have defined the key program deliverables that will assure the successful delivery of final program scope and facilitate program planning and communication. We have outlined the deliverables by tasks in order to provide clear linkage to SBAC’s strategic goals. CTB’s final program schedule will detail delivery dates, including all dates for review cycles and approval, for all final task deliverables.

EXHIBIT K: Deliverables Matrix

<table>
<thead>
<tr>
<th>Deliverable ID</th>
<th>Deliverable</th>
<th>Description of content to be included</th>
<th>RFP Requirement(s)</th>
<th>Completion Date</th>
<th>Reviewers</th>
<th>Communications Plan</th>
</tr>
</thead>
</table>
| 1              | Draft Initial ALDs (version 1) | CTB has proposed a workshop that will result in draft Initial ALDs. The workshop will allow for multiple iterations of the draft ALDs by the panelists. The committee members will represent the SBAC member states as well as other expert members required by the proposal. CTB and College Board content experts will edit these draft Initial ALDs following the workshop. | Draft Initial ALDs (version 1) based on the development work with committee members. | Workshop: 6/20-24/12  
Deliver Draft Initial ALDs: 6/27/12  
Communications Plan:  
a. 6/14/12  
b. 6/25/12 | a. Contractors for Item/Task Specifications (#04), Item/Task Writing/Review – Pilot (#14), Test Design and CAT Specifications (#09), Accessibility and Accommodations Policy Guidelines (#06), Psychometric Services (#05), and Report Development (#15) | a. Email notification 10 days prior  
b. Email notification 3 days prior |
<table>
<thead>
<tr>
<th>Deliverable ID</th>
<th>Deliverable</th>
<th>Description of content to be included</th>
<th>RFP Requirement(s)</th>
<th>Completion Date</th>
<th>Reviewers</th>
<th>Communications Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Revised draft initial ALDs (version 2)</td>
<td>CTB will provide a revised version of the ALDs for each grade level and content area based on review comments from the identified reviewers for Version 2.</td>
<td>Revised draft initial ALDs (version 2) based on feedback from a. Contractors for Item/Task Specifications (#04), Item/Task Writing/Review – Pilot (#14), Test Design and CAT Specifications (#09), Accessibility and Accommodations Policy Guidelines (#06), Psychometric Services (#05), and Report Development (#15) b. Item Development, Test Design, Performance Task, Accessibility and Accommodations, and Validation and Psychometrics Work Groups, (including Lead Psychometrician and EC Liaison)</td>
<td>Deliver Revised Draft Initial ALDs: 7/19/12 Communications Plan: a. 7/17/12 b. 7/16/12; 7/18/12</td>
<td>a. Consortium Executive Committee for approval to share with states b. Consortium higher education state membership</td>
<td>a. Email notification 3 days prior b. Email notification 10 days prior; reminder notification 2 days before deadline</td>
</tr>
<tr>
<td>Deliverable ID</td>
<td>Deliverable</td>
<td>Description of content to be included</td>
<td>RFP Requirement(s)</td>
<td>Completion Date</td>
<td>Reviewers</td>
<td>Communications Plan</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Revised draft initial ALDs (version 3)</td>
<td>CTB will provide a revised version of the ALDs for each grade level and content area based on review comments from the identified reviewers for Version 3.</td>
<td>Validation and Psychometrics Work Groups (including Lead Psychometrician and EC Liaison)</td>
<td></td>
<td></td>
<td>a. Email notification 10 days prior; reminder notification 2 days before deadline</td>
</tr>
<tr>
<td>3</td>
<td>Revised draft initial ALDs (version 3)</td>
<td></td>
<td>Revised draft initial ALDs (version 3) based on Consortium feedback from a. Consortium Executive Committee b. Consortium higher education state membership</td>
<td>Deliver Revised Draft Initial ALDs: 8/12/12 Communications Plan: a. 7/26/12; 8/7/12 b. 7/26/12; 8/7/12</td>
<td>a. The CONSORTIUM’s TAC b. Consortium K-12 state membership</td>
<td>b. Email notification 10 days prior; reminder notification 2 days before deadline</td>
</tr>
<tr>
<td>4</td>
<td>Revised draft initial ALDs (version 4)</td>
<td>CTB will provide a revised version of the ALDs for each grade level and content area based on review comments from the identified reviewers for Version 4.</td>
<td>Revised draft initial ALDs (version 4) based on feedback from a. The Consortium’s TAC b. Consortium K-12 state membership</td>
<td>Deliver Revised Draft Initial ALDs: 8/28/12 Communications Plan: a. 8/15/12; 8/27/12 b. 8/15/12; 8/27/12</td>
<td>a. PARCC b. Other key stakeholder groups, as appropriate</td>
<td>a. Email notification 10 days prior; reminder notification 2 days before deadline</td>
</tr>
</tbody>
</table>

© 2012 CTB/McGraw-Hill LLC (Unpublished)
<table>
<thead>
<tr>
<th>Deliverable ID</th>
<th>Deliverable</th>
<th>Description of content to be included</th>
<th>RFP Requirement(s)</th>
<th>Completion Date</th>
<th>Reviewers</th>
<th>Communications Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Revised draft initial ALDs (version 5)</td>
<td>CTB will provide a revised version of the ALDs for each grade level and content area based on review comments from the identified reviewers for Version 5.</td>
<td>Revised draft initial ALDs (version 5) based on Revisions that enhance desired consistency with PARCC and feedback from key stakeholder groups</td>
<td>Deliver Revised Draft Initial ALDs: 9/18/12 Communications Plan: a. 9/14/12 b. 9/14/12</td>
<td>a. Validation and Psychometrics Work Group leadership b. Executive Committee</td>
<td>a. Email notification 3 days prior b. Email notification 3 days prior</td>
</tr>
<tr>
<td>6</td>
<td>Final draft initial ALDs (version 6)</td>
<td>CTB will provide a final version of the ALD documents in print ready format and digital formats that can be incorporated into a variety of databases.</td>
<td>Final draft initial ALDs (version 6) based on approval by Consortium Executive Committee, Lead Psychometrician, and Validation and Psychometrics Work Group leadership</td>
<td>Deliver Final Draft Initial ALDs: 10/8/12 Communications Plan: a. 10/4/12 b. 9/25/12; 10/5/12</td>
<td>a. V &amp; P work group b. Consortium K-12 state membership</td>
<td>a. Email notification 3 days prior b. Email notification 10 days prior; reminder notification 2 days before deadline</td>
</tr>
<tr>
<td>7</td>
<td>Technical Report</td>
<td>CTB will develop a draft and a final technical report that documents the workshops used to develop the initial ALDs, the review cycles that followed the workshops, and the validity evidence that was collected throughout the development process. Each technical report will be organized into three major sections: Section I will provide information about the applications of the initial ALDs; Section II describes key operations, including the workshop design, review cycles (final only), and the initial ALDs; and Section III presents evaluative information, including validity evidence</td>
<td>Technical report including documentation of the development process, incorporation of feedback, and validity evidence</td>
<td>Deliver Draft Technical Report: 7/11/12 Deliver Final Technical Report: 11/15/12 Communications Plan: a. 10/25/12 b. 10/16/12; 10/26/12</td>
<td>a. Validation and Psychometrics Work Group leadership b. Consortium’s TAC</td>
<td>a. Email notification 3 days prior b. Email notification 10 days prior; reminder notification 2 days before deadline</td>
</tr>
</tbody>
</table>
D. Outcomes and Performance Measurement

The Smarter Balanced Assessment Consortium has designed a balanced assessment system to monitor student growth along learning progressions and achievement continua so that states can monitor students who are on track to reach college and career readiness in high school. Since the scores from the summative and interim assessments are to be used to establish both the status and progress of student learning and progress toward college and career readiness as well as to evaluate the performance of schools and teachers, items and tasks must provide clear evidence of individual status and progress. To this end, we have defined two types of outcomes that we propose to measure. First, we will discuss the outcomes and performance measures associated with the Project Management of this work, and second, we will discuss the task-related outcomes and performance measures for this work.

Project Management

The timelines for this project are very aggressive, which means that the project will need strong oversight. While CTB will manage all parts of this project, the College Board will contribute to achieving identified outcomes that are discussed in the next section. Table 23 shows the following outcomes and performance measures related to the project management activities describe in the task-related outcomes.

Table 23. Outcomes and Performance Measures

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Description/Performance Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective collaboration with Smarter Balanced</td>
<td>Close collaboration with the Smarter Balanced Assessment Consortium and our partners is critical to the success of the project. We envision that our initial planning meeting will include discussions on Consortium expectations for preferred processes for communication and identification of key contact persons, decision makers, and roles, so that a communication plan can be established for the project.</td>
</tr>
<tr>
<td>Effective management between CTB and College Board</td>
<td>We are proposing a program management team with vast experience in managing collaborative partnerships to meet successful outcomes and delivery through use of PMI methodologies and industry best practices. Clear set expectations, requirements, defined metrics in collaboration with SBAC for approval of each deliverable, and communication with all stakeholders will help drive on-time and high-quality delivery. These partnerships coupled with effective program management will allow us to provide SBAC with leading industry experience and expertise to meet project expectations.</td>
</tr>
<tr>
<td>Maintenance of all timelines and deliverables</td>
<td>As part of our program management team, a program schedule analyst has been assigned to provide continual oversight and monitoring of key tasks that lead to final deliverable timelines. The schedule analyst works hand in hand with the project team and PM to track project status through key &quot;real time&quot; and scheduled project benchmark checkpoints. The Program Manager will report to project schedule status to SBAC through regular reporting mechanisms including the weekly management reports. Any risk to the final deliverable schedule will be mitigated by addressing internal timelines so that SBAC deliverable timelines are not impacted.</td>
</tr>
<tr>
<td>Communication and record-keeping</td>
<td>Program Management will maintain key project documentation including Meeting Agendas, Records and Minutes, Weekly Status Reports, and Records of Decision Making and Action Logs. CTB will propose a format for all project communication/record keeping, but is amenable to adopting formats currently in use with current vendors based upon SBAC-preferred formats.</td>
</tr>
<tr>
<td>Fiscal and organizational responsibility for meetings</td>
<td>Program Management will maintain fiscal and organizational responsibility for all project meetings. Logistics information will be available through established collaboration channels. Key fiscal and meeting summary information will be provided through weekly management reports.</td>
</tr>
</tbody>
</table>
Being strong proponents of the belief that we can only “manage what we can measure,” key performance metrics will be identified, tracked, and reported through regular project team “stand-up” meetings instituted to promote collaboration and assure the team is fully aware of project status and can remain focused on critical tasks and deliverables. The project team will be supported by project status dashboards and metric reporting tools that report key performance indicators (displayed through collaboration software) in a manner that supports the successful management, tracking, and reporting project status. Stand-up meetings have been part of CTB’s management culture for years, and have been used very effectively with our business partners in the SBAC proposal development process. The team building and ability to focus on critical issues quickly and effectively make this an important tool in our proposed team management plan.

Outcomes and Performance Measurement of Project Tasks
Table 24 shows each of the major outcomes of this project along with means for the consortium to measure the success of the project.

Table 24. Major Outcomes of the Initial Achievement Level Descriptors Development project

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Description/Performance Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panelist Recruitment</td>
<td>CTB will create a sampling plan and recruit panelists to complete the sampling matrix desired by SBAC.</td>
</tr>
<tr>
<td>ALD-Development Workshop</td>
<td>CTB will design and implement an ALD-Development workshop allows for panelist discussion, cross-grade discussion, and iterative drafts. During this workshop, CTB will collect panelists’ evaluations on important features of the workshop as well as on the ALDs themselves. After the workshop, CTB will analyze the panelist evaluations. These evaluations will serve as an important indicator of the validity of the method. Another important aspect of procedural validity is that the workshop is implemented according to specified design.</td>
</tr>
<tr>
<td>First drafts of general ALDs are created</td>
<td>At the close of the workshop, general and grade-specific ALDs will be created for all grades and content areas. CTB will analyze the panelist evaluations to ascertain panelist support for the ALDs, panelist belief that students who attain the KSAs in the Grade 11 ALDs will be CCR, and panelist support that there are progressions of KSAs throughout the ALDs.</td>
</tr>
<tr>
<td>First draft grade-specific ALDs are created</td>
<td></td>
</tr>
<tr>
<td>Draft Technical Report</td>
<td>The draft technical report will include description of panelist recruitment, methodology of workshop, description of round by round implementation of workshop, description of quality control evidence, initial general policy ALDs for all grade/content areas, initial grade-specific range and target ALDs for all grade/content areas that have been edited by CTB and College Board, descriptions of procedural validity, discussion of panelist feedback, and appendices with training material, evaluation results, and planned review cycles. CTB will provide this draft to the SBAC TAC, and CTB will incorporate TAC feedback into the technical report as desired by the Consortium.</td>
</tr>
</tbody>
</table>
 Revised drafts of general ALDs are created
 Revised draft grade-specific ALDs are created

**Validity Studies**
CTB will conduct validity studies that evaluate the utility of initial ALDs, in particular the range ALDs, for item writers. Another set of validity studies will evaluate the utility of initial ALDs, in particular the target ALDs, for standard setters. College Board will align CCR aspects of the Grade 11 ALDs with existing College Board indicators of CCR (unless SBAC desires that the College Board indicators be used to guide ALD development work). CTB will present the results of the validity studies to the SBAC TAC.

**Final Technical Report**
The final Technical Report will include a description of panelist recruitment, methodology of workshop, description of round-by-round implementation of workshop, description of quality control evidence, discussion of review cycles, initial general policy ALDs for all grade/content areas, initial grade-specific range and target ALDs for all grade/content areas, descriptions of procedural validity, discussion of panelist feedback, usefulness studies, and appendices with training material, evaluation results, and overview of review cycles. CTB will maintain regular correspondence with the Consortium to assure that the final Technical Report contains information desired by SBAC. CTB will provide this technical report to the SBAC TAC.

**E. Risks**
CTB and the College Board will implement processes to assure that there are effective controls in place to consistently deliver customer requirements. We strive to implement process improvements to enhance customer satisfaction and reduce costs to the customer.

To help support these efforts, the CTB program teams are trained in and implement effective risk management techniques focused on risk management planning; risk identification, analysis, responses, and monitoring; and controlling risks on a project. Risk management is critical because it equates to preventing problems, which is fundamental to increased customer satisfaction and increased efficiency and quality.

CTB conducts a risk assessment prior to beginning work. The goal is to identify potential non-conformances in advance and address them by putting controls in place to prevent them from occurring. Our standard procedure is to do this in conjunction with the customer to assure we define and understand the “Critical to Satisfaction” requirements of a project and listen to the voice of the customer. In preparation for the initial program meeting, the program manager will prepare a standard risk analysis for the Smarter Balanced Initial Achievement Level Descriptor Development program.
Through initial discussions and follow-up with SBAC, risks and controls planned to prevent them will be clarified and agreed upon with SBAC leadership. The risk plan will be managed in conjunction with SBAC as the program work progresses.

In addition to these overall risks, we have identified specific risks associated with the overall development of the ALDs. These risks fall into two categories: risks associated with the schedule for this work, and risks associate with the budget. Both risks affect the quality of the final product. Aggressive schedule dates can impact both the qualifications of the participants that are available for the committee meeting and the quality of the final documents, given the very short turnaround time between reviews.

**Schedule risks**

The schedule for this work is extremely aggressive. The required review windows for each panel review, backed up from the end of the contract, requires that the initial ALD committee meet within the first two weeks after contract award. This poses risk for both the initial committee and the versioning of the reviews. It will be difficult to recruit committee members in such a short time and secure their attendance at a meeting with little advance notice. Smarter Balanced state members require additional notice for travel approvals, and other potential panel members may have conflicts with participating on such short notice. We propose several actions to mitigate this risk:

- Smarter Balanced leadership can alert their membership of this upcoming ALD development activity and begin the identification of potential governing state participants. (This activity will most likely be required regardless of the particular vendor selected to carry out the contract given the schedule constraints).
- CTB will begin preliminary planning and identification of additional experts immediately upon contract award and before the official contract stat. While arrangements cannot be finalized, we can begin identifying appropriate individuals and making preliminary logistical arrangements.
- CTB and the College Board will begin preplanning the activities of the ALD development Committee. As previously described, we have already planned specific agendas for the face-to-face meeting and for our internal processes.

In the event that the initial committee meeting is delayed, we suggest the following mitigation activities associated with the review panels.

Reduce the time allowed for the reviews for two weeks to seven working days. This should provide sufficient time for reviewers to be introduced to the materials and to review them accordingly.

Develop the next version of the ALD documents with comments received during the first five days of the review. This would allow the reviews to be conducted back-to-back and still allow time for thoughtful editing of the documents. Comments received during the last two days of the review window would be incorporated into the next revision round. Additional time could then be allotted for the preparation of Version 5 to assure that the final documents reflect all reviewer comments.

This time saved at the end of the schedule could allow additional time for recruiting qualified participants at the beginning of the contract.

**Budget Risks**

We recognize that Smarter Balanced has limited funds and that adherence to budget guidelines is essential. We also recognize that this particular activity was not part of the original Master Work Plan and that Smarter Balanced may not have budget flexibility in this area.

We see the primary risks here in the area of stipends for participants in the review committees. This effort will be the most effective if we have participation from highly qualified individuals who have significant expertise in the development of achievement level descriptors at all levels. While many of
these individuals are SBAC work group members, others are in university or private settings and would expect a stipend for participation that is greater than the $500/day suggested in the RFP. We are hoping that these individuals will have an interest in working with this significant effort since it will have such a large national impact, and will therefore be willing to participate for a nominal stipend. We have also assumed that the representatives from the Smarter Balanced governing states would not require stipends, only travel reimbursements. There is also an associated budget risk in making travel arrangements for panel members outside of the optimal two- or three-week window for less expensive airline fares.

We have designed the ALD Development Committee with the minimum number of participants that we think is appropriate. Many experts would feel that eight members per group is a more appropriate number than the four we have proposed. Under the current design, we will be able to accommodate additional panelists only if they are able to participate without stipend or travel reimbursement.

Our mitigation strategies for these budget risks are similar to those for the schedule risks. CTB and the College Board have already developed significant details around the implementation of our methodology during the preparation of this proposal. We will initiate additional logistical planning as soon as is allowable during contract negotiation, and will be ready to finalize plans and begin recruitment of committee and panel members immediately upon contract award.
Management Proposal

A. Project Management

1. Project Team Structure
CTB/McGraw-Hill has formed a dynamic Collaborative to meet the requirements set forth in the SBAC-12 RFP. CTB and the College Board have a history of success with regard to the deliverables outlined in this proposal. Our Collaborative is led by CTB/McGraw-Hill (CTB). CTB has 85 years of demonstrated success in serving clients in the national and international assessment market. We are well known for the development of valid and reliable assessments energized by nationally recognized psychometrics/research and technological innovations. CTB has expertise in all areas of the required deliverables.

CTB/McGraw-Hill has a proven history of successfully creating initial achievement level descriptors (ALDs). CTB’s research in development of valid ALDs for the standard setting process has led many states to Full Approval without Recommendation Status with the Federal Peer Review Process as required by United States Department of Education. CTB has experience in all areas of large-scale assessment. We work exceptionally well with the College Board and understand and respect that the Collaborative must work with Smarter Balanced leadership, state members, and other partners.

The College Board has a goal of ensuring that every student has the opportunity to prepare to become college ready as well as to enroll in and graduate from college. The College Board’s college readiness initiatives promote curricula, assessment tools, and guidance resources that help K-12 students prepare for the academic rigors of higher education. The College Board has championed innovation, equity and excellence for generations of students. It works to empower teachers by training them to become skilled, dedicated, and inspiring to the students they teach in a variety of content areas. The College Board brings 100 years of experience to the test design/development industry, as well as its familiarity with the Common Core State Standards (CCSS), which is crucial to designing and developing ALDs. The College Board will bring this expertise to meet the Smarter Balanced Assessment Consortium’s goal of college- and career-readiness (CCR) for students.

CTB and the College Board are highly motivated and excited to continue a partnership with the Consortium. More importantly, we believe we can complete all of the pioneering work set forth in this proposal within the timelines. In order to complete the work, we have built strong experienced program and project management teams across CTB and the College Board. CTB will manage the work of the Collaborative while working closely with Smarter Balanced leadership and member states.

In order to meet the goals and timelines of the SBAC-12 project, close collaboration and a sense of common purpose will need to be developed immediately and maintained throughout the project. We have begun this work internally and across our alliance organization, and will extend it to Smarter Balanced leadership and work groups immediately upon contract award.

The combined capabilities of CTB and the College Board provide the expertise necessary to facilitate, develop, review, build consensus, and finalize initial ALDs based on the Consortium approval. Our proposed management structure provides the depth and breadth of experience to assure completion of tasks, services, activities, and outcomes required to deliver the complete project scope. The teams
possess the skills and expertise to assure successful planning, implementation, monitoring, and delivery of high-quality services to meet the contractual requirements as outlined in the SBAC-12 RFP. Our program team will also work hand in hand with the Smarter Balanced Leadership work groups and will be available to provide support for all Consortium stakeholders. Please see Figure 1 below for an organizational chart that shows reporting structure.

**Figure 1. Initial Achievement Level Descriptors Development Organizational Chart**

---

**Project Management Plan**

As the prime vendor, CTB provides full-service project management using an established and mature Project Management Office (PMO). Full-service project management includes demonstrable processes that are aligned with the Project Management Institute (PMI), the International Standards Organization (ISO), the Lean Enterprise Institute, Six Sigma, and Business Process Management.

Additionally, we ensure adherence to the best practices from the Operational Best Practices for Statewide Large Scale Assessment Programs developed by The Council of Chief State School Officers and the Association of Test Publishers. Using these principles provides the basis for delivering the products and services for each deliverable of the Smarter Balanced Assessment Consortium Initial Achievement Level Descriptor Development project.
2. Project Management Capabilities

a. Collaboration with all the groups necessary to complete the contract. Examples of include: work groups, state leads, advisory committees, and other Consortium contractors.

CTB understands the importance of collaborating with all groups necessary to successfully complete the requirements of the SBAC-12 RFP. This project will require the coordination and facilitation of multiple reviews of the draft of initial achievement level descriptors that are created by a committee of educators. These reviews will involve multiple levels of Consortium members including but not limited to, other Consortium contractors, work groups, state leads, Consortium Executive committee members, and the Consortium's Technical Advisory Committee. CTB will collaborate with the necessary Consortium partners to develop a communication plan and schedule to assure a successful review process.

b. Ability to deliver engaging webinar presentations and provide coherent web conferencing services for effective interactive online workshops that facilitate structured discussions or brainstorming with collaboration meeting attendees.

Due to the nature of the development work for the initial draft of the achievement level descriptors, we feel it is critical to develop these initial drafts in a face-to-face setting. Our plan described in Sec. 2.7 of the Technical Proposal proposes a plan for the development of ALDs that provides consistency across grade levels and across a framework of policy and range ALDs. Considering the depth of conversation that will be required, we feel strongly that a face-to-face meeting is necessary. We have proposed a thoughtful configuration of participants and meeting days to minimize the cost of this essential meeting.

The panel reviews can be easily conducted in an online webinar setting. The initial introduction to the review will be conducted via live webinar, and individual reviews can be completed asynchronously. CTB has extensive experience in using webinars both internally and externally and is skilled in using the various features provided within commercial webinar software to engage participants (chat, raise hand, text annotations, etc.).

To the extent possible, CTB will utilize WebEx conferencing services to conduct meetings that require the attendance of Consortium work groups, state leads, advisory committees, and other Consortium contractors.

3. Project Management Deliverables

We have established a strategic Collaborative explicitly to respond to the Smarter Balanced Assessment Consortium's vision for the development for a new generation assessment system and the opportunity to work with the Consortium to develop the initial achievement level descriptors. In forming this collaborative business partnership, we can extend CTB and the College Board's complementary strengths and available resources, disciplines, and experience and expertise in a way that best leverages our combined resources in the development of the initial ALDs, as well as comprehensive coordination across Smarter Balanced leadership and management of all project deliverables.

We have formed the management plan, design, and consultative services structure to ensure that all program tasks and deliverables are developed, reviewed, revised, and finalized within the allowable project timeline and in accordance to SBAC-12 program requirements. We have established a management and team governance structure, and internal controls to ensure that Smarter Balanced has access to a single point of contact and that lines of authority are well understood.

Close collaboration and a sense of common purpose will need to be developed immediately and maintained throughout the project cycles in order to meet the goals and timelines of the Smarter Balanced Initial Achievement Level Descriptors project. We have begun preparation work internally and will extend to Smarter Balanced leadership and work groups immediately upon contract award.
CTB has a long history of deploying the expertise necessary to facilitate, organize, coordinate, and implement the development of policies, procedures, and guidelines to develop the initial achievement level descriptors. We are attentive to the overall purpose of this procurement and the designs, services, and products to fulfill the objectives of the Smarter Balanced comprehensive assessment system. Our proposed management structure provides the depth and breadth of experience to assure completion of tasks, services, reviews, and outcomes required to deliver the complete project scope. Our team possesses the skills and expertise to ensure successful planning, implementation, monitoring, and delivery of services to meet the contractual requirements as outlined in the SBAC-12 RFP. Our program team will also work hand in hand with the Smarter Balanced Leadership and will be available to provide support for all Consortium stakeholders, as well as coordination with previously selected Smarter Balanced vendors and the Consortium’s Project Management Partner.

The structure of our project team was carefully considered to assure we provide the coordination across Consortium stakeholders and the oversight functions to ensure project completion. We have created an organizational hierarchy that provides clarity in the lines of authority necessary to ensure project task completion that meets the Smarter Balanced schedule, quality, and service requirements.

The management structure of our alliance is based on CTB maintaining overall responsibility as Prime Vendor and the College Board serving as a subcontractor. However, our organizations will work as a unified team to fulfill the full obligations outlined in the SBAC-12 RFP, in such a way that each company can contribute its highest value to the final development and delivery of the Smarter Balanced Assessment Consortium’s Initial Achievement Level Descriptors project.

We understand that across the Smarter Balanced procurement process, the Consortium can select multiple vendors to deliver each project scope. We are fully committed to provide our exceptional resources and energies to establish a unified and intensively collaborative partnership with all selected vendors.

**Our proposed team leadership structure is as follows:**

**Program Manager – Ms. Lindy Wienand.** Ms. Wienand will maintain management and oversight responsibility including decision-making, documentation and record keeping, management reporting, schedule and risk management, and communication management within our Collaborative and Smarter Balanced management. Ms. Wienand will have responsibility for the overall management of the program, and will maintain continuous communication with the Smarter Balanced Leadership and the Consortium’s Program Management Partner, ensuring accountability and transparency of our processes and deliverables. In her role as Program Manager, Lindy will maintain responsibility for leading and managing the program team in delivering a unified, coordinated, and integrated work plan that ensures the successful delivery of all proposed tasks, services, activities, review cycles, and final deliverables necessary to meet the comprehensive project requirements.

Ms. Wienand has more than eight years of program management experience at CTB/McGraw-Hill. She has successfully managed a number of custom assessment programs such as the Department of Defense Education Activity contract. As the Smarter Balanced Program Manager, Lindy is responsible for the management and oversight of all program activities including schedule creation, risk identification and mitigation, decision-making, documentation and record keeping, monitoring project implementation progress, establishing priorities, acting as a team leader for representatives of all departments involved, and interfacing with and ensuring the satisfaction of the Consortium. She collaborates with and manages a team of cross-functional professionals to execute contract requirements. She will plan and control all contract events to meet required Smarter Balanced timelines, and will monitor and control costs to complete the initial achievement level descriptors project within budget, while ensuring deliverable quality.

Ms. Wienand will be assisted by Angelica Gordon, Program Office Coordinator, and Kristal Taylor, a Program Schedule Analyst who is responsible for scheduling and tracking progress of program work on
an ongoing basis.

**Project Manager - Mr. Darin Homer.** Mr. Homer has 15 years project management experience. Darin will work with the Smarter Balanced content development and program management teams. He will manage and oversee project tasks, scope, schedules, and budget for the initial achievement level descriptor project. He will coordinate with Ms. Gretchen Schultz, the Smarter Balanced Content Development Lead, and CTB and the College Board staff in the development of high-quality ALDs. Mr. Homer will assess risks involved and implements mitigation plans to avoid problems or delays with the project. He is involved in continuous improvement programs to help further create, develop, and define processes and procedures within content development operations.

**Smarter Balanced Program Team.** CTB and the College Board will provide cohesive program teams, which will be responsible for the delivery of the development of the initial achievement level descriptors based on the deliverables specified in the SBAC-12 RFP. For each deliverable, we have assigned a program team with both excellent depth and breadth of experience that will maintain responsibility for the completion of tasks, services, and activities required to deliver the complete project scope. The teams possess the skills and expertise to ensure successful planning, implementation, monitoring, and delivery of high quality services to meet the contractual requirements as outlined in the SBAC-12 RFP. The combined teams will perform the activities/tasks outlined within our work plan and proposal schedule. The work plan and schedule, finalized upon award of the contract, are based upon customer and contract requirements. The team will work collaboratively with the Smarter Balanced Leadership to deliver the complete program scope.

**Program Work Plan**
As the prime vendor, CTB provides full-service program management using an established and mature Program Management Office (PMO). Full-service program management includes demonstrable processes that are aligned with:

- Project Management Institute (PMI),
- International Standards Organization (ISO),
- Lean Enterprise Institute,
- Six Sigma, and
- Business Process Management.

Additionally, we ensure adherence to the best practices from the Operational Best Practices for Statewide Large Scale Assessment Programs developed by The Council of Chief State School Officers and the Association of Test Publishers. Using these principles provides the basis for delivering the products and services for the Smarter Balanced Initial Achievement Level Descriptors project.

Upon contract award, Ms. Wienand will deliver program initiation documentation in the form of a Program Work Plan. This top-level program-planning document provides the details of the program scope and defined roles and responsibilities of all participants to ensure everyone is able to be productive and effective right from the start. The Program Work Plan provides the answers to the who, what, when, and how, questions related to key activities, milestones, deliverables, timeline, resources, risks, program controls, and quality controls. This document will become the central control document for all teams working on the program, and the basis for change management. The work plans for each program part align with the overarching Program Work Plan, and will define at a greater level of detail the tasks, services, and activities to accomplish the scope.

We will develop and validate the program plan through a series of discussions, meetings, and reviews with the Consortium. As changes arise during the course of the program, we will update the Program Work Plan with the latest information, to ensure that all stakeholders are in agreement with the current scope.
Within one week of contract signing, the Program Manager will submit the Program Work Plan to Smarter Balanced to begin the review and approval process. Ms. Wienand will update the plan with all changes to program scope, schedule, and final deliverables, and will be responsible for gaining approval for all such changes. No changes can take effect without the Consortium approval.

**Master Schedule**

With the magnitude and complexity of the Consortium responsibilities, it is vital that the selected vendors provide complete transparency regarding the status of the program against plan. The Consortium must be confident that vendors can deliver products and services of the highest quality and that these will be delivered per the master schedule agreement, in order to meet the dependencies of stakeholder groups and all subsequent work. The Consortium has our commitment to meet all dates as outlined in the RFP and in accordance with the final, approved Master Program Schedule. Our teams have reviewed the preliminary task schedule, and have created a draft working schedule. We have analyzed the critical path dependencies for the program, including project scope activities that will occur under separate contract by other vendors. We understand the importance of delivering on time and have detailed processes and comprehensive management plans in place that consistently satisfies our customers.

A complete draft master schedule will be reviewed upon contract award to provide the identification, organization, and sequencing of all program tasks, deliverables, and milestones. Our team will work with Smarter Balanced leadership to finalize that schedule to ensure it encompasses all Consortium requirements, and aligns with related project timelines. The program schedule will take the key elements of the program and translate them into a time-based plan to help drive and manage the work. The complete schedule will include a work breakdown structure, all tasks and activities associated with the program, and the interdependencies of the tasks to be performed. The schedule is an important tool used by the management team to communicate progress of the work to the Consortium leadership on an ongoing basis and to help drive/monitor each internal deliverable to ensure customer deliverable dates are met.

The Program Master Schedule will be created using CTB’s program management scheduling software, and will be continuously monitored and analyzed by Ms. Wienand and Mr. Darin Homer, who work daily with the team to assess progress toward delivery dates.

At our first meeting, Ms. Wienand, Program Manager, will lead a walk-through of the schedule to review all key dates and seek Consortium approval for the final schedule. We have included a draft master schedule in our proposal based on the RFP tasks and timelines. This schedule will serve as a starting point for the schedule review as we collaborate to create the final comprehensive program schedule that best meets the needs of your program. We have provided a sample schedule in Appendix A of this proposal.

**The Schedule as a Management and Reporting Tool**

The Master Schedule in combination with the comprehensive plan will be used as tools to help program management drive each phase and task of the work on the critical path to ensure that the timelines are met for each deliverable of the program.

Once the master schedule is approved by the Consortium, the program team will use it to conduct deep dives to track work daily to ensure that our program team’s work is moving forward allowing us to remain on schedule for our final delivery dates for each deliverable to Smarter Balanced. Given the collaborative nature of the deliverables on this project, close tracking of progress is essential.

Our enterprise schedule software allows our Program Manager to develop both standard and custom reports to aid the team in tracking to the required dates, conducting analyses, and reporting on status while confirming that all scope remains on schedule. Customer-facing reports on the schedule of deliverables are also prepared using this software, highlighting all key tasks, dates, and milestones. These customer reports are available in a number of formats to fit the needs of the Consortium or your
constituents. Types of reports may include Gantt charts, banded reports, standard look-ahead reports, status reports, and management reports. The report format and content will be reviewed with Smarter Balanced leadership, ensuring this will be a useful schedule tool to meet your needs. The program team will work with your representatives to determine the type and frequency of the reports you would like to receive. We understand that the Consortium will be working with multiple vendors across related and simultaneous projects. We are committed to developing a complete communication plan that provides you with the information you need, and in the format you need it, to complete future planning and reporting.

**Deliverable Matrix**

We will define the key program deliverables that will ensure the successful delivery of final program scope and facilitate program planning and communication. In our proposal, we have identified the preliminary deliverables by part and provided clear linkage to the Smarter Balanced Assessment Consortium’s strategic goals. CTB’s final Program Work Plan will include the complete Deliverable Matrix (aligned with the Master Schedule) and will provide definition of all major deliverables and assigned resources. The matrix will detail delivery dates, including all dates for review cycles and approval, for all final task deliverables.

**Management Meetings**

Holding productive management meetings between the Consortium and our program team provides an avenue for ongoing communication and collaboration. These meetings are an essential part of the implementation of the program and will allow us to build a strong collaborative relationship focused on achieving the successful contract outcomes and meeting the needs of the Consortium.

Our proposed plan for the management meetings includes a Program Start-up meeting to occur as soon after contract award as is possible. Additionally, we will participate fully in online, teleconference, or videoconference meetings with stakeholder groups, including work group meetings, Executive Committee meeting, and general membership meetings. For the Program Start-up meeting, we will maintain financial responsibility for the travel and meeting costs to support project-specific and cross-consortium engagements, as required by Smarter Balanced.

The initial objective of the Program Start-up meeting is to build a foundation upon which to build the detailed program plan and ensure a congruent understanding of requirements and specifications for deliverables and for the methodology to create the deliverables for each project phase and task. At this meeting reviews of the initial Program Plan and Master Schedule will occur in order to gain Smarter Balanced input and feedback into developing the final versions of each. It will be important that we identify and reach agreement on any necessary adjustments to scope of work and schedule.

**Deliverables from the Launch Meeting:**

- Updated Master Schedule
- Updated Program Plan including the management plan addressing communication, meetings, and management reporting
- Meeting Notes/Decision Log/Open Action List.

Equally important is that, as the work progresses and the designs are “operationalized,” adherence to program plan and work plans is monitored and reviewed at the subsequent management meetings.

**Critical Program Documentation**

The project team will be responsible for comprehensive record keeping and program documentation. The timeline for delivery of records and documentation will be determined in collaboration with Smarter Balanced and will occur in the format and at the frequency the Consortium determines is responsive to the needs of its stakeholder groups. Our project team understands from our emerging partnership with Smarter Balanced that specific formats for critical program documentation have already been identified and created. We will adopt these templates and use them on the SBAC-12 project so
that consistency is maintained for Consortium leadership, members, and partners. We have identified several forms of documentation that we would consider critical to this project.

**Meeting Agendas and Records and Minutes** – Ms. Wienand will provide Smarter Balanced with a meeting agenda for review prior to each meeting. The agenda will be designed to make certain the teams are able to address critical issues. Meeting notes will include a summary of the meeting, action items with owners and dates, and a decision log of any decisions that occurred during the meeting.

**Weekly Status Reports** – CTB will provide a detailed weekly management report to Smarter Balanced. The weekly management report will drive agenda topics for weekly meetings and will address how the project is progressing against the schedule. The weekly status report will also address issues and risks that have been identified and how they are being managed. The weekly status reports are intended to provide the necessary communication to all Consortium stakeholders groups about the progression of the project work.

**Records of Decision Making** – CTB will maintain complete record keeping for all decisions made during contract activities, including during work group meetings and management meetings. CTB will maintain the dynamic decision log throughout the program to ensure comprehensive detailed decision records are maintained, including decision being made, decision date, decision status, and the responsible party or decision authority. Decision logs may become part of the weekly status reports to provide timely communication to all working groups of all decisions.

**Quality**
CTB/McGraw-Hill will certify the adherence to CTB’s high quality control procedures for Smarter Balanced program deliverables. CTB maintains the Quality Management System (QMS) to document and confirm that customer needs and expectations are satisfied. This system supports CTB’s commitment to provide the development of assessments and reports that are consistently of high quality and reliability. The QMS is a comprehensive set of standardized policies, procedures, and processes that define the planning and execution of our core business functions, using guiding principles that include: Customer Focus— understand and meet expectations, Process Approach— activities and resources managed as a process, Continual Improvement— permanent objective, and Factual Approach to Decision Making— based on data and analysis. Following this process, customer needs are defined and translated into operational specifications. Customer requirements flow directly from contract negotiations and are vetted with the customer throughout the process, and customer change requests are clearly documented and formally approved. Corrective and preventive action systems are in place, and there are continuous activities to improve the system itself and overall customer satisfaction and loyalty.

**Proactive Risk Management**
Our team will monitor the program to identify potential problems before they happen. We plan ahead and act timely and responsibly to address any unexpected issues as they arise. We strive to continually enhance customer satisfaction. To do this effectively, we focus on our processes that help to support the customer, ensuring we have effective controls in place to consistently deliver customer requirements while implementing process improvements.

Our program team is trained to implement effective risk management techniques focused on risk management planning; risk identification, analysis, and responses; and monitoring and controlling risks on a project. We use tools that are known globally to be highly effective tools for risk management and are used in Business Process Management (BPM) and Lean Six Sigma.

We will conduct comprehensive risk assessment prior to work beginning, identify potential non-conformances in advance, and address them proactively by putting controls in place to prevent them from ever occurring. This allows us to continually improve our process and our performance. The benefits are simple: problem prevention equals increased customer satisfaction and increased efficiency and quality on the Smarter Balanced Initial Achievement Level Descriptors project.
Our program and project team will conduct a Failure Modes Effect Analysis (FMEA) for the Smarter Balanced program and will develop controls to help eliminate risks. The Program Manager can share our analysis with the Consortium for additional feedback on the control plans to best support the project. By applying proven tools and best practices, and working together to support processes that eliminate risks, we have a more formal and consistent way of conducting risk assessment, which will allow us to increase the probability and impact of positive events in our projects, and decrease the probability and impact of negative events in order to provide our customers with the highest quality outcome.

Our team will do everything possible to identify potential problems early and to provide quick solutions and/or contingency plans to meet the need and to implement corrective actions immediately.

**Monitoring and Tracking**

CTB has implemented a companywide non-conformance tracking system in an effort to control and eliminate non-conformances. The system allows our team to enter non-conformances, track root cause analysis and corrective action status, capture statistics by non-conformance type, and identify trends that require immediate attention. The ultimate goal of this system is to understand where strategic quality-improvement investments must be made to reduce non-conformances for our customers, to improve the overall experience of our customers, and to maximize the quality of our products and services. This system has been very conducive to helping us meet these goals. We have been able to put quality process improvements in place in key areas and to drastically reduce non-conformances.

Should a non-conformance occur, our team immediately alerts the customer, evaluates the severity of the problem, works with the customer to determine what controls are best to provide least impact to the field, puts controls in place, conducts root cause analysis to determine what caused the problem, what corrective and preventive actions need to occur, and develop and implement control plans and or revised procedures. Each step of the way the Consortium will be included as part of the discussions and decisions. All approved changes are documented and applied in our process documentation. Our team will address unanticipated issues as they occur and without delay to ensure the least impact on our customer.

4. **Staff Qualifications/Experience**

Our staffing plan provides experienced staff for the leadership of each activity of the proposal. CTB and the College Board are ideally suited to provide the expertise necessary to support the Smarter Balanced Assessment Consortium in developing achievement level descriptors for the next generation of assessments. We have put together a highly qualified team that will work closely with the Consortium and its various partners, and will provide superlative support in the development of the ALDs. Our team has experience developing ALDs, expertise in the content areas of English Language Arts and mathematics, a deep understanding of the Common Core State Standards, and demonstrated ability to complete this work in the context of evidence-centered assessment design. Dr. Karla Egan, Dr. Christina Schneider, and Mr. Ricardo Mercado are experienced in working with states to draft ALDs. Their work is often done in conjunction with a standard setting workshop; however, they have designed or led ALD-development workshops in Bermuda, Qatar, Minnesota, and Indiana. The operational work that Drs. Egan and Schneider conducted along with their research on ALDs directly led to the framework that was proposed for this RFP. Even though this framework is new (Egan, Schneider, & Ferrara, 2012), the team has started implementing it in North Dakota when target and reporting ALDs were written for that contract. The team’s work on standard setting and ALDs, ALD development, and ALD validation has been presented at the American Educational Research Association and the National Council on Measurement in Education. In addition, it has been published in the Peabody Journal of Education, Handbook of Accessible Achievement Tests for All Students, and Setting Performance Standards: Foundations, Methods, and Innovations.
Résumés for all staff can be found in Appendix C of this proposal. The following tables provide an overview of the project management team from each company and the staff that each of the alliance members will provide.

Table 1. CTB/McGraw-Hill ALD development project responsibilities

<table>
<thead>
<tr>
<th>Name and Title</th>
<th>Responsibilities</th>
<th>Percentage of Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marjorie Bryant, Mathematics Assessment Editor</td>
<td>Review and edit Versions 1-6 of ALD documents for Mathematics.</td>
<td>30%</td>
</tr>
<tr>
<td>Karla Egan, Lead Research Scientist</td>
<td>Design methodology for ALD development; create sampling design for panelist recruitment; facilitate workshop for ALD development, review ALD drafts, write technical report and coordinate contributions for other departments and with the College Board, contribute status report.</td>
<td>40%</td>
</tr>
<tr>
<td>Heather Farina, Mathematics Content Lead</td>
<td>Facilitate initial ALDs committee meeting, oversee revision of Math ALD documents after panel review sessions, and provide Mathematics expertise for ALD development process.</td>
<td>20%</td>
</tr>
<tr>
<td>Angelica Gordon, Program Office Coordinator</td>
<td>Record keeping, maintain communication with project team and Smarter Balanced management, and coordinate workshop logistics.</td>
<td>3%</td>
</tr>
<tr>
<td>Darin Homer, Content Project Lead</td>
<td>Report directly to Program Manager, Lindy Wienand. Maintain responsibility for developing and maintaining the Smarter Balanced project schedule; manage and report project metrics for scheduling, costs, risks and resources; oversee and track project budget, ensure project risks are identified and management plans are established; manage the documentation of scope documents as well as the change-control process; establish and execute quality improvement plans.</td>
<td>12.50%</td>
</tr>
<tr>
<td>Dan Macomber, English Language Arts Assessment Editor</td>
<td>Review and edit Versions 1-6 of ALD documents for English Language Arts.</td>
<td>30%</td>
</tr>
<tr>
<td>Ricardo Mercado, Senior Standard Setting Specialist</td>
<td>Implement design for panelist recruitment; assist with technical report.</td>
<td>6%</td>
</tr>
<tr>
<td>Christina Schneider, Research Scientist</td>
<td>Assist with methodological design, assist with technical port; review ALD drafts.</td>
<td>12%</td>
</tr>
<tr>
<td>Gretchen Schultz, Overall Content Lead and English Language Arts Content Lead</td>
<td>Facilitate initial ALDs development committee meeting, oversee revision of all ALD documents after panel review sessions, and provide ELA expertise for ALD development process.</td>
<td>20%</td>
</tr>
<tr>
<td>Kristal Taylor, Program Schedule Analyst</td>
<td>Maintain project schedule, track progress of all project work.</td>
<td>3%</td>
</tr>
<tr>
<td>Lindy Wienand, Program Manager</td>
<td>Overall management of the Smarter Balanced initial ALDs project, maintain communication with project team and Smarter Balanced management; record keeping, schedule and risk management.</td>
<td>8%</td>
</tr>
</tbody>
</table>
Biographies for CTB’s SBAC-12 project team follow in alphabetical order.

Marjorie Bryant
Assessment Editor, Mathematics
CTB/McGraw-Hill employee since 2000

Job Description: Ms. Bryant selects, develops, and writes mathematics items and scoring rubrics. Her responsibilities include preparing art manuscripts, supervising item writers, and coordinating with production, research, and editing departments. She also participates in content review and item selection with customers.

Qualifications: Ms. Bryant has over 8 years of teaching experience in mathematics in grades 5 through 8. She has also been a music director and teacher.

Contracts assigned: Washington, DC, Comprehensive Assessment System

Past work: New York State and Connecticut Mastery Test, Generation 4

Karla L. Egan
Research Manager
CTB/McGraw-Hill employee since 1999

Job Description: Dr. Egan manages the research portion of CTB custom contracts, including psychometric design, data analyses, and standard setting. She consults with internal and external clients regarding psychometric issues.

Education and Qualifications: Dr. Egan received her Doctorate in Sociology with emphasis in psychometric methods and her Master’s degree in Sociology from the University of Massachusetts in Amherst, Massachusetts. She also holds a Bachelor’s degree in Sociology from Truman State University in Kirksville, Missouri.

Successes and Innovations: Dr. Egan is an investigator in several different CTB Research & Development initiatives. She is the principal investigator on Detection of Cheating through Hierarchical Growth Models, a project that explores the ways that aberrant scores can be detected through statistical modeling. She is a co-investigator on Aligning Achievement Level Descriptors to Mapped Item Demands across Achievement Levels to Enhance Valid Interpretation of Scale Scores and Inform Item Development, a cross-disciplinary effort to explore the use of achievement level descriptors (ALDs). She was also a co-investigator on Detecting and Deleting Anchor Set Items: Guidelines and Consequence, a project that explored the consequences of removing anchor items on equating.

In addition, Dr. Egan conceived of and led the development effort to create Bookmark Pro, CTB’s proprietary software that significantly reduced turnaround time for producing standard setting materials and increased the quality and efficiency of the feedback reports produced during the standard setting workshop. She has been named CTB Employee of the Month twice since 1999 for work on standard settings.

Relevance of these to present work: Dr. Egan has ten years of experience working on all aspects of large-scale assessment.

Contracts assigned: Missouri, DoDEA, Bermuda (secondary), North Dakota (secondary), standard setting (secondary)

Heather Farina

Content Development Lead/Mathematics Assessment Editor
CTB/McGraw-Hill employee since 2004

Job Description: Ms. Farina oversees all content areas (English language arts, mathematics, science, etc.) to ensure accurate and timely completion of content area responsibilities. She provides and/or organizes needed training to enable completion of item and form development procedures by the content development team. Ms. Farina determines schedules and tracks progress in relation to the schedule for all content areas. She coordinates communication with internal and external customers for all content areas. Ms. Farina facilitates item content review meetings and range finding meetings, and facilitates standard setting committees using CTB’s Benchmark procedures.

Education and Qualifications: Ms. Farina holds both a Master’s and Bachelor’s degree in Elementary Education from Indiana University in Bloomington, Indiana. She has been a classroom teacher in Indiana and California.

Successes and Innovations: Ms. Farina has had on-time completion of tasks on a consistent basis and customer sign-off on forms with no comments or corrections requested. She has provided support for other content editors on other projects. Ms. Farina has facilitated improved communication between departments and team members.

Relevance of these to present work: Deadlines are met, customer satisfaction is provided, and a team atmosphere is cultivated to ensure quality products.

Contracts assigned: Content Development Lead/Mathematics Lead for Acuity Philadelphia; Mathematics Lead for Acuity Omaha

Past work: Common Core Summative Item Bank, Grade 7; Support for Acuity Common Core, all grades; Support for Acuity Power of You, Grades 6–8; Wisconsin CRT Grades 5 and 6; Washington, D.C. Grades 3, 4, 5 and 6; Connecticut support for Grades 3-8, 10.

Angelica Gordon

Program Office Coordinator
CTB/McGraw-Hill employee since 2003

Job Description: Ms. Gordon supports the program manager in the fulfillment of program responsibilities. She is responsible for planning, establishing, and maintaining program communications with internal and external customers. She identifies gathers, organizes, and maintains program information, specifications, deliverables, and records. Ms. Gordon plans, coordinates, and facilitates program meetings with the internal team and customer, and documents decisions. She also assists the program manager in monitoring program costs and budgets across departments, and communicates suspected problems and issues to the program manager.

Education: Ms. Gordon received an Associate’s degree in Liberal Arts from Los Angeles Harbor College in Wilmington, California.

Successes and Innovations: Ms. Gordon was the recipient of the CTB Team Achievement Award, First Quarter, in 2011. Ms. Gordon successfully planned and coordinated several workshops for the CSAP/TCAP contract, including Standard Setting, Instructional Impact, Anchor Review, Technical Advisory Committee Meeting, Start-Up Planning Meeting, and Sensitivity and Bias Review workshops. She negotiated site fees with the hotel, including catering and room costs. She arranged travel and rooming for all participants attending the workshops (approximately 20 to 90 participants per workshop; usually two to five days each).

Relevance of these to present work: Team Achievement Award was a result of working on the Colorado Program Team (CSAP/TCAP). Ms. Gordon assumed responsibility for coordination of the
workshops as On-Site Program Management Specialist.

Contracts assigned: Transitional Colorado Student Assessment Program, Colorado Alternate Assessment, Colorado English Language Acquisition.


Darin Homer

Content Development Project Manager
CTB/McGraw-Hill employee since 1997

Job Description: Mr. Homer works with content development teams and program management to develop solutions for custom assessment programs. He manages and oversees project tasks, scope, schedules, and budgets, and coordinates with content development leads and staff in the development of high quality testing materials. Mr. Homer assesses risks and implements mitigation plans to avoid problems or delays. He is involved in continuous improvement programs to help further create, develop, and define processes and procedures within publishing operations.

Education and Qualifications: Mr. Homer has a bachelor’s degree in Business Administration/Marketing from California State University in Chico, California.

Contracts assigned: Indiana; Washington, DC; and North Dakota

Past work: Florida, Missouri, Louisiana, New York State, Maryland, Tennessee, Mississippi, and Ohio

Dan Macomber

Assessment Editor II
CTB/McGraw-Hill employee since 2003

Job Description: Mr. Macomber develops item and passage specifications, reviews and edits items according to item acceptability and item quality rubrics, facilitates passage and item reviews, selects items for forms development according to research and content standards and processes, and provides scoring and reporting support.

Qualifications and Education: Mr. Macomber earned both his Master's degree in English and Bachelor's in English Literature with a minor in teaching English to speakers of other languages from Humboldt State University in Arcata, California. He is a highly skilled user of Monarch, CTB's internal item database and development tool, as well as Adobe Acrobat, Microsoft Word, and Excel. He participates in Toastmasters International, and has held officer roles in CTB's Toastmasters Club from 2006 to the present.

Successes and Innovations: Mr. Macomber has broad-based knowledge of evidence design (test parameters, blueprint, test design), framework specifications, research statistics, and design models. He has extensive experience facilitating at standard setting meetings.

Contracts assigned: His current contract assignments include Colorado (TCAP), and Indiana ISTEP and IMAST.

Past work: Mr. Macomber has also been assigned to Alaska, Wisconsin WKCE, Maryland, and the District of Columbia. He has teaching experience, and is a Returned Peace Corps Volunteer.

Ricardo L. Mercado

Sr. Standard Setting Specialist
CTB/McGraw-Hill employee since 2001

Job Description: Mr. Mercado is responsible for sample acquisition and research studies of CTB/McGraw-Hill shelf and custom assessments, and for the implementation of CTB/McGraw-Hill
standard setting activities. He works closely with CTB/McGraw-Hill Research Scientists, state
departments of education, and stakeholders to develop and implement successful solutions for sample
acquisition and standard settings.

Education and Qualifications: Mr. Mercado received his bachelor's degree in Psychology from University
of California, Davis on 2000, and received certification from the Project Management Institute in 2007.

Successes and Innovations: Mr. Mercado assisted in acquiring a sample of more than 10,000 examinees as
part of a multi-year, multi-faceted research study involving LAS Links. He cultivated partnerships between
CTB/McGraw-Hill, vendors, and stakeholders to implement sample acquisition plans; and participated in
the planning and implementation of over 60 standard settings since 2001 using a variety of standard
setting methods, including the Bookmark Standard Setting Procedure; and helped develop Bookmark Pro,
a proprietary software package that speeds the production of Bookmark standard
setting materials and reports.

Relevance of these to present work: Mr. Mercado will use the experience garnered to develop
successful sample acquisition and standard setting solutions on future contracts.

Contracts assigned: Mr. Mercado serves in a project management capacity for CTB/McGraw-Hill sample
acquisition efforts for LAS Links, LAS Links Online, LAS Links Español, and TABE Adaptive; and works on
several different contracts that call for standard setting, including the Washington English Language
Proficiency Assessment (WELPA) and the District of Columbia Comprehensive Assessment System (DC
CAS).

Past work: Research Associate (2001-2003). Mr. Mercado assisted in test analyses and standard settings
for North Dakota, Washington DC, and DoDEA.

M. Christina Schneider

Research Scientist
CTB/McGraw-Hill employee since 2006

Job Description: Dr. Schneider works on both custom contracts and standard setting. She designs
standard-settings activities for multiple statewide assessments, conducts item tryout analyses and item
response theory scaling and equating using partial-credit models, and designs research studies for the
statewide testing programs and produces corresponding reports.

Education and Qualifications: Dr. Schneider holds a doctorate in Music Education, master's degrees in
Educational Research and Music Education, and a bachelor's degree in Music Education from the
University of South Carolina in Columbia, South Carolina. She was formerly a managing psychometrician
at the state level and has direct experience with the peer review process. She wrote the
technical/psychometric sections for the state's peer review documentation and implemented research
studies related to the use of accommodations and modifications on statewide assessments.

Successes and Innovations: AERA Division D Award for Significant Contribution to Educational
Measurement and Research Methodology, 2006 contributing author of the special issue of Applied
Measurement in Education on vertically moderated standards.

Teacher Quality Research Grant titled “Investigating the Efficacy of a Professional Development Program
in Classroom Assessment for Middle School Reading and Mathematics,” 2005 Principal Investigator: $1.8
million dollar, federally funded four-year grant (2005–2009) to work with classroom teachers on building
standards-based classroom assessments and study the efficacy of the professional development on teacher
assessment skill and student achievement.

Relevance of these to present work: Dr. Schneider has a proven ability to develop high quality
experimental designs that are necessary in the peer review environment.

Past work: Psychometric and Data Analysis Group Coordinator; National Assessment of Educational Progress (NAEP) Coordinator.

**Gretchen R. Schultz**

*Principal Assessment Editor, English Language Arts*

*CTB/McGraw-Hill employee since 2001*

Job Description: Ms. Schultz is responsible for contributing to the conceptualization, planning, and implementation of complex assessment contracts/programs/products in a specific content area; developing test specifications and guiding the integration of industry-leading assessment practices in all phases of item development, item selection, and test construction in a specific content area; providing key product support to customers, both internal and external within CTB and across MHE; interacting with customers and traveling to customer sites as needed; producing high quality exemplary materials within the established timelines; and serving as subject matter expert in the response to requests for proposals, representing concerns, best practices, and thought leadership.

Qualifications and Education: Ms. Schultz has had a professional career as an educator, state assessment specialist, and developer of high-stakes tests. She has a Master of Liberal Arts, with emphasis on History of Ideas and English studies; a Bachelor of Arts in Humanities; and additional credits earned toward Supervision and Administration Certificate as well as an Advanced Certificate in Humanities. She has been pivotal in developing materials to guide and support item and test development as a teacher, as a Maryland State Department of Education assessment specialist, and as Principal Assessment Editor for CTB.

Successes and Innovations: Ms. Schultz directed the development of the Comprehensive Content Standards Framework for English Language Arts, which has become CTB’s primary system framework; delivered and developed materials for professional development relative to standards, assessment, and instruction; designed and developed prototypes for statewide English Grades 9 and 10, high-stakes assessments; developed blueprints for formative, interim, and summative assessments for shelf and custom products; and developed prototypes for performance assessments and tasks. Ms. Schultz’s contributions to CTB were acknowledged by a quarterly Silver Award in the category “Lead by Example.”

Relevance of these to present work: Current responsibilities include 1) providing training to internal staff, customers (upon request), and vendors on item and test development, content-specific topics, and production processes; 2) supporting teams through development and production of high-stakes, custom assessments; 3) supporting proposal teams with text, materials, and presentations. Presently, Ms. Schultz provides leadership primarily to the English and social studies groups, but has also provided support to math and science teams, and serves on key company committees and teams.

Past work: Content manager of assessment programs for Maryland, Connecticut, Kentucky, Massachusetts, Bermuda, Wisconsin, and the District of Columbia.

**Kristal Taylor**

*Program Schedule Analyst*

*CTB/McGraw-Hill employee since 1999*

Job Description: Ms. Taylor is responsible for supporting program and project managers in creating, maintaining, and sustaining resource-loaded schedules for multiple simultaneous programs at CTB.

Education: Ms. Taylor holds a Bachelor's degree in English and with a minor in Spanish from California Polytechnic State University in San Luis Obispo, California.
Successes and Innovations: Ms. Taylor worked for more than two years in the Colorado Department of Education as an on-site CTB representative. 

Relevance of these to present work: Working on site with the customer enabled Ms. Taylor to get contractual information quickly and to deliver it to the program and project managers for prompt application. 

Contracts assigned: Colorado: CSAP, CSAPA, CELApro, CELAplace, and Department of Defense Education Activity U.S. History. 

Past work: Colorado Program Office Coordinator, Program Management Specialist In-State Colorado. 

Lindy Wienand 

Program Management 
CTB/McGraw-Hill employee since 2004 

Job Description: Ms. Wienand has successfully managed a number of statewide assessment programs, as well as various shelf programs such as the Department of Defense Education Activity contract. As a Program Manager, she is responsible for the management and oversight of all program activities including schedule creation, risk identification and mitigation, decision documentation and record keeping, monitoring project implementation progress, establishing priorities, acting as a team leader for representatives of all departments involved, and interfacing with and ensuring the satisfaction of the customer. She collaborates with and manages a staff of cross-functional professionals to execute contract requirements, plans and controls all contract events to meet required timelines, and monitors and controls costs to complete projects within budget, while ensuring deliverable quality. 

Education: Ms. Wienand has a Bachelor’s degree in Human Environment Sciences from the University of Alabama 

Contracts assigned: Department of Defense Education Activity, Ohio/AIR 

Past work: Alaska, Arizona, Cincinnati Public Schools, various Diocesan Shelf Contracts, Hawaii/AIR, Missouri MAP and Missouri LAS Links.

The College Board Key Staff Assignments: 
Dr. Andrew Wiley and Dr. Pamela Kaliski will serve as lead psychometric facilitators from the College Board. Dr. Wiley has extensive experience working within the educational environment of college and career readiness. His experience includes the identification of benchmarks for college readiness on the SAT, and working in the development of validation of assessments designed to assess college readiness. Dr. Kaliski has led standard settings for several programs at the College Board. She has co-facilitated the development of ALDs for several Advanced Placement exams, and worked on developing ALDs used for standard setting purposes for ACCUPLACER. Dr. Kaliski is also a co-facilitator of a full day evidence-centered assessment design training workshop at the annual meeting of the National Council of Measurement in Education, where she will be facilitating the component about ECD-based ALDs. Additionally, Dr. Kaliski conducts research related to ALD development, validation, and item alignment with ALDs (Kaliski, Huff, & Barry, 2011; Reshetar, Chajewski, Lionberger, Plake, & Kaliski, 2012). 

References: 
Table 2. College Board ALD development project responsibilities

<table>
<thead>
<tr>
<th>Name and Title</th>
<th>Responsibilities</th>
<th>Percentage of Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kelly Godfrey, Associate Research Scientist</td>
<td>ALD development, review drafts, contribute to report, monitor deliverables and timelines</td>
<td>5%</td>
</tr>
<tr>
<td>Pamela Kaliski, Assistant Research Scientist</td>
<td>Facilitate workshop, ALD development, review drafts, contribute to report</td>
<td>10%</td>
</tr>
<tr>
<td>Kevin Sweeney, Executive Director of Psychometrics</td>
<td>Review drafts, contribute to report</td>
<td>1%</td>
</tr>
<tr>
<td>Natasha Vasavada, Executive Director, Standards and Curriculum Alignment Services</td>
<td>Review drafts of ALDs</td>
<td>2%</td>
</tr>
<tr>
<td>Andrew Wiley, Executive Director, Research and Development</td>
<td>Facilitate workshop, Achievement Level Descriptors (ALD) development, review drafts, contribute to report</td>
<td>10%</td>
</tr>
<tr>
<td>Data Analyst(s)</td>
<td>Data merging, matching, cleaning</td>
<td>10%</td>
</tr>
</tbody>
</table>

Biographies for the College Board SBAC-12 project team follow in alphabetical order.

Kelly E. Godfrey
Associate Research Scientist, Research & Development
College Board since 2007

Job Description: Dr. Godfrey is a member of Research & Development’s Psychometrics team. Her job duties entail psychometric support for College Board programs such as Advanced Placement, ACCUPLACER, SAT, and PSAT/NMSQT. She conducts both operational and empirical work, including validity research, IRT parameter estimation, test equating, and advanced statistical modeling such as HLM.

Education: Dr. Godfrey holds a Ph.D. from the Educational Research Methodology program at the University of North Carolina at Greensboro, specializing in research, measurement, and evaluation. She also holds a degree in Psychology from the University of North Carolina at Chapel Hill.

Accomplishments and Experience: Dr. Godfrey is an active member of the American Educational Research Association (Division D), as well as the National Council of Measurement in Education (NCME), where she is currently co-chair of the Membership Committee. She has published her findings on grade inflation and teacher professional development, and has co-authored a chapter in the upcoming Handbook on Measurement, Assessment, and Evaluation in Higher Education on responsive evaluation. Kelly is also a trainer for QSR’s NVivo software, a qualitative data analysis platform, and frequently visits institutions of higher education and research organizations to lead workshops on its uses.

Prior to her work at the College Board, Kelly held psychometric research graduate intern positions at Measured Progress (Dover, NH) and Educational Testing Service (Princeton, NJ). She also worked as a reader for a high school algebra range-finding at Measurement Incorporated in Charlotte, NC.

Relevant Experience: Dr. Godfrey is proficient in measurement programs such as BILOG, MULTILOG, and PARSCALE, in addition to statistical analysis software packages SPSS, HLM and SAS. She also has experience facilitating standard setting activities for diagnostic exams.
**Pamela Kaliski**  
*Assistant Research Scientist*  
*College Board, since August 2008*

**Job Description:** Pamela is an Assistant Research Scientist at the College Board, on the Psychometrics team. Her role is to collaborate with others to provide psychometric services and conduct relevant research studies to College Board testing programs.

**Education:** Pamela Kaliski received her Ph.D. from James Madison University in May 2009, in Assessment and Measurement. Prior to that, she received a Master’s degree in Psychological Sciences in 2006, also from James Madison University, and a Bachelor’s degree in Psychology in 2003 from Appalachian State University.

**Accomplishments and Experience:** Beginning in August 2008, Pamela worked 30 hours per week for the College Board in Research and Development. Once she completed her Ph.D. in May 2009, she was hired as a full-time Assistant Research Scientist at the College Board. From August 2008 to November 2009, Pamela was on the Assessment Design team, supervised by Kristen Huff, and worked mostly on evidence-centered assessment design (ECD) work for the Advanced Placement (AP) program. She then moved to the Psychometrics team, supervised by Kevin Sweeney and Maureen Ewing. She continues to work on ECD projects for AP and conducts research on ECD issues (e.g., identifying difficulty drivers, how best to generate task models). Aspects of this project involve developing form assembly specifications, working with the vendor to review item analysis results, and making recommendations for task model improvement. Over the past year, she has been involved with two standard settings and one achievement level description (ALD) development meeting. She will facilitate at least one standard setting in the upcoming fiscal year, and at least one ALD development meeting. Additionally, Pamela conducts reliability studies (with a focus on generalizability theory models to understand rater reliability) and validity research studies (with a focus on hierarchical linear models) for College Board assessments.

**Kevin P. Sweeney**  
*Executive Director of Psychometrics, Research and Development*  
*College Board since 2009*

**Job Description:** Dr. Sweeney is responsible for the management of the Psychometrics team at the College Board. In that capacity he is responsible for the oversight and direction of psychometric efforts for College Board assessment products and programs. This work includes ensuring that the high-quality technical characteristics of the College Board assessments, including programs such as the AP, SAT, SAT Subject Tests, PSAT-NMSQT, Accuplacer, and CLEP, are maintained as the programs evolve. He supervises over a dozen research scientists and psychometricians in offices in Newtown, PA and New York.

**Education:** Dr. Sweeney has a Ph.D. in Psychometrics from Fordham University. He also holds a Master’s Degree in Psychology from Fordham University and a Bachelor of Arts degree in Psychology from St. Bonaventure University.

**Accomplishments and Experience:** Dr. Sweeney has participated in invited panels, coauthored a book chapter, and made numerous presentations at national and regional conferences on the topic of standard setting. Most recently he was a discussant at the 2011 Large Scale Assessment Conference where he discussed the standard setting methodology proposed for using NAEP as a measure of career readiness. He has designed and conducted standard setting workshops in over a dozen states for statewide assessments. These standard settings have ranged across grades from 2 through 12, and content areas including math, ELA, writing, and science; they have included general assessments as well as ELL and alternate assessments. He has made presentations of standard setting results to policy makers, laymen, technical audiences, and other stakeholders. Dr. Sweeney developed a significant improvement to the body of work methodology that both streamlined the process and improved the
accuracy of the cut scores. In addition, since joining the College Board, he has overseen standard setting work for Accuplacer and AP, as well as overseeing field trials and analyses to develop College Ready Benchmarks for Readistep and PSAT. His research and publications focus on higher standard setting and technical psychometric issues including equating and differential item functioning.

Previous Experience: Before joining the College Board, Dr. Sweeney’s psychometric career began at the American Institute of Certified Public Accountants where, for 5 years, he oversaw and implemented the psychometric work on the CPA licensing exam, including the exam’s first equating and standard settings. He then joined Measured Progress where he was responsible for all psychometric analyses, including item and test analyses, equating and scaling, as well as reporting of results for customized statewide assessments. During his more than ten years there he also contributed significantly to the development of standard setting methodologies for customized state-wide tests, including alternate assessments and English language proficiency tests.

**Natasha Vasavada**

*Executive Director, Standards and Curriculum Alignment Services, Research and Development*

*College Board since 2007*

Job Description: Natasha Vasavada serves as Executive Director of Standards and Curriculum Alignment Services in Research and Development at the College Board. She leads the College Board's standards and curriculum alignment initiatives to link College Board assessments and curricula to state, national, and international standards, including the Common Core State Standards.

Education: Ms. Vasavada holds an M.Ed. in Educational Leadership and Policy Studies from Teachers College, Columbia University, and an M.A. in Educational Leadership from Fairleigh Dickinson University. She received her B.A. (summa cum laude) in history and German from Davidson College.

Accomplishments and Experience: Ms. Vasavada's areas of expertise include standards and curriculum development as well as standards alignment with assessments, curriculum, and other content frameworks. She served on both the content teams and advisory committee for the development of the Common Core State Standards, and she has designed and authored studies examining the relationship between the Common Core State Standards and the Advanced Placement, SAT, and PSAT/NMSQT programs. Ms. Vasavada also oversees the development and implementation of the College Board Standards for College Success, including their use as the foundation for College Board curriculum and assessment programs. Ms. Vasavada has supervised content teams in providing standards alignment expertise to states seeking to build new assessments aligned with the Common Core State Standards (e.g., New York) and states seeking to revise and strengthen their own state standards and align them with college readiness expectations (e.g., Colorado, Texas, Virginia). She has an extensive curriculum and assessment background utilizing the Understanding by Design and Evidence Centered Design models, and her content expertise and training is in the areas of humanities, history, and social studies.

Previous Experience: Prior to joining the College Board, Ms. Vasavada worked in New Jersey public schools as a K-12 social studies supervisor and as a high school social studies and history teacher. Under her leadership, she has worked with educators to design standards-based curricula in the areas of U.S. history, Modern World History, Middle Eastern Studies, Historical Research, World Geography, and U.S. and Comparative Government and Politics. She served on the national AP History Commissions in 2006-2008 to define the core knowledge, skills, and achievement levels expected of students to be successful in college-level history courses.

**Andrew Wiley**

*Executive Director in Research & Development*

*College Board since 2002*

Job Description: Dr. Wiley is responsible for a team of test developers and research scientists who are
charged with the design, development, and implementation of all new assessment initiatives at The College Board.

Education and Qualifications: He received his doctorate in psychometrics from Fordham University, and his work experience includes time spent in both the education and certification/licensure field. Prior to joining the College Board in 2002, Dr. Wiley worked at the American Board of Internal Medicine where he worked on the scoring, standard setting and equating of medical certification examinations. He has also worked for the American Institute of Certified Public Accountants where he worked on equating and standard setting for the licensing examination for CPAs. Dr. Wiley has worked for the Association of American Medical Colleges, where he completed research on the validity of the Medical College Admission Test (MCAT), and the American Council on Education, where he conducted research investigating the reliability and validity of the General Equivalency Diploma (GED).

As an active member of the measurement community, Dr. Wiley serves on the Board of Directors for the Association of Test Publishers (ATP). In previous years, he has served as both the program chair and executive committee chair for ATP.

B. Experience of the Vendor

Alliance Members General Capabilities

CTB /McGraw-Hill

Founded in 1926, CTB/McGraw-Hill is a leading provider of high-quality assessment products and services that help learners of all ages meet their potential. To achieve this, we provide successful paper/pencil and online solutions that play a vital role in education nation- and worldwide.

As experts in educational measurement and assessment, CTB’s goal is to facilitate sound policy-making by providing complete, accurate, and unbiased information regarding the scientific and measurement implications associated with valid alternatives, and to conduct the ongoing research to support sound decision-making. We do not make policy decisions, but we can provide information to indicate the best direction for the assessment or program. As an initial means of response to a request for information to support policy decisions, our customers may draw on our many years of experience in K-12 educational measurement and the extensive knowledge of our research staff concerning educational policies that relate to the design and consequences of many of the nation’s testing programs.

Empirically-based research often provides a sound platform from which to make policy decisions. Though the existing research base can inform assessment policy, custom-designed research studies may better inform unique policy considerations. We have provided many custom studies in support of client needs to evaluate policy intended to foster compliance with the Standards for Educational and Psychological Testing.

CTB provides technical reports, as required, that quantify each program’s processes, methodologies, and quality control procedures, and which clearly document our adherence to the highest psychometric standards. We work with states to prepare the necessary documentation for USDOE reviews by ensuring that technical reports provide the type of information required. We are proud of the fact that over half of the first states to pass Peer Review worked with CTB.

CTB works with educators at every step of the assessment and reporting cycle. We provide these services with large and small programs tailored to meet the goals and budgets of schools, districts, states, private businesses, and countries. In addition to our custom work, we also publish innovative norm-referenced and criterion-referenced assessments that provide inexpensive, reliable, and valid information to support student growth for school districts, adult education facilities, and other users.

For 85 years, CTB/McGraw-Hill has distinguished itself as a leader in educational assessment and reporting with the highest commitment to excellence and service. Our clients’ students are part of a
family of 18 million, in 8,700 school districts and dioceses, in 50 states, and in 46 countries including Bermuda, Canada, China, France, Germany, Guam, India, Japan, Peru, Puerto Rico, Qatar, Singapore, South Korea, the United Kingdom, and the U.S. Virgin Islands that benefit from CTB-developed assessments, scoring and reports.

CTB collaborates with and provides services to a number of state assessment programs at any given time. Our best references may well be the long-term relationships we have built with many states, districts, and schools and our continued ability to meet their changing program needs while continuously providing quality assessments, scoring, and reporting. Table 3 summarizes our wide-ranging experience with large-scale assessments.

Table 3. CTB’s Experience with Large-Scale Assessment

<table>
<thead>
<tr>
<th>State / District</th>
<th>Program Name</th>
<th>Custom / Existing</th>
<th>ELA</th>
<th>M</th>
<th>S</th>
<th>SS</th>
<th>H</th>
<th>ELL</th>
<th>Online</th>
<th>Grades</th>
<th>Students</th>
<th>Contract Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>State Assessment</td>
<td>E</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3-10</td>
<td>18,000</td>
<td>1997-2012</td>
<td></td>
</tr>
<tr>
<td>Alabama</td>
<td>Alternate Assessment Program</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>K-8, 10, 11</td>
<td>5,500</td>
<td>2006-2013</td>
<td></td>
</tr>
<tr>
<td>Bermuda</td>
<td>Norm &amp; Criterion Referenced</td>
<td>C/E</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>3-8, 11</td>
<td>3,700</td>
<td>2006-2012</td>
<td></td>
</tr>
<tr>
<td>Bermuda</td>
<td>Middle School Assessment</td>
<td>E</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td>6-8</td>
<td>4,000</td>
<td>2005-2012</td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>STAR</td>
<td>E</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2-12</td>
<td>4,500,00</td>
<td>2002-2009</td>
<td></td>
</tr>
<tr>
<td>Colorado</td>
<td>Colorado Student Assessment Program (CSAP)</td>
<td>C/E</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3-10</td>
<td>475,000</td>
<td>1996-2011</td>
<td></td>
</tr>
<tr>
<td>Colorado</td>
<td>CSAP Alternate</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>3-10</td>
<td>4,544</td>
<td>2006-2011</td>
<td></td>
</tr>
<tr>
<td>Colorado</td>
<td>English Language Assessment (CELA)</td>
<td>C</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K-12</td>
<td>90,000</td>
<td>2007-2011</td>
<td></td>
</tr>
<tr>
<td>Department of Defense</td>
<td>End-of-Course Testing</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td>3,5,6,7</td>
<td>25,000</td>
<td>2004-2013</td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td>Administration FCAT</td>
<td>E</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td>3-11</td>
<td>1,900,00</td>
<td>2006-2010</td>
<td></td>
</tr>
<tr>
<td>Georgia</td>
<td>Criterion Referenced Competency Tests (CRCT)</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td>1-8</td>
<td>980,000</td>
<td>2006-2011</td>
<td></td>
</tr>
<tr>
<td>Hawaii</td>
<td>Hawaii State Assessment</td>
<td>E</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3-8, 10</td>
<td>97,400</td>
<td>2006-2013</td>
<td></td>
</tr>
<tr>
<td>State / District</td>
<td>Program Name</td>
<td>Custom / Existing</td>
<td>ELA</td>
<td>M</td>
<td>S</td>
<td>SS</td>
<td>H</td>
<td>ELL</td>
<td>Online</td>
<td>Grades</td>
<td>Students</td>
<td>Contract Period</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>-----</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>---</td>
<td>-----</td>
<td>--------</td>
<td>--------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>Indiana</td>
<td>ISTEP+</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-8</td>
<td>350,000</td>
<td>1984-2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indiana</td>
<td>English Language Proficiency Assessment</td>
<td>C</td>
<td>9</td>
<td>K-12</td>
<td>31,942</td>
<td>2006-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indiana</td>
<td>Statewide Diagnostic Assessments</td>
<td>C/E</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-8</td>
<td>350,000</td>
<td>2008-2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maryland</td>
<td>English Language Proficiency</td>
<td>C/E</td>
<td>9</td>
<td>K-12</td>
<td>30,000</td>
<td>2006-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missouri</td>
<td>Custom and Terra Nova Survey</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-8</td>
<td>240,000</td>
<td>1995-2013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missouri</td>
<td>Benchmark Tracker ScoreSmart</td>
<td>E</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>teacher</td>
<td>5000</td>
<td>2008-2010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Jersey</td>
<td>Private Labeled Benchmark Tracker</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-8</td>
<td>230,000</td>
<td>2007-2012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>NCLB</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>3-8</td>
<td>1,500,000</td>
<td>1997-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York City</td>
<td>Periodic Assessment Program</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-11</td>
<td>200,000</td>
<td>2007-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nevada</td>
<td>English Language Development Program</td>
<td>C/E</td>
<td>9</td>
<td>K-12</td>
<td>90,000</td>
<td>2005-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Dakota</td>
<td>Statewide Assessment</td>
<td>C/E</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-8</td>
<td>36,000</td>
<td>1997-2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio</td>
<td>Custom Development</td>
<td>C</td>
<td>9</td>
<td>5</td>
<td>100,000</td>
<td>2004-2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qatar</td>
<td>QCEA</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>4-12</td>
<td>45,000</td>
<td>2003-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qatar</td>
<td>QSSC</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>12</td>
<td>4,700</td>
<td>2008-2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utah</td>
<td>Online Writing</td>
<td>C</td>
<td>9</td>
<td>K-12</td>
<td>22,600</td>
<td>2005-2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington DC</td>
<td>NRT/CRT</td>
<td>C/E</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-12</td>
<td>42,000</td>
<td>2004-2014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West Virginia</td>
<td>Westest &amp; EOI</td>
<td>C</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>3-12</td>
<td>300,000</td>
<td>2002-2013</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CTB provides statewide assessments that are fully performance-based, supports multiple measures of student growth and continues to pioneer English-language development proficiency assessment. We deliver summative, interim, and formative assessments online, on paper or via student response devices to deliver the timeliest results possible. Our flexible infrastructure allows assessments to be incorporated into classrooms and makes the integration of multiple student measures practical.

Our ability and our history show in our work across assessments. We help states and educators focus on what students should know and can do by aligning assessments with standards to insure accurate results that measure and support student growth.

CTB works with educators at every step of the assessment and reporting cycle. We provide these services with large and small programs tailored to meet the goals and budgets of schools, districts, states, private businesses, and countries. CTB provides a range of solutions to support key education goals, among them:

- assistance in the design and development of summative and formative assessment programs at every level—state, district, and school—that meet federal and state requirements
- involving educators in item development, review and scoring, providing valuable professional development to help educators successfully use student data to effect change
- diagnostic results that inform and guide instruction, including early diagnosis of reading and language difficulties
- language assessments and assessments in other languages for students whose language is other than English
- student reports designed to fill the information needs of administrators, teachers, and parents to evaluate student growth from year to year and throughout the year.

At CTB, we know that coordinating and managing multiple subcontractors and handoffs across components and vendors can be challenging.

CTB has proven experience with the full spectrum of partnering relationships, including co-vendors, working as a sub-contractor, and working as the prime contractor to deliver on the on all requirements of complex assessment programs. We have successfully partnered states with other vendors, such as AIR in Ohio and ETS in California. In all partnering relationships, it is critical that the principals begin with and maintain a sense of common purpose across organizations and teams. As the prime vendor for
the Smarter Balanced Initial Achievement Level Descriptors Development program, it is our Program Management team that will maintain primary responsibilities to ensure that the project teams, made up of individuals from our partner organizations, work cooperatively and effectively to successfully deliver all program requirements. Through the use of proven tools and techniques, we will drive the success of the project across the partnership teams. We will develop and implement:

- Clearly identified requirements, expectations, and key performance measures
- Continuous communication and collaboration across these partnerships with a clearly established communication plan to guide the process
- Identification of dependencies and real-time monitoring of those dependencies to readily mitigate risks to ensure quality and delivery are not impacted
- Frequent checkpoints meetings (stand-ups) and monitoring of progress toward outcomes/deliverables
- Collaborative technologies and a team portal, telecommunications, WebEx, video-conferencing
- Weekly Progress Reports.

Collaboration will be aided through a team hierarchy where all participating team members understand project reporting relationship, escalation paths, and status reporting mechanisms. Daily/frequent stand-up meetings, aided by project reporting dashboards, will keep critical objectives/deliverables visible to all contributors and provide transparency to management staff. Secure team portal sites will aid the team in ongoing communication and collaboration, while maintaining the security and limited accessibility to project contributors and stakeholders.

The College Board

The functions of the Research and Development (R&D) Department at the College Board encompass providing the scientific knowledge and processes in developing and maintaining the quality of the College Board assessments and educational initiatives. In undertaking these responsibilities for assessments, R&D has expertise and capacity in all phases of the test development process (including articulation of content and curriculum standards, alignments, assessment design, and test development), psychometrics (item analysis, scaling/equating, standard setting and validation studies), formative assessments, and statistical analysis of large-scale test data. With respect to educational initiatives, R&D has expertise in learning theory and in economic, social, educational, and psychological theories of education, and both the expertise and capacity to design and implement all types of program evaluations and other applied research studies that examine the impact of programs in schools. The department is comprised of 75 staff members and 10 part-time graduate student interns with over 50 percent of staff having doctorates in economic, education, psychology, and sociology. Annually, R&D staff members present at national and regional professional conferences, hold leadership positions in these professional organizations, and publish articles in peer-reviewed journals. Each year, R&D undertakes almost 300 projects representing the full spectrum of measurement and educational research.

1. Indicate other relevant experience that indicates the qualifications of the Vendor, and any subcontractors, for the performance of the potential contract.

CTB has extensive experience in working with states to develop achievement level descriptors. We have selected an experience program team for the Smarter Balanced initial achievement level descriptor development project. Dr. Karla Egan, Dr. Christina Schneider, and Mr. Ricardo Mercado are experienced in working with states to draft ALDs. Their work is often done in conjunction with a standard setting workshop; however, they have designed or led ALD-development workshops in Bermuda, Qatar, Minnesota, and Indiana. The operational work that Drs. Egan and Schneider conducted along with their research on ALDs directly led to the framework that was proposed for this RFP. Even though this framework is new (Egan, Schneider, & Ferrara, 2012), the team has started implementing it in North Dakota when target and reporting ALDs were written for that contract. The team’s work on
standard setting and ALDs, ALD development, and ALD validation has been presented at the American Educational Research Association and the National Council on Measurement in Education. In addition, it has been published in the Peabody Journal of Education, Handbook of Accessible Achievement Tests for All Students, and Setting Performance Standards: Foundations, Methods, and Innovations.

The College Board has experience with using Evidence-Centered Design (ECD) in the development of assessments. In recent years, the Advanced Placement program began a revision to its three history subjects (US History, World History, European History) and four sciences (Biology, Chemistry, Environmental Science, Physics). In 2007, College Board Research and Development staff, along with several consulting psychometricians, used an ECD framework to lead the development of Achievement Level Descriptors (ALDs) for all seven subjects. Acknowledging the uniqueness of each subject, as well as the similarities across disciplines, College Board researchers were able to implement various approaches to ALD development for the support of item development, score interpretation and maintenance of a high quality assessment program. In addition, College Board Research and Development leadership has been involved in leading and facilitating ALD/PLD development for several states, including Texas and Tennessee.

College Board’s experience developing achievement level descriptors and implementing Evidence-centered Design involves its own Advanced Placement program. Working alongside well-known psychometricians in the field of educational measurement, the methodology was thoroughly evaluated and documented in a special issue of Applied Measurement in Education, published in 2010.

References:


2. Include a list of contracts the Vendor has had during the last five (5) years that relate to the Vendor’s ability to perform the services needed under this RFP. List contract reference numbers, contract period of performance, contact persons, telephone numbers, and fax numbers/e-mail addresses.

<table>
<thead>
<tr>
<th>Programs</th>
<th>Program Type</th>
<th>Contract Reference Number</th>
<th>Contact Information</th>
<th>Contract Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bermuda</td>
<td>Bermuda Criterion Reference Test (BCRT) Custom assessment</td>
<td>05143-KG05-000</td>
<td>Name: Kalreta Conyers-Steede Tel: (441) 278-3352 Email: <a href="mailto:ksteede@gov.bm">ksteede@gov.bm</a></td>
<td>2005-2011</td>
</tr>
<tr>
<td>Indiana</td>
<td>Indiana State Assessment Custom assessment</td>
<td>7-93</td>
<td>Name: Michele Walker Tel: (317) 232-9050 Email: <a href="mailto:mwalker@doe.in.gov">mwalker@doe.in.gov</a></td>
<td>1984-2012</td>
</tr>
<tr>
<td>North Dakota</td>
<td>North Dakota State Assessment Custom assessment</td>
<td>2017001-0619-33</td>
<td>Name: Greg F. Gallagher Tel: (701) 328-1838 Email: <a href="mailto:ggallagher@nd.gov">ggallagher@nd.gov</a></td>
<td>1997-2015</td>
</tr>
<tr>
<td>Qatar</td>
<td>Qatar Assessment Custom assessment</td>
<td>C1102110</td>
<td>Name: Dr. Hamda Al-Sulaiti Tel: +974 4455-9401 Email: <a href="mailto:hamda.alsulaiti@sec.gov.qa">hamda.alsulaiti@sec.gov.qa</a></td>
<td>2004-2012</td>
</tr>
</tbody>
</table>

C. References

CTB: Indiana State Assessment
Contact: Michele Walker
Title: Director of Student Assessment Organization: Indiana Department of Education Address: 151 W. Ohio Street, Indianapolis, IN 46204 Telephone: (317) 232-9050 Email: mwalker@doe.in.gov
Contract Period: 1984-2012

The Indiana Department of Education (IDOE) collaborated with CTB/McGraw-Hill (CTB) to conduct a cut score setting for the Indiana Statewide Testing for Educational Progress-Plus (ISTEP+) tests on July 14-17, 2009, in Indianapolis, Indiana. The Bookmark Standard Setting Procedure (BSSP; Lewis, Mitzel & Green, 1996) was implemented to set performance standards for grade and content areas: Grades 3-8 English/Language Arts, Grades 3-8 Mathematics, Grades 4 and 6 Science, and Grades 5 and 7 Social Studies.

The original ISTEP+ testing program was administered annually in fall to students throughout the state of Indiana. Students were tested on the academic standards associated with the previous school year (e.g., Grade 5 students were tested in fall on the standards taught in Grade 4). Cut scores for the original fall ISTEP+ tests were established across years, coinciding with the development and administration of grade and content areas as they were added to Indiana’s assessment program. Starting with the 2008-09 school year, the test window for the ISTEP+ moved to spring, testing students on the standards associated with the same school year (e.g., Grade 4 students were tested in spring on the standards taught in Grade 4). To establish cut scores for the spring ISTEP+ that were reasonable and
comparable to the cut scores for the fall ISTEP+, a cut score review was held for English/Language Arts, Mathematics, and Science. Social Studies was tested for the first time in the spring of this same year and therefore required new cut scores to be set.

During the workshop, two activities were conducted: cut score review for English/Language Arts, Mathematics, and Science; and cut score setting for Social Studies. The methods used for the cut score review and the cut score setting were very similar and are described in this report. For the cut score review, participants considered preliminary cut scores, which were based on the existing fall cut scores, and recommended new cut scores for the spring ISTEP+. For the cut score setting, workshop participants engaged in the BSSP to recommend cut scores for the Grade 5 and 7 Social Studies tests, which were new in the 2008-09 administration. Participants recommended cut scores associated with three performance levels: Did Not Pass, Pass, and Pass Plus. Pass Plus represents the highest level of knowledge, skills, and abilities expected. The 2009 ISTEP+ Cut Score Setting consisted of orientation, training, three rounds of judgments, smoothing discussions, and refinement of the performance level descriptors (PLDs).

After the workshop, the IDOE presented the participant-recommended cut scores and associated impact data to the National Technical Advisory Committee (TAC) and the state’s Education Roundtable. Both of these committees supported the participant-recommended cut scores without modification.

The Indiana State Board of Education approved cut scores for the ISTEP+ on July 28, 2009. Indiana has received Full Approval Status with the United States Department of Education which includes evaluation of development and descriptors and standard setting process. In addition, Michele Walker Director of Student Assessment is available as a verbal reference to this work at 317-232-9050

**CTB: North Dakota State Assessment**

**Contact:** Greg F. Gallagher  
**Title:** Standards and Achievement Director  
**Organization:** North Dakota Department of Public Instruction  
**Address:** 600 E. Boulevard Avenue, Bismarck, ND 58505-0440  
**Telephone:** (701) 328-1838  
**Email:** ggallagher@nd.gov  
**Contract Period:** 1997–2015

CTB McGraw-Hill facilitated the North Dakota Language Arts Standard Setting in Bismarck, North Dakota, on November 13-16, 2011. Panelists used the Bookmark Standard Setting Procedure (BSSP; Lewis, Mitzel, & Green, 1996) to recommend standards for seven grades of the North Dakota State Assessment (NDSA) in Language Arts: Grades 3-8 and 11. This report describes the implementation of the BSSP to recommend cut scores, the recommendations from the across-grade discussion, and the development of Reporting achievement level descriptors (ALDs).

The North Dakota Department of Public Instruction (DPI) convened 79 participants from across North Dakota to recommend cut scores. The DPI divided the participants into seven grade groups, each with 10-15 participants. Within each grade group, the DPI divided participants into three tables that were balanced in terms of relevant demographic characteristics (e.g., gender geographic location). The participants in each grade group participated in three rounds of activities in which they recommended three cut scores that defined four achievement levels: Novice, Partially Proficient, Proficient, and Advanced.

The workshop was conducted over three and one-half days. The first half-day was devoted to Table Leader training, and three days for standard setting and ALD writing. The workshop consisted of training, orientation, three rounds of judgments, across-grade discussion, and Reporting ALD writing. North Dakota has received Full Approval Status with the United States Department of Education, which
includes evaluation of development and descriptors and standard setting process.

Greg F. Gallagher, Standards and Achievement Director, is available as a verbal reference to this work at 701-328-1838.

**CTB: Minnesota State Assessment**

**Contact:** Jennifer Dugan  
**Title:** Director of Student Assessment  
**Organization:** Minnesota Department of Education  
**Address:** 1500 Highway 36 West, Roseville, MN 55113  
**Telephone:** 651-582-8654  
**Email:** jennifer.dugan@state.mn.us  
**Contract Period:** 2006

On March 8-9, 2006, staff members from CTB/McGraw-Hill worked in collaboration with the Minnesota Department of Education (MDE) to conduct the achievement level description (ALD) and content sorting workshops in Minneapolis, Minnesota. Two separate committees of Minnesota educators were convened to define new achievement levels for the Minnesota Comprehensive Assessments – Series II (MCA-II) and operationalize them as per the assessments.

Thirty-seven Minnesota educators convened on March 8, 2006, to recommend the number and names of the achievement levels to be implemented in the MCA-II program, and then to write descriptions of these achievement levels using information taken from the Minnesota Academic Standards for Reading and Mathematics. Participants recommended four achievement levels for the MCA-II. The recommended names for these four achievement levels are:

- Does Not Meet the Standards,
- Partially Meets the Standards,
- Meets the Standards, and
- Exceeds the Standards,

Participants were divided into eight subgroups to write descriptions of each of these achievement levels. Separate groups worked on ALDs for Reading and Mathematics: within each content area, participants were divided into groups focusing on Grades 3 and 4, 5 and 6, 7 and 8, and High School. Participants referred to the Minnesota Academic Standards as they crafted the ALDs, and were given opportunities to discuss their descriptions across grades to promote grade-to-grade articulation in the expectations of students.

A separate committee of 43 Minnesota educators convened on March 9, 2006, to study the ALDs written by the previous committee. Participants worked in eight subgroups for the Content Sorting activity. Separate groups worked on Reading and Mathematics within each group; participants were divided into groups focusing on Grade 3, 5, 8, and High School.

After studying the ALDs, participants studied test items from their assigned grade and content area. Participants worked individually and in concert to determine which achievement level best described the knowledge, skills, and abilities measured by each item. Items were sorted into one of the four achievement levels—labeled during the Content Sorting workshop as 1st Level, 2nd Level, 3rd Level, and 4th Level—or into borderline categories between these levels, such as Borderline 1st/2nd. Participants’ item level judgments were captured using formal rating forms. Participants then wrote Target Student descriptions. These descriptions summarize the knowledge, skills, and abilities held by students on the cusp between two achievement levels. To do this, participants summarized the knowledge, skills, and
abilities measured by items they rated as on the borderline between achievement levels during the content sorting process.

Minnesota has received Full Approval Status with the United States Department of Education which includes evaluation of development and descriptors and standard setting process.

Jennifer Dugan, Director of Student Assessment, is available as a verbal reference to this work at 651-582-8654.

**CTB: Qatar Assessment**

**Contact:** Dr. Hamda Al-Sulaiti  
**Title:** Director, Evaluation Institute  
**Organization:** Supreme Education Council  
**Address:** PO Box 35111, E Ring Road, Doha, Qatar  
**Phone:** +974 4455-9401  
**Email:** hamda.alsulaiti@sec.gov.qa

**Contract Period:** March 2011 to February 2012

In Spring 2009 cut scores (passing scores) were calculated for the Qatar Senior School Certificate (QSSC) assessments in Arabic, Biology, Chemistry, English, Islamic Studies, Mathematics, and Physics. Two cut scores were used to define three performance levels: Low, Satisfactory, and High. The QSSC represents a high-stakes testing program given to Grade 12 students before their exit from high school.

Performance level descriptions (PLDs) were written by the Evaluation Institute, and summarized the knowledge, skills, and abilities expected of students in each performance level. As part of a week-long workshop facilitated by CTB, content experts from the Evaluation Institute worked independently and in concert to examine the QSSC content standards, define the expected characteristics of students in each performance level, and develop PLDs that could be used as part of the QSSC standard setting.

Selected teachers from each of the independent schools were gathered together to study these PLDs and discuss their expectations of students. CTB led these teachers through this examination of the PLDs and, as needed, worked with the teachers to refine the PLDs to promote clarity, concision, and overall usefulness by Qatari educators.

The Borderline Groups procedure (Livingston & Zieky, 1982) was used to determine the cut scores for the QSSC. In this procedure, teachers studied lists of their students, examined the PLDs, and indicated which performance level best described each student. The performance level rating made by each teacher of each student was matched with the test score earned by the student on the QSSC, and the cut scores were then calculated using logistic regression. Through this process, educators from across the State of Qatar understood the connections between the various elements of the QSSC performance standards: the performance levels, the PLDs, the cut scores, and the ultimate classification of students into performance levels.

The scale score ranges associated with each of the performance levels for the QSSC, as determined through the Borderline Groups procedure, were determined. The percentage of students classified in each performance level (impact data) on the spring 2009 administration of the QSSC were also presented as part of the scope of the work.

Dr. Hamda Al-Sulait, Director, Evaluation Institute, is available if you would like an international verbal reference to this work, at +974 4455-9401.
College Board: Maine College and Career Readiness

Name of Contact: Dan Hupp, Director of Standards and Assessment
Telephone: 207-624-6827
E-Mail: dan.hupp@maine.gov

Time Period: 2004 to present

Description of Program and Services: The College Board, working in partnership with the state of Maine, has provided college and career access to Maine’s students through multiple pathways and rigorous programs of study. These pathway programs increased awareness of, preparation for, and connection to college and career success for Maine’s students while also ensuring continued compliance with NCLB requirements. The state of Maine recognized the challenge of preparing students to meet 21st century learning challenges. By addressing these challenges, students are prepared with Common Core learning standards that lead to success in school and future jobs. The comprehensive statewide approach to measure these standards and skills include the delivery of the following assessments in Maine:

- In fall 2004, all public high school began offering the PSAT/NMSQT to their sophomores.
- In spring 2006, the Maine Department of Education began administering the SAT to all public high school juniors, not only to meet federal testing requirements, but also to increase awareness of, and opportunities for, higher education among Maine’s students.
- In fall 2007, the PSAT/NMSQT program was expanded to include all public high school juniors.
- The state of Maine has long offered the College Board’s Advanced Placement Program (AP) course in English, math, science, history and other areas to its public high school students.
- AP course became available online during the 2007-08 academic year, making college-level studies available to even more of Maine’s high school students.

Please contact Dan Hupp for verbal reference about the Maine’s College and Career Readiness Program at 207-624-6827.

D. Contractor Intake Form

Vendors must submit a completed Contractor Intake Form, Exhibit E. Please sign and include any attachments that are necessary.

Provided on the following page is a fully executed Contractor Intake Form, Exhibit E, signed by CTB/McGraw-Hill's Chief Operating Officer, Sandor Z. Nagy.
Appendix A: Schedule
## SBAC RFP NO. 12
### Work Breakdown Structure (WBS)

<table>
<thead>
<tr>
<th>Task #</th>
<th>Task Name</th>
<th>Owner</th>
<th>Duration</th>
<th>Start Date</th>
<th>Finish Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SBAC RFP NO. 12</td>
<td></td>
<td>134d</td>
<td>05/09/12</td>
<td>11/15/12</td>
</tr>
<tr>
<td>2</td>
<td>START SBAC RFP NO. 12</td>
<td></td>
<td>1d</td>
<td>05/09/12</td>
<td>05/09/12</td>
</tr>
<tr>
<td>3</td>
<td>Program Management</td>
<td></td>
<td>130d</td>
<td>05/15/12</td>
<td>11/15/12</td>
</tr>
<tr>
<td>4</td>
<td>Final Detailed Schedule</td>
<td></td>
<td>1d</td>
<td>05/15/12</td>
<td>05/15/12</td>
</tr>
<tr>
<td>5</td>
<td>Submit Final Detailed Schedule within One Week of Contract Award</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>05/15/12</td>
<td>05/15/12</td>
</tr>
<tr>
<td>6</td>
<td>Contractor Kick-Off Meeting</td>
<td></td>
<td>1d</td>
<td>06/12/12</td>
<td>06/12/12</td>
</tr>
<tr>
<td>7</td>
<td>Contractor Kick-Off Meeting</td>
<td></td>
<td>1d</td>
<td>06/12/12</td>
<td>06/12/12</td>
</tr>
<tr>
<td>8</td>
<td>Weekly Status Reports</td>
<td></td>
<td>109d</td>
<td>06/14/12</td>
<td>11/15/12</td>
</tr>
<tr>
<td>9</td>
<td>Weekly Status Reports</td>
<td>Lindy Wienand</td>
<td>109d</td>
<td>06/14/12</td>
<td>11/15/12</td>
</tr>
<tr>
<td>10</td>
<td>TAC Meeting Presentation</td>
<td></td>
<td>12d</td>
<td>07/11/12</td>
<td>07/26/12</td>
</tr>
<tr>
<td>11</td>
<td>Draft Technical Report Available</td>
<td>Karla Egan</td>
<td>1d</td>
<td>07/11/12</td>
<td>07/11/12</td>
</tr>
<tr>
<td>12</td>
<td>Present Preliminary Products at TAC Meeting in Denver</td>
<td>Karla Egan</td>
<td>2d</td>
<td>07/25/12</td>
<td>07/26/12</td>
</tr>
<tr>
<td>13</td>
<td>ALD Development</td>
<td></td>
<td>9d</td>
<td>06/12/12</td>
<td>06/24/12</td>
</tr>
<tr>
<td>14</td>
<td>CTB Pre-Work</td>
<td></td>
<td>9d</td>
<td>06/12/12</td>
<td>06/24/12</td>
</tr>
<tr>
<td>15</td>
<td>Define Methodology</td>
<td>Karla Egan</td>
<td>1d</td>
<td>06/13/12</td>
<td>06/13/12</td>
</tr>
<tr>
<td>16</td>
<td>Recruit Content and/or Curriculum Specialists in ELA &amp; Math</td>
<td>Lindy Wienand</td>
<td>5d</td>
<td>06/12/12</td>
<td>06/18/12</td>
</tr>
<tr>
<td>17</td>
<td>Convene the Committee Members to Follow the Approved Methods for Developing the Initial ALDs</td>
<td>Lindy Wienand</td>
<td>5d</td>
<td>06/20/12</td>
<td>06/24/12</td>
</tr>
<tr>
<td>18</td>
<td>Deliverables - Development of General and Grade-Specific Initial ALDs in ELA and Math (Grades 3-8 &amp; 11)</td>
<td>Lindy Wienand</td>
<td>96d</td>
<td>06/14/12</td>
<td>10/29/12</td>
</tr>
<tr>
<td>19</td>
<td>Draft Initial ALDs - (Version 1)</td>
<td></td>
<td>25d</td>
<td>06/14/12</td>
<td>07/19/12</td>
</tr>
<tr>
<td>20</td>
<td>Draft Initial ALDs (version 1) Based on the Development Work with Committee Members</td>
<td>Darin Homer</td>
<td>3d</td>
<td>06/25/12</td>
<td>06/27/12</td>
</tr>
<tr>
<td>21</td>
<td>Email Notification 10 Days Prior to Review - (Contractor for Item/Task Specs, Item/Task Writing/Review--Pilot, Test Design and CAT Specs, Accessibility &amp; Accommodations Policy Guidelines, Psychometric Services &amp; Report Development)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>06/14/12</td>
<td>06/14/12</td>
</tr>
<tr>
<td>22</td>
<td>Email Notification 3 days Prior to Review (Item Development, Test Design, Performance Tasks, Accessibility and Accommodations, and Validation and Psychometrics Work Groups (including Lead Psychometrician and EC Liaison)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>06/25/12</td>
<td>06/25/12</td>
</tr>
<tr>
<td>23</td>
<td>Review</td>
<td>Contractors/Work Groups</td>
<td>10d</td>
<td>06/28/12</td>
<td>07/12/12</td>
</tr>
<tr>
<td>24</td>
<td>Implement Customer Comments</td>
<td>Darin Homer</td>
<td>2d</td>
<td>07/13/12</td>
<td>07/16/12</td>
</tr>
<tr>
<td>25</td>
<td>Provide Status of Version 1 Deliverable to CONSORTIUM (within a week after review)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>07/19/12</td>
<td>07/19/12</td>
</tr>
<tr>
<td>26</td>
<td>Revised Draft Initial ALDs - (Version 2)</td>
<td>Darin Homer</td>
<td>3d</td>
<td>07/17/12</td>
<td>07/19/12</td>
</tr>
<tr>
<td>27</td>
<td>Revised Draft Initial ALDs (version 2) Based on Feedback from the Contractors and Item Development, Test Design, Performance Task, Accessibility and Accommodations, and Validation and Psychometrics Work Groups (including Lead Psychometrician and EC Liaison)</td>
<td>Darin Homer</td>
<td>3d</td>
<td>07/06/12</td>
<td>08/09/12</td>
</tr>
<tr>
<td>28</td>
<td>Email Notification 10 Days Prior to Review (CONSORTIUM Higher Education State Membership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>07/06/12</td>
<td>07/06/12</td>
</tr>
<tr>
<td>29</td>
<td>Email Notification 3 days Prior to Review (CONSORTIUM Executive Committee for Approval to Share with States)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>07/17/12</td>
<td>07/17/12</td>
</tr>
<tr>
<td>30</td>
<td>Reminder Notification 2 days Prior to Review (CONSORTIUM Higher Education State Membership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>07/18/12</td>
<td>07/18/12</td>
</tr>
</tbody>
</table>
### SBAC RFP NO. 12

#### Work Breakdown Structure (WBS)

<table>
<thead>
<tr>
<th>Task #</th>
<th>Task Name</th>
<th>Owner</th>
<th>Duration</th>
<th>Start Date</th>
<th>Finish Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>Review</td>
<td>CONSORTIUM Higher Ed/CONSORTIUM Executive Committee</td>
<td>10d</td>
<td>07/20/12</td>
<td>08/02/12</td>
</tr>
<tr>
<td>32</td>
<td>Implement Customer Comments</td>
<td>Darin Homer</td>
<td>2d</td>
<td>08/03/12</td>
<td>08/06/12</td>
</tr>
<tr>
<td>33</td>
<td>Provide Status of Version 2 Deliverable to CONSORTIUM (within a week after review)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/09/12</td>
<td>08/09/12</td>
</tr>
<tr>
<td>34</td>
<td>Revised Draft Initial ALDs - (Version 3)</td>
<td></td>
<td>25d</td>
<td>07/26/12</td>
<td>08/29/12</td>
</tr>
<tr>
<td>35</td>
<td>Revised Draft Initial ALDs (version 3) Based on CONSORTIUM Feedback from the CONSORTIUM Executive Committee and the CONSORTIUM Higher Education State Membership</td>
<td>Darin Homer</td>
<td>2d</td>
<td>08/07/12</td>
<td>08/08/12</td>
</tr>
<tr>
<td>36</td>
<td>Email Notification 10 Days Prior to Review (CONSORTIUM’s TAC)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>07/26/12</td>
<td>07/26/12</td>
</tr>
<tr>
<td>37</td>
<td>Reminder Notification 2 days Prior to Review (CONSORTIUM’s TAC)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/07/12</td>
<td>08/07/12</td>
</tr>
<tr>
<td>38</td>
<td>Email Notification 10 days Prior to Review (CONSORTIUM K-12 State Membership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/26/12</td>
<td>07/26/12</td>
</tr>
<tr>
<td>39</td>
<td>Reminder Notification 2 days Prior to Review (CONSORTIUM K-12 State Membership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/07/12</td>
<td>08/07/12</td>
</tr>
<tr>
<td>40</td>
<td>Review</td>
<td>CONSORTIUM’s TAC/CONSORTIUM K-12 State Membership</td>
<td>10d</td>
<td>08/09/12</td>
<td>08/22/12</td>
</tr>
<tr>
<td>41</td>
<td>Implement Customer Comments</td>
<td>Darin Homer</td>
<td>2d</td>
<td>08/23/12</td>
<td>08/24/12</td>
</tr>
<tr>
<td>42</td>
<td>Provide Status of Version 3 Deliverable to CONSORTIUM (within a week after review)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/29/12</td>
<td>08/29/12</td>
</tr>
<tr>
<td>43</td>
<td>Revised Draft Initial ALDs - (Version 4)</td>
<td></td>
<td>25d</td>
<td>08/15/12</td>
<td>09/19/12</td>
</tr>
<tr>
<td>44</td>
<td>Revised Draft Initial ALDs (version 4) Based on Feedback from the CONSORTIUM’s TAC and the CONSORTIUM K-12 State Membership</td>
<td>Darin Homer</td>
<td>2d</td>
<td>08/27/12</td>
<td>08/28/12</td>
</tr>
<tr>
<td>45</td>
<td>Email Notification 10 Days Prior to Review (PARCC)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/15/12</td>
<td>08/15/12</td>
</tr>
<tr>
<td>46</td>
<td>Reminder Notification 2 days Prior to Review (PARCC)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/27/12</td>
<td>08/27/12</td>
</tr>
<tr>
<td>47</td>
<td>Email Notification 10 days Prior to Review (Other Key Stakeholder Groups, as Appropriate)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/15/12</td>
<td>08/15/12</td>
</tr>
<tr>
<td>48</td>
<td>Reminder Notification 2 days Prior to Review (Other Key Stakeholder Groups, as Appropriate)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>08/27/12</td>
<td>08/27/12</td>
</tr>
<tr>
<td>49</td>
<td>Review</td>
<td>PARCC/Other Key Stakeholder Groups</td>
<td>10d</td>
<td>08/29/12</td>
<td>09/12/12</td>
</tr>
<tr>
<td>50</td>
<td>Implement Customer Comments</td>
<td>Darin Homer</td>
<td>2d</td>
<td>09/13/12</td>
<td>09/14/12</td>
</tr>
<tr>
<td>51</td>
<td>Provide Status of Version 4 Deliverable to CONSORTIUM (within a week after review)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>09/19/12</td>
<td>09/19/12</td>
</tr>
<tr>
<td>52</td>
<td>Revised Draft Initial ALDs - (Version 5)</td>
<td></td>
<td>18d</td>
<td>09/14/12</td>
<td>10/09/12</td>
</tr>
<tr>
<td>53</td>
<td>Revised Draft Initial ALDs (version 5) Based on Revisions that Enhance Desired Consistency with PARCC and Feedback from Key Stakeholder Groups</td>
<td>Darin Homer</td>
<td>2d</td>
<td>09/17/12</td>
<td>09/18/12</td>
</tr>
<tr>
<td>54</td>
<td>Email Notification 3 Days Prior to Review (Validation and Psychometrics Work Group Leadership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>09/14/12</td>
<td>09/14/12</td>
</tr>
<tr>
<td>55</td>
<td>Email Notification 3 days Prior to Review (Executive Committee)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>09/14/12</td>
<td>09/14/12</td>
</tr>
</tbody>
</table>
### SBAC RFP NO. 12

#### Work Breakdown Structure (WBS)

<table>
<thead>
<tr>
<th>Task #</th>
<th>Task Name</th>
<th>Owner</th>
<th>Duration</th>
<th>Start Date</th>
<th>Finish Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>56</td>
<td>Review</td>
<td>Validation and Psychometrics Work Group Leadership/Executive Committee</td>
<td>10d</td>
<td>09/19/12</td>
<td>10/02/12</td>
</tr>
<tr>
<td>57</td>
<td>Implement Customer Comments</td>
<td>Darin Homer</td>
<td>2d</td>
<td>10/03/12</td>
<td>10/04/12</td>
</tr>
<tr>
<td>58</td>
<td>Provide Status of Version 5 Deliverable to CONSORTIUM (within a week after review)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>10/09/12</td>
<td>10/09/12</td>
</tr>
<tr>
<td>59</td>
<td>Final Initial ALDs - (Version 6)</td>
<td></td>
<td>25d</td>
<td>09/25/12</td>
<td>10/29/12</td>
</tr>
<tr>
<td>60</td>
<td>Final Draft Initial ALDs (version 6) Based on Approval by CONSORTIUM Executive Committee, Lead Psychometrician, and Validation and Psychometrics Work Group Leadership</td>
<td>Darin Homer</td>
<td>2d</td>
<td>10/05/12</td>
<td>10/08/12</td>
</tr>
<tr>
<td>61</td>
<td>Email Notification 10 Days Prior to Review (CONSORTIUM K-12 State Membership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>09/25/12</td>
<td>09/25/12</td>
</tr>
<tr>
<td>62</td>
<td>Email Notification 3 days Prior to Review (V &amp; P Work Group)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>10/04/12</td>
<td>10/04/12</td>
</tr>
<tr>
<td>63</td>
<td>Reminder Notification 2 days Prior to Review (CONSORTIUM K-12 State Membership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>10/05/12</td>
<td>10/05/12</td>
</tr>
<tr>
<td>64</td>
<td>Review</td>
<td>V &amp; P Work Group/CONSORTIUM K-12 State Membership</td>
<td>10d</td>
<td>10/09/12</td>
<td>10/22/12</td>
</tr>
<tr>
<td>65</td>
<td>Implement Customer Comments</td>
<td>Darin Homer</td>
<td>2d</td>
<td>10/23/12</td>
<td>10/24/12</td>
</tr>
<tr>
<td>66</td>
<td>Provide Status of Version 6 Deliverable to CONSORTIUM (within a week after review)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>10/29/12</td>
<td>10/29/12</td>
</tr>
<tr>
<td>67</td>
<td>Technical Report</td>
<td></td>
<td>87d</td>
<td>07/17/12</td>
<td>11/15/12</td>
</tr>
<tr>
<td>68</td>
<td>Technical Report</td>
<td></td>
<td>87d</td>
<td>07/17/12</td>
<td>11/15/12</td>
</tr>
<tr>
<td>70</td>
<td>Email Notification 10 Days Prior to Review (CONSORTIUM's TAC)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>10/16/12</td>
<td>10/16/12</td>
</tr>
<tr>
<td>71</td>
<td>Email Notification 3 days Prior to Review (Validation and Psychometrics Work Group Leadership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>10/25/12</td>
<td>10/25/12</td>
</tr>
<tr>
<td>72</td>
<td>Reminder Notification 2 days Prior to Review (CONSORTIUM K-12 State Membership)</td>
<td>Lindy Wienand</td>
<td>1d</td>
<td>10/26/12</td>
<td>10/26/12</td>
</tr>
<tr>
<td>73</td>
<td>Review</td>
<td>Validation and Psychometrics Work Group Leadership/CONSORTIUM K-12 State Membership</td>
<td>10d</td>
<td>10/30/12</td>
<td>11/12/12</td>
</tr>
<tr>
<td>74</td>
<td>Implement Customer Comments</td>
<td>Darin Homer</td>
<td>2d</td>
<td>11/13/12</td>
<td>11/14/12</td>
</tr>
<tr>
<td>75</td>
<td>Provide Final Tech Report Deliverable to CONSORTIUM</td>
<td>Karla Egan</td>
<td>1d</td>
<td>11/15/12</td>
<td>11/15/12</td>
</tr>
<tr>
<td>76</td>
<td>END SBAC RFP NO. 12</td>
<td></td>
<td>1d</td>
<td>11/15/12</td>
<td>11/15/12</td>
</tr>
</tbody>
</table>
The operational work that Drs. Egan and Schneider conducted along with their research on ALDs directly led to the framework that was proposed for this RFP. Even though this framework is new (Egan, Schneider, & Ferrara, 2012), the team has started implementing it in North Dakota when target and reporting ALDs were written for that contract. The team's work on standard setting and ALDs, ALD development, and ALD validation has been presented at the American Educational Research Association and the National Council on Measurement in Education. In addition, it has been published in the Peabody Journal of Education, Handbook of Accessible Achievement Tests for All Students, and Setting Performance Standards: Foundations, Methods, and Innovations.

For an in-depth discussion of the ALD framework, please see the Performance Level Descriptors chapter written by Dr. Karla Egan, M. Christina Schneider, and Dr. Steve Ferrara on the following pages.

Reference:
Performance level descriptors (PLDs) are ubiquitous in K-12 student achievement testing, appearing on websites of state departments of education, on score reports, and in test interpretation guides. PLDs define the knowledge, skills, and processes (KSPs) of students at specified levels of achievement and often include input from policy makers, stakeholders, and content experts. An uneasy relationship frequently exists between the content expectations expressed in the PLDs and the policy they are supposed to represent. This may be due to the way that PLDs are developed. Instead of being a conceptualization point that connects and organizes a state's standards and assessment system (Bejar, Braun, & Tannenbaum, 2007; Hansche, 1998), PLDs are often developed immediately prior to standard setting (Cizek & Bunch, 2007; Hambleton & Pitoniak, 2006), resulting in PLDs that have neither guided test development nor been guided by the content of test. The consequence is that the expected KSPs of the PLDs may not reflect the actual KSPs as represented by the final cut scores.

To rectify this situation, a framework is proposed for developing PLDs so that they are foundational to test development. This framework is based on the intended uses of PLDs for (a) test development, (b) standard setting guidance, and (c) score interpretation. The proposed framework comprises four types of interrelated PLDs that are meant to explicate the construct being measured and the intended test score interpretations in regard to that construct. That is, the framework is meant to drive a construct-centered approach to assessment design (Messick, 1994). The four types of proposed PLDs include:

- **Policy PLDs** that articulate the policy makers' vision of the goals and rigor for the final performance standards;
- **Range PLDs** that are created by test developers to identify which aspects of items align to a particular performance level in regard to the cognitive and content rigor that has been defined;
- **Target PLDs** that are created in conjunction with the Range PLDs and are used by standard setting panelists to represent just how much a threshold or borderline student in a particular performance level should know and be able to do; and
- **Reporting PLDs** that are developed by a sponsoring agency once cut scores are finalized, to define the appropriate, intended interpretations regarding what a test score means.
Such a linked system of PLDs would serve to define the construct that is being measured and to describe what students should know and be able to do in relation to the construct. When a clear definition of the target of measurement exists, a more fully aligned assessment system is created.

This chapter is divided into the following sections. The first section defines various standard setting terms that may have different meanings in other articles or contexts. Next, the historical development of PLDs is traced in concert with their uses. Then, current practice involving PLDs is reviewed. This is followed by the section that describes threats to the validity of PLDs. The main portion of the chapter presents the proposed PLD framework. The chapter ends with a discussion on possible future trends in the development of PLDs and their uses.

PLDs and Standard Setting Terminology

The term performance standard is used to describe an organization's high-level expectations regarding what students should know and be able to do. Performance levels provide for classification of examinees' test performance according to ordered evaluative categories (Haertel, 1999). If test performance is labeled Proficient instead of Basic, it implies that a student possessed more knowledge and skills than her or his peers who attained the Basic status. Cut scores are the specific points on the test scale that separate students into performance levels. Cut scores are the numeric operationalization of the PLDs. The term standard setting is often used to refer to the activity whereby educators or other qualified participants recommend cut scores. In this chapter, the term standard setting is used to refer to a multi-phase process used by a sponsoring agency to define PLDs and their accompanying cut scores. In this chapter, the term cut score recommendation workshop is used to refer to a workshop where educators and appropriate stakeholders recommend cut scores.

Emergence of PLDs and Their Associated Uses in the Literature

The term performance level descriptor began appearing frequently in the literature in the 1990s; nonetheless, there are precursors to PLDs which can be found in the psychometric literature. For example, the early advocates of criterion-referenced tests (CRTs) touted their role in helping teachers focus their instruction on "those items which can be proved to be one of importance in preparation for intelligent living and future usefulness in life" (Monroe, DeVoss, & Kelly, 1924, as cited in Zieky, 1995, p. 4). This type of score reporting is similar to reporting the KSIs in PLDs on modern score reports. In tracing the development of PLDs, three different, yet intended, uses emerge in the literature, including (a) standard setting guidance, (b) score interpretation, and (c) test development.

Standard Setting Guidance

Bourque (2000) has asserted that the "most important function" of the PLD "is to provide a mental framework or structure for standard setting panelists" (p. 8). The PLDs are the heart of the standard setting process; thus, it is not surprising that much of the research literature focuses on this aspect of their use. What is surprising is the length of time it took for formal PLDs to be used as part of the standard setting process.

Nedelsky's (1954) introduction of the F-D student represents one of the first efforts to formalize the student for whom cut scores were being set. In essence, Nedelsky asked panelists to eliminate the distractors from an item that the D-student would know to be incorrect. After this, the probability that they would guess the correct answer was computed for each item, and
the probabilities were summed to find the cut score. Over time, the concept of the F-D student morphed into a discussion of the minimally competent examinee or borderline candidate. In the 1970s, there was widespread use of minimum competency tests by states (see Garth & Perkins, 1979, for a comprehensive summary of these programs) and by licensure and certification programs. Researchers recommended that the standard setting panelists discuss the necessary skills expected of the minimally competent examinee (or minimally competent candidate) (Nassif, 1979). An early article by Perkins (1979) attempted to formalize procedures for defining the minimal competency. She also documented the steps that some states took to formally define the minimally competent examinee for their state tests.

Nonetheless, there seems to have been disagreement in the field during the 1980s regarding the need to define minimal competency. In a 1984 article, Cross, Impara, Frary, and Jaeger investigated Jaeger's (1978) standard setting method and observed that "... a desirable feature of this method is that it avoids the troublesome problem of defining what is meant by 'minimally competent'" (p. 158). Throughout the 1980s, the standard setting literature referred to the need to define the minimally competent candidate, student, or borderline group; however, these definitions were neither written down nor formalized, and they were usually no more than a conversation among the panelists (e.g., Norcini, Lipner, Langdon, & Strecker, 1987; Norcini & Shea, 1992; Norcini, Shea, & Kanya, 1988). Even so, it was this use of the minimally competent candidate that led to an important period in the modern conceptualization of PLDs and its ultimate centralized role in standard setting.

Written PLDs were rarely used during actual cut score recommendation workshops until the 1990s (Bourque, 2000), when the National Assessment Governing Board (NAGB) decided to adopt performance standards in May of 1990 (see Vinovskis, 1998, for an informative review of the political machinations behind the rise of achievement standards on the National Assessment of Educational Progress [NAEP]). According to Brown (2000):

NAGB believed there were procedures to establish performance expectations using input from a variety of interested stakeholders, such as teachers, principals, curriculum specialists, and business leaders. The process had been satisfactory for establishing licensure cut-scores, so it was assumed that such a process could be expanded to establish multiple performance levels that would be useful in reporting NAEP results. (p. 14)

States soon followed NAGB's lead and began using a PLD process when developing cut scores for their own statewide student assessments.

The 1992 NAEP standard setting represented the first time that formal PLDs were produced to guide NAEP standard setting panelists (Bourque, 2000). Beginning with generic policy-based descriptions, standard setting panels reviewed sample items (NAEP used the term exercise pool) and exemplar responses for constructed-response items to operationalize the policy descriptions for the purposes of standard setting. These PLDs were continually refined throughout the standard setting process, and they were designed not to be specific to the NAEP item pool.

Following this lead from NAGB, a philosophical shift appears in the literature in the 1990s, and researchers increasingly discussed formalizing the definitions associated with performance levels prior to or at the beginning of a cut score recommendation workshop. Instead of mere discussions of the KSPs of the minimally competent examinee, formal processes that documented these KSPs were undertaken (Mills, Melican, & Ahluwalia, 1991; Plake, Impara, & Potenza, 1994).

In addition to formalizing the process of writing PLDs, researchers began experimenting with the effect PLDs have on panelists' judgments. Researchers found that when panelists used
PLDs to guide their judgments, the variability of panelists' ratings decreased (Giraud, Impara, & Plake, 2005; Hurtz & Auerbach, 2003; Impara, Giraud, & Plake, 2000). This line of investigation provided a strong research base for the content-based PLDs that were eventually required for standard setting through the peer review process associated with the No Child Left Behind (NCLB; 2002) legislation. The NCLB peer review process moved states away from using generic policy-based PLDs for standard setting to using content-based competency PLDs.

Score Interpretation

The idea that PLDs may be used for reporting purposes to provide the general public with appropriate interpretations may be traced to the use of the scale anchoring process used with NAEP. From the mid-1980s to early 1990s, NAGB used a scale anchoring process to describe student performance at specific points (anchors) along the NAEP scale using items from the NAEP test (Jaeger, 1998). These anchor descriptions were intended to give the public a more concrete idea of what students along the scale could do.

In 1991, NAEP released the first set of PLDs for reporting purposes (Koretz & Deibert, 1993). Unfortunately, the vague language used in these PLDs was found to be misinterpreted when the PLDs were reported by the press (Koretz & Deibert, 1993). When NAEP tried again with the 1992 assessments, researchers found inconsistencies between the PLDs that described what students should know and actual student performance (Burstein et al., 1996). Nonetheless, the NAEP PLDs were popular with the public because they gave the end user an idea of what the student could accomplish (Brown, 2000).

Throughout the rise of PLD use for the NAEP, there was ongoing debate regarding what the performance levels should reflect. It was unclear, even to the proponents of performance standards, whether they should reflect what students should be able to do or what they can do (Vinovskis, 1998). This murky beginning has yet to be resolved satisfactorily. Although not articulated, this debate reflects the different uses of PLDs: When PLDs are used for reporting purposes, then proponents believe that PLDs should reflect the KSPs actually possessed by students (Burstein et al., 1996; Crane & Winter, 2006; Linn, 1998). When PLDs are used for standard setting purposes, proponents advocate that the PLDs should reflect what students should know and be able to do (Bourque, 2000; Perie, 2008).

As a result of the NCLB peer review requirements, states administering K-12 tests must use PLDs for both standard setting and score reporting. Frequently, organizations want to use the same PLDs to serve both purposes; however, standard setting PLDs frequently do not reflect the final cut scores, and reporting PLDs are not available to guide standard setting. This has led to much discussion regarding when PLDs should be finalized (Bourque, 2000; Crane & Winter, 2006; Hansche, 1998; Mercado & Egan, 2005; Perie, 2008; Schneider, Egan, Siskind, Brailsford, & Jones, 2009). One point of view holds that PLDs should be written and finalized prior to the cut score recommendation workshop (Bourque, 2000; Perie, 2008) so that the expectations of student performance are known and may be disseminated prior to testing. Bourque (2000) has noted that in the case of NAEP, the a priori PLDs are not suitable for reporting results to the public, and she has recommended the use of exemplar items to communicate the KSPs of students in each performance level.

Lewis and Green (1997) argued that PLDs need not be written until the cut score recommendation workshop is conducted and cut scores are finalized. Although this approach ensures good alignment between the cut scores and PLDs, it does not provide standard setting panelists with the content-based policy guidance that they need at the cut score recommendation workshop. In addition, if a single test form is used to develop such PLDs, they may not generalize to
the actual performance of new items when future test forms are implemented over time (Schneider, Egan, Kim, & Brandstrom, 2008). To this end, Crane and Winter (2006) have recommended that PLDs be monitored and updated as new tests are brought online so that the PLDs continue to reflect actual student performance.

Test Development

While some researchers were examining the evolving use of PLDs for standard setting and reporting purposes, others were making arguments regarding the utility of PLDs for guiding test development (Bejar et al., 2007; Kane, 1994; Mills & Jaeger, 1998; Pellegrino, Jones, & Mitchell, 1999). Egan, Schneider, and Ferrara (2011) contend that PLDs developed in conjunction with the development of a state's content standards (or shortly thereafter) can be used as a tool to aid in the development of test blueprints, item specifications, and the items themselves.

Although the call for the use of PLDs in test development is not new, it has only recently been attempted. Ferrara, Svetina, Skucha, and Murphy (2009) referenced one state testing program for which item writers developed a mathematics test using both content standards and PLDs with less than stellar results. The authors found that only 40% of Grade 3 items, 54% of Grade 4 items, and 14% of Grade 5 items aligned to the item writers' intended performance levels once the cut scores were determined.

In related work, the College Board began using evidence-centered design (ECD; Mislevy, Steinberg, & Almond, 2002) for the development of their Advanced Placement tests. College Board staff and consultants detailed the use of ECD to create PLDs (Plake, Huff, & Reshetar, 2010) and described how ECD was used to guide the development of the test blueprint and test items (Hendrickson, Huff, & Luecht, 2010). Here, again, initial attempts to write items to specific performance levels were not as successful as hoped (K. Huff, personal communication, October 2, 2010). The results from these studies suggest that extensive efforts are needed to reconceptualize the process of writing items and building tests.

When item writers are asked to target items to particular performance levels, it is assumed that the item writers understand the factors influencing item difficulty to such a degree that they can target items to particular regions of the test characteristic curve (TCC). Results from research on the modified Angoff standard setting method have been mixed regarding how accurately standard setting panelists estimate the difficulty of items (Fehrmann, Woehr, & Arthur, 1991; Impara & Plake, 1998; Plake & Impara, 2001). Furthermore, research in the field of social psychology suggests that, while training helps, people have difficulty making accurate, probability judgments (Nickerson, 2004; Pious, 1993). Until the profession has a comprehensive, empirically validated framework of factors that affect item difficulty, item writers' attempts to target particular performance levels will be hindered (Schneider, Huff, Egan, Tully, & Ferrara, 2010).

This section summarized three uses of PLDs for standard setting guidance, score interpretation, and test development. The next section examines the way states approach PLDs for score interpretation.

Current Approaches to PLDs for Score Interpretation

Throughout the 1990s, states began the process of creating PLDs to accompany their content standards. Once NCLB was enacted in 2002, all states were required to develop PLDs and to use those PLDs for standard setting and score reporting; however, the form and functionality of the PLDs was not mandated and, hence, departments of education were left to their own devices to create PLDs. At the time, there were few high-level guidelines for creating PLDs (beyond descriptions
of procedures used by NAGB to produce PLDs for NAEP and the NAEP PLDs to use as models)_.
Hansche (1998) provided a checklist that advised on the usability and clarity of PLDs. Crane and Winter (2006) advised on practical concerns for writing PLDs (such as the diversity of the PLD-writing panel). In addition, they discussed the need for the coherence of PLDs across grades and content areas. And, in 2008, Perie provided guidance on a system of PLDs that included policy definitions for PLDs, full PLDs that define content KSPs, and final PLDs that reflect the content KSPs as defined by the cut scores. Rabinowitz and his colleagues provided advice that the developers of PLDs consider "the assessment, instructional, and reporting uses to be made of these descriptors" (Rabinowitz, Roeber, Schroeder, & Sheinker, 2006, P-26).

Even with this guidance, the composition and usability of PLDs varies greatly across states. This section summarizes the ways that state departments of education currently define Proficient, with specific attention paid to the structure and content of those PLDs. For this analysis, PLDs for Grade 6 mathematics from 47 of the 50 states plus the District of Columbia were examined. The 48 PLDs were located (not always easily) on state websites or were e-mailed by staff from the state departments. This examination is limited to publicly available PLDs. In theory, these PLDs should be used for score interpretation. An examination of each state's PLDs used for standard setting guidance is beyond the scope of this chapter.

Most states provided grade- and content-specific PLDs for Grade 6 mathematics, while a small number provided only high-level policy descriptors. As we describe later, seven of the state PLDs explicitly describe "typical" students, while several imply that they describe borderline students. Table 5.1 summarizes the labels for Proficient (e.g., "Proficient," "Meets the Standard"), the content and syntactic structure of the definitions, the defining phrase for Proficient (e.g., "satisfactory performance," "grasp of the curriculum"), and the warranted inference. Analysis of the content and syntactic structure of the definitions focuses on whether (a) nouns and noun phrases are present to refer to knowledge or understanding, (b) verb phrases with a list of content area skills are present, and/or (c) a bulleted list of knowledge and skills based on the state assessment program content standards is present to support the noun or verb phrases. Warranted inferences are defined as (a) affirmative statements (i.e., what students know and can do), (b) probabilistic statements (i.e., what students are likely to know and be able to do), (c) hybrid affirmative-probabilistic statements, or (d) aspirational statements (i.e., what students should know and be able to do).

Labels for the Proficient Level

"Proficient" is the most widely used label in state assessment programs (28 states). This is likely due to NAEP, which has significantly influenced conceptions of and methods for defining performance standards. Three additional states refer to performing At Proficiency or at a numbered level (e.g., Level3). Eleven states use a variation on Meets the Proficient standard. Six additional states label the Proficient level as Mastery, Pass, Intermediate (among four or five levels), or Sufficient. Burt and Stapleton (2010) have shown that panelists have considerable variation in their understandings of the Proficient level label and the inferences that it engenders about what examinees know and can do.

Content of PLDs for Proficient

Generally speaking, PLDs most often include a paragraph that is intended as a general, introductory, and primary definition of each performance level. This paragraph is often supported
Table 5.1  Labels, Content and Syntactic Structure, Defining Phrases, and Warranted Inferences for PLDs for 47 State Assessment Programs plus the District of Columbia

<table>
<thead>
<tr>
<th>Label</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Proficient</td>
<td>28</td>
</tr>
<tr>
<td>Proficiency level/level x</td>
<td>3</td>
</tr>
<tr>
<td>Meets/met the standard</td>
<td>11</td>
</tr>
<tr>
<td>Master's</td>
<td>2</td>
</tr>
<tr>
<td>Pass</td>
<td>2</td>
</tr>
<tr>
<td>Intermediate</td>
<td></td>
</tr>
<tr>
<td>Sufficient</td>
<td></td>
</tr>
</tbody>
</table>

**Content and Syntactic Structure**

*Noun or Noun Phrase Refers to:*

- Performance                              4
- Knowledge                                 21
- Conceptual understanding, grade level concepts  2
- Grade level mastery (and in addition to...)  2
- N.A.                                     20

*Verb Phrase:*

- With list of skills                      39
- About performance                        4
- About logical reasoning, strategies, and errors  2
- About using concepts and understanding processes  2
- N.A.                                     2

*Supporting, bulleted list of knowledge and skills:*

- Knowledge and skills                     16
- N.A.                                     32

**Defining Phrase**

- Satisfactory/solid performance            9
- Competency, mastery, attainment, ability, progress  9
- Proficient                                2
- General/good/broad understanding, grasp of the curriculum  9
- Solves problems                           2
- Adequate knowledge/skills, applies knowledge/skills  9
- Partial success                           7
- N.A.                                     2

**Warranted Inference**

- Affirmative                               28
- Probabilistic                             8
- Both or hybrid                            10
- Aspirational                              2

N.A. = not applicable
by a bulleted list of content area skills. A syntactic structure for defining and describing performance levels has emerged organically over time. Nouns and noun phrases explicate knowledge and conceptual understanding described in PLDs. Examples of noun phrases from state PLDs include "knowledge of grade level content" and "mathematical concepts." In PLDs, verb phrases delineate skills identified in the assessment program's content standards, usually in bulleted or numbered form or in strings of phrases or sentences. Examples of verb phrases from state PLDs include "convert fractions" and "solve problems."

**Nouns and noun phrases to specify content and conceptual knowledge.** Twenty-eight of the 48 states refer to mathematics content knowledge and conceptual understanding, and most of these states (21) refer to understanding of grade-level mathematics concepts. Other states' definitions of Proficient refer to performance, mathematics knowledge, and grade-level mastery. Definitions of Proficient for 20 states do not refer to mathematics knowledge and conceptual understanding, indicating that these definitions rely almost exclusively on references to mathematics skills and broader mathematics processes.

**Verb phrases to specify skills-based knowledge.** The most common strategy for conveying mathematics content requirements to reach the Proficient level is to list skills in verb phrases. This strategy appears often in both introductory paragraphs and supporting, bulleted lists of skills. Thirty-nine of these states define Proficient by specifying mathematics skills in an introductory paragraph using verb phrases embedded in complete sentences. The phrases or sentences may simply be lists or formatted as a paragraph. Seven other state definitions refer to performance, logical reasoning and strategies, and using mathematics concepts and understanding mathematics processes. Only two states do not use verb phrases to list mathematics skills.

Sixteen states include lists (usually bulleted or numbered) of mathematics skills and knowledge to support the introductory definition of Proficient or in place of a general, introductory definition. References to mathematics knowledge and conceptual understanding are relatively infrequent compared to the highly frequent references to mathematics skills.

**Defining Phrase for Proficient**

PLDs are devices that communicate a state's intended definition of Proficient to the end users. The defining phrase explicates the intended or take-away message in definitions of Proficient. Because standard setting panelists' (and most likely other end users') interpretations of labels for Proficient vary considerably (see Burt & Stapleton, 2010), defining phrases are important. States' uses of defining phrases vary widely. Nine states each (36 altogether) refer to satisfactory or solid performance; competency, mastery, attainment, ability, or progress; general, good, or broad understanding of the mathematics content standards or grasp of the curriculum; and adequate knowledge and skills or applies knowledge and skills. Two states refer to proficient performance, and two other states use the term "solves problems." One state refers to partial success. No concise, defining phrase for Proficient is locatable in seven states' definitions.

**Warranted Inferences from the PLDs for Proficient**

In addition to the defining phrase, it is necessary to deconstruct the PLDs to understand the student for whom the PLDs are written and the types of KSPs described in the PLDs. Because a performance level encompasses a range of ability, when writing PLDs, it is helpful to identify the
A Note on Usability

As mentioned previously, we searched the websites for the 50 states and the District of Columbia in order to conduct this analysis. In some cases, it was quite difficult to find the PLDs on state websites because they were buried in documentation (e.g., technical reports), which made the PLDs less accessible to the public. PLDs are a useful tool for helping parents and teachers understand students' test performance; therefore, states should make it easy to locate the PLDs on their websites.

In addition, many of the PLDs we reviewed would not be usable by a parent without a background in mathematics education. When writing PLDs, it is important to remember who the end users will be and to target language appropriately. This may mean that short, conceptual PLDs should be written for parents, while long, detailed PLDs should be written for educators (and informed parents). Only a handful of states provided interpretative guidance on how to use the PLDs.

In reviewing the PLDs, it became apparent that some PLDs were better written and more accessible than others. For example, the Colorado Department of Education provides policy descriptors for the testing program; generic content descriptors that discuss mathematics concepts that students know; short, grade-specific content descriptors that summarize mathematics skills possessed by students in each performance level; and long, grade-specific content descriptors that detail the KSPs possessed by students within a grade level and content area. By creating multiple types of descriptors, Colorado may target each type at different audiences. For example, the grade-specific content descriptors may be used by teachers, while the short, grade-specific content descriptors may be more appropriate for parents.

Viewed together, the current state of K-12 PLDs shows the need for improved methods for future PLD development. Currently, most states provide grade- and content-specific PLDs that align to their content standards. This analysis also shows that the vast majority of states currently provide PLDs that reflect skills possessed by students rather than skills expected of students. There is much room for improvement in terms of helping end users access, interpret, and use the PLDs.
Threats to the Validity of PLDs

The previous section addressed current approaches to PLDs for the purpose of score interpretation. It is unknown how many of these PLDs were also used for other purposes, such as standard setting guidance and/or test development. Theoretically, organizations that are cognizant of the various uses of PLDs can link the different PLD types together and use them as tools in the development of a fully aligned assessment system. In principle, PLDs guide test development so that the tests are well aligned to the intended constructs of interest. PLDs guide standard setting so that panelists operationalize the PLDs in terms of cut scores with the same conceptualization used by item writers. Because the tests are already well aligned to the PLDs, the PLDs reflect the KSPs possessed by the students at or above each cut score. Therefore, the PLDs used in score reporting actually represent the observed skills of students within a particular performance level category. In reality, though, this seldom (if ever) occurs.

Disconnection of the Cut Score from the PLD Foundation

As previously noted, a sponsoring agency may desire to use the same set of PLDs for guiding the cut score recommendation workshop and interpreting tests scores. Once cut scores are finalized, however, a disconnect may occur between the descriptions of student expectations developed prior to setting standards and the KSPs associated with the cut score along the test scale. One of the earliest examples of this disconnect occurred in the 1996 NAEP standard setting when NAGB rejected panelist-recommended cut scores and adopted alternative cut scores. Once this occurred, it was necessary to adjust the PLDs so that they were aligned to the cut scores enacted through that policy decision (Bourque, 2000).

In the K-12 arena, policy makers may also adjust cut scores following a cut score recommendation workshop. Because cut scores are supposed to be operationalizations of the PLDs, this adjustment can create a disconnect between the PLDs and the cut scores, posing a threat to the validity of score interpretation (Linn, 1998; Schneider et al., 2009). This also points to the need to validate PLDs prior to their release to the public (see, e.g., Linn, 1998).

The KSPs represented in the PLDs should correspond to the cut scores that were implemented by the test's sponsoring agency. PLDs developed for standard setting are but one manifestation of an organization's policy; the final cut scores the agency approves and implements is another. If the PLDs are not well aligned to the cut scores, then their validity for score interpretation and future item development is called into question.

If policy body cut score adjustments are made, it is unlikely that the PLDs used initially for test development and standard setting guidance will fully reflect the final cut scores. Policy cut score adjustments may set the stage for erroneous interpretations regarding what examinees are actually able to do. For these reasons, a single set of PLDs is unlikely to serve all purposes equally well. Therefore, an interdependent system of PLDs is proposed that will likely need to be adjusted following the standard setting process to ensure that the intended construct is clearly defined for score reporting and for item development purposes beyond the baseline year of the assessment system.

A Proposed Framework for Developing PLDs

In this section, we describe a proposed framework for developing an interrelated system of PLDs designed to support the intended inferences of the test scores. Using the proposed framework, a sponsoring agency would create four interrelated PLDs that are developed in an iterative
sequence. It is essential to develop a congruous set of PLDs that reflect the final cut scores the sponsoring agency approved. This is consistent with the notion that a defensible standard setting process requires performance standards to be prospective, progressive, and predictive. The four PLDs are (a) Policy PLDs, (b) Range PLDs, (c) Target PLDs, and (d) Reporting PLDs. Table 5.2 provides an overview of the framework, delineating each PLD type along with its primary use, purpose, potential author, when it should be written, and its intended audience. In the following subsections, the framework will be explicated with particular attention paid to the development of each PLD type.

**Policy PLDs**

Policy PLDs are high-level descriptions that anchor the meaning of the performance labels that policy makers have selected to use with a testing program. Using approximately two or three sentences, the sponsoring agency describes its vision of the complexity of the KSPs that students should know to meet its goals (e.g., to be ready for college). Once written and accepted by the sponsoring agency, the Policy PLDs are not altered throughout the course of the testing program. They should only be revisited once a new testing program is to be put into place. For example, a state that established Policy PLDs with the advent of NCLB may reinvent those PLDs once it begins test development under the Common Core State Standards (Council of Chief State School Officers, CCSSO, & National Governor’s Association, NGA, 2011).

Table 5.3 shows Policy PLDs from two testing programs’ California and NAEP. These PLDs do not describe the specific content-based competencies expected of student performance; rather, these examples set the stage for the cognitive complexity with which students should be able to manipulate the content within a performance level by using phrases such as "complex
Tallie 5.3  Policy PLDs for Two Testing Programs

<table>
<thead>
<tr>
<th>California</th>
<th>NAEPu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced. This category represents a superior performance. Students demonstrate a comprehensive and complex understanding of the knowledge and skills measured by this assessment, at this grade, in this content area.</td>
<td>Advanced. Superior performance.</td>
</tr>
<tr>
<td>Proficient. This category represents a solid performance. Students demonstrate a competent and adequate understanding of the knowledge and skills measured by this assessment, at this grade, in this content area.</td>
<td>Proficient. Solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.</td>
</tr>
<tr>
<td>Basic. This category represents a limited performance. Students demonstrate a partial and rudimentary understanding of the knowledge and skills measured by this assessment, at this grade, in this content area.</td>
<td>Basic. Partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.</td>
</tr>
<tr>
<td>Far Below I Below Basic. This category represents a serious lack of performance. Students demonstrate little or a flawed understanding of the knowledge and skills measured by this assessment, at this grade, in this content area.</td>
<td></td>
</tr>
</tbody>
</table>


understanding," "application .... to real-world situations," or "analytical skills." Although these high-level PLDs are a necessary starting place to set the tone for a testing program, policy-based claims such as "well prepared for work at the next grade level" are difficult to substantiate (Linn, 1998) and require validity studies to support such claims.

It is important to note that the Policy PLDs in Table 5.3 were used by the respective states for all grades and content areas. Only one set of Policy PLDs should be developed for a testing program. Policy PLDs should use a similar set of memorable words that distinguish performance across levels, clearly state the degree of knowledge and skills across performance levels, and are carefully worded to clearly articulate the sponsoring agency's vision for examinees (Perie, 2008).

**Practical considerations.** Given the uses of Policy PLDs, these descriptors are the starting point for a testing program and may even be part of the legislation mandating the testing program. Policy PLDs should be developed prior to writing content standards because they set the tone for the expected skills students should possess. Also, given that Policy PLDs represent policy statements, their development will vary with jurisdiction and sponsoring agency. These descriptors may be written by policy makers in conjunction with high-level staff at the sponsoring agency. Alternatively, they may be written by content experts and approved by a policy board. In short, there are various ways to write Policy PLDs that provide high-level direction for a testing program. When developing the Policy PLDs, the number and names of the performance levels will need to be decided.

**Number of performance levels.** If a sponsoring agency has the latitude to select the number of performance levels, then the agency may want to consider the types and stakes of the decisions that will be made from the results of the testing program. Special consideration may be given to
the importance of false positives versus false negatives and the associated cost of misclassifying examinees. In general, the greater number of performance levels for a test, the lower the accuracy of correctly identifying the true performance level of the student (Ercikan & Julian, 2002).

Performance level labels. Nothing within the realm of PLDs may prove quite as contentious as naming the performance levels. In K-12 testing, labels such as Unsatisfactory or Failing can provoke strong reactions from educators, parents, and students. Chosen by policy makers, sponsoring agency staff, a special committee, or a combination of all three, sponsoring agencies often seek to provide meaningful names for their performance levels to communicate their policy agenda through the labeling process. In choosing labels, a sponsoring agency should be cognizant of the connotations associated with the labels to prevent unintended consequences (Burt & Stapleton, 2010).

Some evidence exists that well-developed PLDs may mitigate the connotations associated with performance level labels at cut score recommendation workshops (Burt & Stapleton, 2010); however, there is no evidence that PLDs are used by parents, students, and teachers to contextualize performance level labels outside of a cut score recommendation workshop. Even the use of generic performance level labels does not completely avoid the problem that labels evoke certain reactions. As end users become aware, for example, that Level 4 is the highest level of attainment within a testing program, it is easy to imagine teachers and administrators referring to Level 4 students in the same way that they might refer to Advanced students. This points to the need for the sponsoring agency to provide guidance to end users for using and interpreting performance level information, regardless of the labeling system.

Range PLDs

The Range PLDs describe the spectrum of performance for all examinees within each performance level. The Range PLDs define the construct that is to be measured by the assessment, specifically defining the "processes, strategies, and knowledge structures that are involved in item solving" (Embretson & Gorin, 2001, p. 349) for each level of achievement. The Range PLDs, then, become the cognitive theory that will underlie test development. Currently, broadly defined test specifications represent the cognitive theories (Leighton & GierJ, 2007b) that underlie test development for most summative CRTs. These specifications provide direction to test developers in terms of the weighting that should be assigned to content strands. Range PLDs contain the breadth of the test specifications, with additional details regarding how students across the ability spectrum should engage the content.

The Range PLDs provide item writers with a structure that denotes the specific evidence of student achievement that needs to be elicited across the range of achievement. As such, this PLD type will be quite lengthy and detailed because it describes the progression of processes, strategies, and knowledge that items should elicit across the performance level for all content standards found on the assessment. The primary purpose of the Range PLDs is to provide guidance to the item writers and test developers. These PLDs are not intended for consumption by the general public; therefore, they may be written for a technical audience.

Because the Range PLDs guide test development, they should deconstruct what evidence of achievement looks like at each performance level, for each standard on the assessment. Table 5.4 shows an example of a portion of a Range PLD for a Grade 8 mathematics test. As demonstrated in Table 5.4, this Range PLD scaffolds KSPs across the ability range of each performance level. For example, a Level 1 student can determine the perimeter of an irregular figure without
<table>
<thead>
<tr>
<th>Level</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>determine perimeter of irregular figure when no context is present</td>
<td>determine perimeter of irregular figure when context is present</td>
<td>given perimeter, determine the missing measurement of an irregular figure when context is present</td>
<td>represent an irregular figure as a regular figure to find the perimeter given length of 2 sides when context is present</td>
</tr>
<tr>
<td>find area of a noncomplex irregular figure when no context is present</td>
<td>find missing measurement of noncomplex irregular figure given area when context is present</td>
<td>find missing measurement of complex irregular figure given area when context is present</td>
<td>apply the formula for the surface area of a cylinder, given the formula and value for pi</td>
</tr>
<tr>
<td>find volume of irregular figure when no context is present</td>
<td>find volume of irregular figure when context is present</td>
<td>find missing measurement of irregular figure given volume when context is present</td>
<td>double the dimensions of a rectangular prism, and calculate the volume, given the formula</td>
</tr>
<tr>
<td>use data in a circle graph to determine if statements are true</td>
<td>use data in a single bar graph to determine if statements are true or false</td>
<td>use data in a line graph to determine if statements are true</td>
<td>use data in a double bar graph to determine if statements are true</td>
</tr>
<tr>
<td>determine the mode of a set of data</td>
<td>determine the range of a set of data</td>
<td>determine the median of a set of five 2-digit numbers presented with context</td>
<td>determine the mean of a set of five 2-digit numbers presented with context</td>
</tr>
</tbody>
</table>

context being present; whereas, a Level 2 student is expected to show how to apply this skill in a real-world situation. Notice that the descriptors do not distinguish the lowest performing student from the mid-level and highest performing students within a performance level.

**Practical considerations.** Ideally, the Range PLDs should be written in conjunction with the content standards; however, they may also be written after the content standards are available and before test development begins. In the realm of K-12 testing, the Range PLDs are not easy to develop. They require knowledge of theories regarding the way and sequence in which students learn particular concepts and the way that students demonstrate knowledge. These are not areas with clear-cut answers. In addition, the use of Range PLDs for test development means a radical departure from current operating procedures for most testing programs.

**Authoring range PLDs.** Range PLDs define the way students approach and process content, in addition to the types of content expected of students in each performance level. To write these PLDs, a sponsoring agency may first need to conduct literature reviews on existing theories of learning within a content area. A literature review may provide a starting point on how to delineate the KSPs within each performance level. In addition, it may prove useful to interview teachers and students about how students of different ability levels approach and process content (Gorin, 2006). Also, a sponsoring agency should invite a committee of educators and cognitive scientists to draft preliminary Range PLDs that take into account the pre-existing knowledge of committee members and the findings of the field work conducted by the sponsoring agency.

**Test development.** It is beyond the scope of this chapter to detail how Range PLDs can be used for item writing and test development; rather, the intent is to introduce a framework of PLDs that links test development to score interpretation. The use of Range PLDs is a departure from current practice, and its feasibility must be questioned given the results of known attempts to
target item writing to PLDs (Ferrara et al., 2009; Huff & Plake, 2010). Nonetheless, the idea of linking PLDs to item writing is still relatively new, and more efforts are needed to truly understand the plausibility of the model.

Articulation of performance levels. The articulation of performance levels refers to the performance levels making sense across grades, content areas, the testing program, or all three. Range PLDs need to be developed as a system so that they are coherent across each of the following:

Grades. This is often the largest area of concern for a sponsoring agency. The descriptors associated with each grade level should show an accumulation of skills across grades. The expectations for students in higher grades should be greater than those for students in lower grade levels. This implies that the content experts who develop the Range PLDs should also have knowledge of educational, developmental curriculum.

Content Areas. In some cases, a sponsoring agency may not be concerned with the articulation of Range PLDs across content areas. It is very difficult to compare KSPs across multiple content areas and to maintain similar types of rigor across content areas. If this type of articulation is desired by the sponsoring agency, then the agency will need to recruit curriculum specialists with cross-grade knowledge of all the content areas.

Testing Program. Many states must develop PLDs for alternate and modified assessments in addition to their general assessment. There should be some cohesion across these assessments: in particular, there should be good consistency between the general and modified assessments (Egan et al., 2011). For this to occur, the sponsoring agency should use the Range PLDs from the general assessment to define the expectations of the modified assessment’s Range PLDs.

Target PLDs

Target PLDs are developed in conjunction with the Range PLDs. They define the expected performance of a student who just enters a performance level, and they represent the lower end of the Range PLDs. Target PLDs help refine the cognitive theory underlying the Range PLDs by delineating which KSPs constitute entry into a performance level. Therefore, Target PLDs help transition from the breadth of the Range PLDs to the narrowly focused Reporting PLDs and represent the sponsoring agency's first attempt to explicate the intended interpretations of the test scores. Target PLDs may be used by item writers to define the KSPs of the student at the threshold of the Range PLDs; however, the essential function of Target PLDs is to provide guidance to participants at cut score recommendation workshops.

The importance of Target PLDs to the standard setting process cannot be underestimated because they provide standard setting panelists with a common understanding of the sponsoring agency's expectations regarding the content-based competencies that students should possess in order to enter each performance level. To clarify, the Target PLDs describe the KSPs of the borderline student, that is, the student just entering a performance level. Panelists arrive at a cut score recommendation workshop with many different conceptions of students and student performance. If queried, each panelist could provide the KSPs he or she believes, for example, the Proficient student should possess. For this reason, standard setting facilitators use Target PLDs to unite panelists in a common understanding of the KSPs expected of students with just the skills necessary to be considered, for example, Proficient. Participants use Target PLDs as guidance from the sponsoring agency to understand its expectations for examinee performance.
Table 5.5 Example of Target PLDs from a High School Social Studies Test

<table>
<thead>
<tr>
<th>Performance Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above the Standard</td>
<td>A student just entering this level should be able to analyze and evaluate the impact of major events, people, and themes in U.S. History and connect them to multiple outcomes. These students explain relationships and draw conclusions and support them with detailed and accurate evidence. They use complex or historical maps and primary source materials (e.g., political cartoons, speeches, photographs) to draw conclusions, analyze cause-and-effect relationships, and to interpret and apply new information.</td>
</tr>
<tr>
<td>At the Standard</td>
<td>A student just entering this level should be able to describe and explain major events, people, and themes in U.S. History and connect them to other historical events. These students draw conclusions and support them with accurate evidence. Students use maps and primary source materials to retrieve information and draw limited conclusions. They demonstrate the ability to understand issues and problems, recognize cause-and-effect relationships, and to compare the present to the past.</td>
</tr>
<tr>
<td>Partially Met Standard</td>
<td>A student just entering this level should be able to recognize and identify major events, people, and themes in U.S. History and use maps and primary source materials to retrieve information. These students draw simple conclusions and support them with limited or no evidence.</td>
</tr>
<tr>
<td>Below the Standard</td>
<td>A typical student performing at this level demonstrates insufficient understanding of major events, people, and themes in U.S. History and an insufficient ability to recognize or identify information found in maps and primary source materials.</td>
</tr>
</tbody>
</table>

Even though some panelists may not agree with the Target PLDs, they generally accept that the Target PLDs are to be used to guide their discussions and recommendations. Table 5.5 shows Target PLDs from a high school Social Studies test. The PLDs in Table 5.5 are written for the student who is just entering each performance level, with each PLD describing the skills those entry students should have.

Practical considerations. The concurrent development of the Range and Target PLDs represents a chicken-and-egg situation. The KSPs delineated within Target PLDs are transitional KSPs that show the shift from one performance level to another. In some cases, these transitional KSPs within a content strand may be clear to the PLD authors. In other cases, though, the PLD authors may be able to broadly identify KSPs into the performance levels and then struggle to determine where the transitional KSPs occur. The Range and Target PLDs may require several iterations before the versions are ready for use by item writers and test developers. Since the Target PLDs are the theoretical lower bound of the Range PLD, it is possible that they could be changed as new information on student learning presents itself, in particular, after the first operational administration of items. The question of whether or not Target PLDs should be adjusted during the cut score recommendation workshop, as new information is provided, is addressed in the next section.

Developing Range and Target PLDs

The method described below has been used to successfully create Target PLDs to guide standard setting, and it is easily understood and utilized by panelists. This method has been expanded so that both Range and Target PLDs can be developed. It assumes that content standards are already developed.

Authoring Range and Target PLDs is best accomplished in a workshop with small groups of educators. To develop Range and Target PLDs, panelists should begin studying the Policy PLDs
to understand the intent of the sponsoring agency regarding the testing program. It is the job of the panelists to operationalize this intent through the Range and Target PLDs.

Next, the panelists should study and begin to annotate the content standards. During this process, the panelists analyze the cognitive expectations, content information, and skills found in the content standards. State standards often imply a range of cognitive complexity within a single statement, as shown in Figure 5.1. Figure 5.1 shows third-grade English Language Arts Standard 1 from New York. A substrand of this content standard says, "Locate and use library media resources to acquire information, with assistance." The skill within this substrand can be separated into two different levels of cognitive complexity:

- Locate library media resources to acquire information with assistance
- Use library media resources to acquire information with assistance

After identifying the range of cognitive complexity within the content standard and substrands, the panelists should annotate the content standards, indicating the performance level to which each skill aligns. The Policy PLDs provide general guidance for analyzing the content and skills in the content standards. Within each performance level, the skill can be further delineated into just Proficient (P-), average Proficient (P), or highly Proficient (P+) using the annotation process shown in Figure 5.1. A similar delineation should be made for the other performance levels. For example, Figure 5.1 shows panelists have indicated that locating library media resources is a skill expected of the just Basic examinee (B-), whereas using library media resources is a skill expected of the average Basic examinee (B).

As this analysis is performed, the content and skills can be compiled into a database showing their alignment to each performance level. Then, a matrix can be created showing the delineation of KSPs into Range PLDs, which include all of the content and skills assigned to a performance level. The Range PLDs should eventually resemble the matrix presented in Table 5.4.

The KSPs classified in the just Proficient (P-) category constitute the Target PLD for Proficient. The KSPs of both the Target and Range PLDs should be examined for fluency across the grade levels within a content area. The KSPs should be compiled into a single set of PLDs for each grade level and content area, resembling those shown in Table 5.5.

Adjusting Target PLDs. The question can be asked whether or not the Target PLDs should be adjusted once the Range PLDs are being used by item writers; in particular, should they be adjusted at the cut score recommendation workshop? One viewpoint is that Target PLDs reflect the expectations that the sponsoring agency has regarding student performance and thus should not be changed. In this viewpoint, the sponsoring agency has standard setting panelists operationalize the expectations delineated in the Target PLDs through the cut score recommendation workshop.

The irony of this viewpoint is that a sponsoring agency develops many types of test-related documents (e.g. content standards, test items, depth-of-knowledge documentation) with input from informed members of the public. For example, committees of educators often help develop content standards, and the sponsoring agency is responsible for refining and approving the content standards. Target PLDs can and should work in much the same way. In standard setting contexts involving K-12 achievement testing programs, the panelists at cut score recommendation workshops are most often educators with, perhaps, other informed members of the public (such as parents or business leaders) also serving as panelists. The information presented at the cut score recommendation workshop is helpful in refining the lower end of the Range PLDs, which are the Target PLDs. At a Bookmark standard setting, for example, the information
Original Grade 3 Content Standard*

Standard 1: Students will read, write, listen, and speak for information and understanding.
- Read and use library media resources to acquire information, with assistance
- Read unfamiliar texts to collect data, facts, and ideas.
- Locate information in a text that is needed to solve a problem
- Identify main ideas and supporting details in informational texts
- Recognize and use organizational features, such as table of contents, indexes, page numbers, and chapter headings/subheadings, to locate information, with assistance

Annotated Standard

Standard 1: Students will read, write, listen, and speak for information and understanding.
- B- B B Locate and use library media resources to acquire information, with assistance
- B+ B B Skills of the student just entering proficiency
- P- P- P- Skills of the student just entering proficiency
- P B+ B Skills of the average Proficient student
- P+ P+ B Skills of the high-performing Proficient student
- P+ P+ P+ Skills represented by P- are compiled for the Target PLD.

Example of Target Proficient PLD

Standard 1: The Grade 3 target Proficient Student will read, write, listen, and speak for information and understanding with the following limitations:
- Read unfamiliar text to collect data
- Use organizational features, such as table of contents and chapter headings/subheadings

Skills represented by P- would be compiled for the Range PLD.


Panelists begin the cut score recommendation workshop by studying and discussing the Target PLDs that were developed prior to the workshop, in concurrence with the Range PLDs. Ideally, the panelists would have studied the Range PLDs prior to the cut score recommendation workshop. Throughout the workshop, panelists continue to discuss the skills in the Target PLDs and to refine them based on new information that is introduced in the multiple rounds of the standard setting process. By the end of the standard setting process, the Target PLDs have been informed by panelist discussions, the test items, the content standards, and impact data. This
should serve to strengthen the theoretical underpinnings of the lower end of the Range PLDs (i.e., the Target PLDs).

Reporting PLDs

The Reporting PLDs describe what students who just enter a performance level should know and be able to do consistently. In addition, the KSPs described in lower performance levels are subsumed by students in the higher performance levels. As the name implies, Reporting PLDs are useful to communicate to examinees (and to parents, educators, or other relevant audiences) about their test performance in terms of KSPs. Reporting PLDs are fluid in the sense that they may be updated following the administration of new test forms. It is important to understand that the KSPs reflected in these PLDs do not change; rather, they are augmented as new information is added to the Reporting PLDs when new test forms are administered. The purpose of adjusting the KSPs is to provide more information to the end user about the examinee.

Table 5.6 shows a Reporting PLD from the South Carolina Palmetto Achievement Challenge Test. As shown in Table 5.6, the Reporting PLD describes the KSPs that students in each performance level can likely do. These particular PLDs also discuss some of the KSPs that students in each performance level likely cannot do.

Practical considerations. Reporting PLDs should be written once cut scores are finalized by a sponsoring agency. As was discussed previously, a sponsoring agency may adjust cut scores following a cut score recommendation workshop based on policy concerns; therefore, the KSPs in the Reporting PLDs need to reflect those adjustments. They may be written by a special committee, sponsoring agency staff, or staff from a testing contractor. In writing the Reporting PLDs, it is important to offer interpretative guidance for the end user so they use the PLDs appropriately.

Updating Reporting PLDs. Reporting PLDs should be reviewed periodically to ensure that they remain aligned with the KSPs measured by the test (Crane & Winter, 2006). Reporting PLDs are often developed from one or two test forms, and the KSPs reflected in the original Reporting PLDs may be broadened to reflect information from new test forms.

Relationship to Target PLDs. Target PLDs represent the lower boundary of Range PLDs and provide the sponsoring agency's expectations for that performance. For score reporting, the agency makes a determination regarding which student's performance the Reporting PLDs represent: the typical student within a performance level or the student at the cut score. Their development should be consistent with the student the PLDs represent, and the score report should make clear whether the skills discussed are representative of a typical student at the cut score or the average student within the performance level. The sponsoring agency should provide guidance, which can be accomplished in the first line of each Reporting PLD, by including a statement such as: (a) "The typical student in this performance level is able to..." or (b) "The student at the cut score for the performance level is able to..."

The interpretative guidance should also state that the Reporting PLDs are only a sampling of the KSPs that students in each performance level know and are able to do. The guidance should inform end users that the Reporting PLDs from summative CRTs are meant to provide a global snapshot of the examinee's KSPs and are not intended to be the sole source of information for curricular decisions. These decisions should be made in conjunction with other educational information about the examinee.
Table 5.6  Example of Reporting PIDs

<table>
<thead>
<tr>
<th>Grade 8 PACT English Language Arts Performance-Level Descriptors</th>
<th>Below Basic</th>
<th>Below Basic likely cannot do:</th>
</tr>
</thead>
</table>
| What below basic students likely can do: | | • locate details in longer passages  
   • make simple inferences from high-interest informational and literary texts  
   • paraphrase the main idea  
   • provide literal interpretations of informational and literary text  
   • combine strategies (e.g., locate details to make an inference) while reading  
   • identify literary elements (e.g., simile, point of view) introduced in elementary school |
| • use key words or phrases in passages that are of high interest to skim text and locate obvious details  
• identify the main idea when the passage provides a stated main idea  
• draw simple conclusions about a passage when the text provides obvious support for those conclusions |
| What basic students likely can do that below basic students likely cannot do: | | • make fine distinctions among details and analyze them to make more complex inferences regarding longer, denser informational and literary texts  
   • understand and analyze both literal and figurative language  
   • combine strategies (e.g., locate details to make an inference) while reading  
   • interpret poetry  
   • make interpretations that go beyond the text to answer constructed-response questions  
   • support their responses with relevant details |
| • locate details in longer passages  
• make simple inferences from high-interest informational and literary texts  
• paraphrase the main idea  
• provide literal interpretations of informational and literary text  
• combine strategies (e.g., using details to make an inference) while reading  
• identify literary elements (e.g., simile, point of view) introduced in elementary school |
| What proficient students likely can do that basic students likely cannot do: | | • make fine distinctions among and analyze details to make more complex inferences regarding longer, denser informational and literary texts  
   • understand and analyze both literal and figurative language  
   • interpret poetry  
   • make interpretations that go beyond the text to answer constructed-response questions  
   • support their responses with relevant details |
| • make fine distinctions among details and analyze them to make more complex inferences regarding longer, denser informational and literary texts  
• understand and analyze both literal and figurative language  
• interpret poetry  
• make interpretations that go beyond the text to answer constructed-response questions  
• support their responses with relevant details |
| What proficient students likely cannot do: | | • make fine distinctions among many details to make more complex inferences regarding longer, denser informational and literary texts  
   • understand, analyze, and evaluate both literal and figurative language  
   • provide detailed, relevant, complete, insightful, and creative answers to constructed-response questions |
| • make fine distinctions among many details and analyze them to make more complex inferences regarding longer, denser informational and literary texts  
• understand, analyze, and evaluate both literal and figurative language  
• provide detailed, relevant, complete, insightful, and creative answers to constructed-response questions |
| What advanced students likely can do that proficient students likely cannot do: | | • make fine distinctions among many details and analyze them to make more complex inferences regarding longer, denser informational and literary texts  
   • understand, analyze, and evaluate both literal and figurative language  
   • provide detailed, relevant, complete, insightful, and creative answers to constructed-response questions |

A Method for Developing Reporting PLDs

Hem-mapping standard setting procedures, such as the Bookmark Standard Setting Procedure (see Lewis, Mitzel, Mercado, & Schulz, Chapter 12 of this volume) or the ID Matching procedure (see Ferrara & Lewis, Chapter 13 of this volume), lend themselves to the development of Reporting PLDs. In these procedures, test items are ordered from easiest to hardest, and an item map can be created that includes selected information about each item. Figure 5.2 shows an example of an item map that can be used in conjunction with the test items to write Reporting PLDs. The item map in Figure 5.2 has the cut scores placed on it.

For each item, the item map in Figure 5.2 shows the item's order of difficulty, scale location, item number on the operational test, item type, score key, and content strand. In addition, it shows two questions that are generally asked and answered in a Bookmark standard setting. These two questions can help item writers analyze the KSPs of each item.

In Figure 5.2, the second, sixth, and eighth items represent the score points of a single construct-response (CR) item worth three points. Unlike a multiple-choice (MC) item, which appears once on the item map, a CR item appears as many times as it has points. When multi-point CR items are included, Reporting PLDs should be written to the KSPs demonstrated by the highest CR point for each item within a performance level. For example, if the second and third CR points for an item are in the range of Proficient performance, the Reporting PLD for that level should describe the KSPs for the third score point of the item.

For each performance level, content experts can use the two questions on the item map in Figure 5.2 to summarize the KSPs of the items whose scale locations fall below the cut score for that level. The KSPs of the items below the cut score represent the performance of the student who just entered the performance level, and they are the KSPs held in common by all students in the performance level. The Reporting PLDs should be written neither too broadly, in which case no information is provided, nor with too much detail, in which case items could be compromised.

To begin writing Reporting PLDs, it is often helpful to start with bulleted lists of KSPs instead of writing paragraphs. Ryan (2003) found stakeholders prefer bulleted lists to narrative paragraphs. It is also useful to choose explicit, action verbs, such as explain, identify, or apply and to

<table>
<thead>
<tr>
<th>Item</th>
<th>Score</th>
<th>Type</th>
<th>Scale Location</th>
<th>CR Points</th>
<th>Score Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>568</td>
<td>MC</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>560</td>
<td>CR</td>
<td>1 of 3</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>562</td>
<td>MC</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>563</td>
<td>MC</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>568</td>
<td>CR</td>
<td>2 of 3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>571</td>
<td>MC</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>575</td>
<td>CR</td>
<td>3 of 3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>580</td>
<td>MC</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>584</td>
<td>MC</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>584</td>
<td>MC</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>596</td>
<td>MC</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>598</td>
<td>MC</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 5.2 Sample Item map with cut scores shown](image)
avoid generic verbs such as know, understand, or recognize (see Mehrens & Lehmann, 1991, for other examples of clarifying language).

**Reporting PLDs should provide contextual characteristics.** Contextual characteristics of items can help explain why a particular skill is located in a particular performance level. For example, Grade 3 Basic students may be able to follow written directions only when the directions are short and have associated figures to help demonstrate the meaning. A reporting descriptor such as "follows written directions when associated figures are present" or "follows dense written directions when no associated figures are present" describes important differences in examinee knowledge and skills across performance levels. The inclusion of these types of contextual differences provides information about the complexity of the tasks in a particular performance level, which teachers can use to support instruction.

When developing the Reporting PLDs, keep in mind that they will be provided to end users who may have little curricular knowledge; therefore, jargon and overly technical language should be avoided.

**A note on the response-probability criterion.** For all item-mapping procedures, a response-probability (RP) criterion must be determined before the items can be ordered from easiest to most difficult by scale location. In practice, RP values of .50 or .67 (RP50 and RP67, respectively) are fairly common (see Lewis et al., Chapter 12 of this volume, for additional information on RP values). When items are ordered using RP67, the scale location associated with the item is the scale score an examinee needs to have a two-thirds chance of answering the item correctly. It is important to consider the implications of the RP value used when writing the Reporting PLDs.

If the Reporting PLDs reflect the KSPs of an examinee, then the cut score should also reflect this to the degree possible. The RP67 value reflects a most likely scenario; that is, a situation in which a examinee who has a scale score equal to the scale location has a two-thirds chance of answering the item correctly. If RP50 is used, an examinee who has a scale score equal to the scale location has a 50% chance of answering the item correctly. In this situation, examinees are equally likely to answer incorrectly as they are correctly. This means that half the examinees who just enter a performance level would have mastered the expected content. This would need to be reflected in the language of the Reporting PLDs.

**Validity Concerns of the PLD Framework**

The proposed PLD framework attempts to address the intended uses of PLDs by parsing them into interrelated sets of PLDs that include Policy PLDs, Range PLDs, Target PLDs, and Reporting PLDs. One concern, however, is that actual examinee performance will not meet the performance predicted by the Range PLDs. In part this may be due to the way items are written, or it may be that the theory explicated in the Range PLDs does not align to reality.

**Item writing.** Item writing is a combination of art and science. As such, items cannot be guaranteed to align accurately to expected performance. Easy items may be written for advanced content and vice versa. This issue may be addressed by developing a very large pool of items from which to develop test forms; however, this may not always be feasible due to economic and/or logistic constraints. In addition, it may be that the underlying cognitive theory of the Range PLDs does not accurately predict the examinees’ KSPs for a particular content area.
Adjusting the Range PLDs. Suppose the underlying cognitive theory of the Range PLDs does not accurately predict the examinees’ KSPs for a particular content area. This can be addressed by adjusting the Range PLDs to match what has been learned about examinee performance on the test or by returning to the item pool to understand whether the cognitive theory or the items should be adjusted. Although adjusting the Range PLDs is tempting, this should only be done after multiple pieces of information indicate that the original hypotheses associated with the underlying cognitive theory are incorrect. It may take several test development cycles before all the items are aligned to the correct Range performance level. Because the Range PLDs guide test development, they should not be moving targets as new forms are created and administered. Even so, it may be necessary to adjust the Range PLDs once cut scores are finalized so that there is alignment between the operational cut scores and the Range PLDs.

Future Directions in PLD Development and Use

The role of PLDs in assessment programs, and how they are conceptualized and used, has evolved considerably in the last 20 years. Given the current movement to adopt and assess Common Core State Standards, there is good reason to expect this evolution to accelerate in the next few years. This section considers potential changing approaches to PLDs, possible roles for PLDs in the next generation of state assessment programs, and needs and ideas for validation research.

Conceptual Approaches to PLDs

The widespread adoption of the Common Core State Standards (CCSS) and the extensive participation in multi-state consortia will significantly influence how PLDs are conceptualized and described. Because the CCSS are organized around readiness for college and careers, the Range PLDs developed based on the CCSS will reflect definitions of readiness for students as they approach high school graduation, and will reflect whether students in earlier grades are on track each school year to achieve readiness by the time they graduate. Range PLDs that define college and career readiness, in early grades and at the end of high school, predict future outcomes such as continuing achievement across grades, following a trajectory toward readiness for college or a career, and subsequent successful performance in college or in the workplace. These Range PLDs may or may not be worded explicitly as predictions, but they will imply predictions that should be evaluated in validation studies.

Given the use of CCSS, it is likely that many states (e.g., those within a consortia) will share PLDs. New questions will arise as common core assessments and PLDs are interpreted in multiple states. For example, will a definition of Proficient performance be interpreted equivalently across states with diverse curricula, instructional materials, teaching forces, student demographics, and rates of poverty, disabilities, and English-language limitations?

Role of PLDs in the Next Generation of State Assessments

The next generation of state assessments encompasses formative, interim (benchmark), and/or through-course assessments in addition to the end-of-year summative assessments. For example, the Partnership for Assessment of Readiness for CoUege and Careers (PARCC) consortium envisions quarterly, through-course assessments that are intended to supplement and be integrated with the achievement information from the state summative assessment. This type of integrated testing system provides states with the opportunity to collect and describe richer
information on student performance than at any time in the past. It also provides an opportunity to more thoroughly develop and test the cognitive theory that underlies the Range PLDs. Not only can the Range PLDs be used to develop summative assessments, they can also be used to develop the other assessment types that form the integrated testing system.

The use of varied assessment types also allows states to more thoroughly develop their Reporting PLDs. States will need to define and describe performance on a composite of summative and quarterly performances. Reporting PLDs should be available for each component of an assessment system, and these PLDs should be developed in conjunction with one another. This also means that cut scores from the summative tests will need to be aligned with cut scores from the other components of the assessment system. In addition, the use of varied assessment types provides an opportunity to provide Reporting PLDs for sub-score information (e.g., number sense and geometry sub-scores in mathematics).

The usefulness of sub-scores is likely to be enhanced as researchers continue to demonstrate applications of multidimensional IRT (MIRT; Reckase, 2009) and cognitive diagnostic models (Leighton & Gierl, 2007a). Both of these approaches provide empirical support or, rather, illustrate the usefulness and necessity of reporting performance on subsets of items in a seemingly unidimensional set of items. Test designers and researchers search for subsets of items that provide useful, differentiated performance profiles of examinees to guide instruction. As MIRT and cognitive diagnostic modeling come into wider use, the PLDs within the proposed framework will need to reflect performance profiles.

Experience in educational achievement testing in writing PLDs has been limited to accounting for the knowledge and skill requirements in multiple-choice and constructed-response items. The next generation of state assessments is expected to include a wider range of constructed-response items, technology-enhanced items, and performance tasks (e.g., essays based on extended reading assignments). Although there are notable exceptions (e.g., PLDs for assessments of the National Board for Professional Teaching Standards; see www.nbpts.org), conceptualization and planning will be required to incorporate the unique demands of these types of assessment tasks into the PLD creation process.

**Validation Research for the PLD Framework**

This chapter has presented a framework for developing four types of PLDs: Policy PLDs, Range PLDs, Target PLDs, and Reporting PLDs. To the degree it is feasible, the usefulness of these PLDs should be demonstrated empirically. For example, the logic of first developing Policy PLDs seems obvious, and the use of Policy PLDs in NAEP demonstrates the utility of doing so. Long-standing practice provides some support for continuing this practice. In contrast, the usefulness of Reporting PLDs can be tested empirically by determining whether end users such as examinees, parents, and educators make warranted inferences about what examinees know and can do based on the PLDs (see, e.g., Hambleton & Slater, 1997). A recent study by Burt and Stapleton (2010) illustrates the need to conduct such research. Those authors found that even content experts who represented typical standard setting panelists interpreted performance labels differently (e.g., Basic and Proficient). They also found that affective preferences for some labels persisted after the labels were defined (e.g., Basic and Apprentice were preferred over Limited, and Proficient was preferred over Satisfactory). Seemingly simple differences between labels may be interpreted in different ways and may evoke important differences in emotional responses.

In addition, it is necessary to evaluate whether standard setting panelists are able to understand and apply the Target PLDs during cut score recommendation workshops. The workshops
should include time for training and discussion of the Target PLDs. Questionnaires should be administered during the workshop to gauge panelists' level of understanding of, usage of, and agreement with the Target PLDs. If panelists do not use the Target PLDs and/or do not understand them, then this could have serious consequences for the validity of the cut scores that they recommend. (See Skorupski, Chapter 7 of this volume, for a review of research on panelists' cognitions.)

Additionally, there is preliminary evidence that Reporting PLDs need to be evaluated and, perhaps, updated every few years as new operational test forms become available. It is not well understood how stable KSPs are across years. This variability in KSPs across operational test forms reflects the difficulty in targeting items to particular scale score ranges. Currently, a statistical indicator for evaluating PLD drift does not exist. Instead, it is necessary to delineate the KSPs for each item, map the items by statistical difficulty, and then examine how often KSPs change performance levels across test administrations. If the KSPs associated with PLDs contradict each other when a new test form is administered, this may undermine the credibility of the Reporting PLDs.

Conclusion

The main purpose of this chapter was to provide a framework for developing four types of PLDs that serve as the foundation of the test development process. Within the proposed framework, the Policy PLDs set the tone for the testing program. The Range PLDs articulate the intended construct so that items are written and tests developed to align with expected achievement from the very beginning of the test development process. The Target PLDs specify the necessary KSPs to enter a performance level, thereby defining the lower limit of the Range PLDs. The Target PLDs are operationalized through the standard setting process, and they are transformed into Reporting PLDs that reflect the final cut scores.

The four PLD types were based on the common uses of PLDs found in the literature. By recognizing and attending to these different uses, the framework addresses the question of when PLDs should be finalized. The framework recognizes that PLD development, like the determination of cut scores, is ultimately a policy question. Policy makers generally reserve the right to adjust the cut scores after panelists have made recommendations at a workshop. When policy makers adjust cut scores, they are also adjusting the underlying content and it is only appropriate that this is reflected in the PLDs.

Finally, it is hoped that the proposed framework provides guidance to sponsoring agencies as they look to develop PLDs for new testing programs. It is hoped that sponsoring agencies will utilize each of the four types of PLDs when building assessments.

Notes

1. PLDs are also referred to as achievement level descriptors (ALOs).
2. Peer review guidance associated with the NCLB Act requires that PLDs are developed prior to standard setting. Thus, for K-t2 testing, PLDs must be developed before the workshop.
3. A similar analysis of definitions of Proficient performance on alternate assessments for students with significant cognitive disabilities appears in Ferrara, Swaffield, and Mueller (2009).
4. A noun phrase is a phrase whose head is a noun that may be amplified by modifiers referring to other parts of speech; see O'Grady, Archibald, Aronoff, and Rees-Miller, 2009.
5. A verb phrase is a phrase whose head is a verb that may be amplified by modifiers delineating other parts of speech; see O'Grady, Archibald, Aronoff, and Rees Miller, 2009.
6. This framework is an expansion of an idea proposed by Egan, Schneider, and Ferrara (in press).
7. We concentrate on CRTs because this is the test for which cut score recommendation workshops are held.
8. Bourque (2000) reported that NAEP standard setter facilitators were nervous the first time PLDs were written for panelists because the facilitators did not know whether panelists would accept PLDs that they had not written. She comments that panelists did not seem to mind; our experience is similar.

References


Appendix C: Résumés

The résumés for the SBAC-12 Initial Achievement Level Descriptor Development proposal are presented below in alphabetical order. The reference table of contents is provided for your convenience.

Janice Barth ..........................................................................................................................................................................2
Marjorie Bryant...................................................................................................................................................................4
Kathryn Dunlap...................................................................................................................................................................5
Karla L. Egan........................................................................................................................................................................7
Heather Farina ................................................................................................................................................................. 11
Kelly E. Godfrey – College Board .................................................................................................................................. 13
Angelica Gordon ............................................................................................................................................................. 16
Darin Homer .................................................................................................................................................................... 17
Pamela K. Kaliski – College Board ................................................................................................................................... 18
Dan Macomber ................................................................................................................................................................ 21
Ricardo Mercado ............................................................................................................................................................. 22
M. Christina Schneider ................................................................................................................................................... 24
Gretchen R. Schultz ........................................................................................................................................................ 28
Kevin P. Sweeney – College Board .................................................................................................................................. 30
Kristal Taylor ................................................................................................................................................................. 33
Natasha Vasavada – College Board ................................................................................................................................... 34
Andrew Wiley – College Board ........................................................................................................................................ 39
Janice Barth

**Academic Background**

- Ed.D.  Educational Leadership / Curriculum & Instruction Minor, West Virginia University, Morgantown, WV
- M.S.  Special Education, Gifted Education, Marshall University, Huntington, WV
- B.A.  Social Studies, Marshall University, Huntington, WV

**Professional Experience**

**CTB/McGraw-Hill, Monterey, CA: 2010 to present**

**State Solutions Manager:** Direct interaction with staff of assigned states to provide assessment solutions that match the needs of the state assessment programs. Seek and find released request for proposals and bring them to the CTB Bid Board for review/bid decisions.

**West Virginia Department of Education, Charleston, WV: 2007 to 2010**

**Special Assignment to State Superintendent:** Directed and managed state assessments, Title I, Special Education Monitoring and Research office staff and budgets. For these four areas, the job responsibilities included the following: developing policies, writing proposals, working effectively with staff from the West Virginia Department of Education, county school district staff, parents, higher education, the Center for Professional Development and the West Virginia Legislature. Worked effectively with the United Stated Departments of Education, Council of Chief State School Officers (CCSSO), and the National Assessment Governing Board (NAGB). Special Assignments included serving as the Federal Liaison to CCSSO, the legislative advocacy group that communicates and corresponds with members of Congress, and serving as a member of the State Superintendent’s Cabinet.

**West Virginia Department of Education, Charleston, WV: 2000 to 2007**

**Executive Director:** Responsibilities included developing and implementing the new state assessment for accountability and all other state assessment programs; directed and managed office staff and budgets; developed policies; wrote assessment proposals working effectively with staff from the West Virginia and United States Department of Education, county school district staff, higher education, Center for Professional Development and the legislature. Worked with the West Virginia Board of Education, Appalachian Educational Lab (AEL), Council of Chief State School Officers (CCSSO), National Assessment Governing Board (NAGB) and parents.

**West Virginia Department of Education, Charleston, WV: 1995 to 2000**

**NAEP Coordinator:** Coordinated National Assessment of Educational Progress (NAEP) statewide testing, prepared statewide administrator trainings to include instructional materials for teacher training on NAEP frameworks and item bank to improve student achievement. Worked with the federal NAEP representatives and state level coordinators.

**West Virginia Department of Education, Charleston, WV: 1989 to 1995**

**West Virginia Challenge Coordinator:** Developed, designed, and implemented new instructional programs in the following areas: mathematics, social studies, and English to include cooperative learning, thinking skills, and worked directly with State, RESA, and county personnel. Grant writing for Challenge Program included successful $500,000 to Benedum Foundation for Innovative Program Design and Development and secured other successful smaller grants for Benedum.
Janice Barth (page 2 of 2)

**Professional Experience (continued)**

**West Virginia University, Charleston, WV: 1991 to 2007**

**Adjunct Associate Professor:** Developed course syllabus and instruction for West Virginia Challenge Program of Studies, including the development and preparation of instructional materials for satellite transmission; adjunct professor for WVDE workshops that secures college credit.

**Kanawha County Board of Education, Charleston, WV: 1981 to 1989**

**Teacher:** Member of collaborative team that designed and implemented the Kanawha County School Integrated Service Model for gifted students in the high school level in Kanawha County.

**Professional Organizations**

Member of WV Partnership to Assure Student Success (PASS)

Federal Liaison Member for WVDE Legislative Advocacy Team

Member of Legislative NCLB Subcommittee B

Member of RESA IV Regional Council

Member of West Virginia Association School Administrators

Member of West Virginia Reading Council

President of West Virginia Council for Exceptional Children

Conference Chairperson and President Elect for West Virginia Council for Exceptional Children

Curriculum Committee Member for National Association for Gifted Children

President for West Virginia Association for Gifted and Talented

Conference Chairperson and President Elect for West Virginia Association for Gifted and Talented

Member of Association for Supervision and Curriculum Development

**Publications**


Annual West Virginia Accountability Workbook (2003–2010)

West Virginia Request for Proposal (RFP) (2007)


Design WESTEST Website (WVDE, 2004)


Statewide Assessment Results Book (WVDE, 2004)

The State of West Virginia Consolidated Application Accountability Workbook, (WVDE, 2003)

Investigation of the Relationship Between School Organizational Health and School Achievement as Mediated by Socioeconomic Status in West Virginia Middle Schools in the Areas of Reading, Mathematics and Writing from 1996 to 1999 (Dissertation, 2001)

© 2012 CTB/McGraw-Hill LLC (Unpublished)
Marjorie Bryant

**Academic Background**
Indiana University, Bloomington, IN - Music Performance studies
University of California Santa Cruz, Santa Cruz, CA - Music studies

**Professional Experience**

**CTB/McGraw-Hill, Monterey, CA: 2000 to present**

*Assessment Editor, Mathematics*: Selected, developed, and wrote mathematics items and scoring rubrics; prepared art manuscripts; supervised item writers; coordinated with Production, Research, and Editing departments; and participated in content review and item selection with customers.

**Winham Street Christian Academy, Salinas, CA: 1992 to 2000**


*Music Director*: Developed instrumental program for grades 6–12.

*Mathematics Tutor*: Tutored high school mathematics.

**Private Music Teacher, Saratoga, CA: 1983 to 1988**

*Teacher*: Weekly music lessons to an average of 30 students.
Kathryn Dunlap

**Academic Background**
Ph.D. General Administration University of Oklahoma, Norman, OK
M.A.T. Bilingual/ESL, Oklahoma City University, Oklahoma City, OK
B.A. History/French, College of Wooster, Wooster, OH

**Professional Experience**

**CTB/McGraw-Hill, Monterey, CA: 2011 to present**

**State Solutions Manager:** Responsible for the development and execution of state-level strategies within a number of assigned states. Includes defining and executing state plans, working in collaboration with Business Solutions, Program Management and Operations, to develop state-level business by anticipating, communicating, and contributing to winning/workable solutions and proposals for new and renewal business, and building and fostering effective long-term customer relationships.

**Pearson: 2008 to 2011**

**Director, Business Development, State Services:** Responsibilities included winning new contracts in states where Pearson did not currently have contracts. Meeting with state leaders, Provided solutions to meet the needs of state customers and responded to Requests for Proposals from these states.

**Harcourt Assessment: 2004 to 2008**

**National Measurement Consultant:** Won and maintained contracts by providing solutions and resources to meet the needs of state customers, assisting states in peer review and providing leadership in the future direction of state assessments.


**Education Program Specialist for Assessment and Accountability:** Worked with states to develop assessment and accountability systems compliant with federal requirements; developed policy and guidance

**Professional Organizations/Presentations**

State Collaborative on Assessment and Student Standards (SCASS) Comprehensive Assessment Systems (CAS), Council of Chief State School Officers (CCSSO)
SCASS Accountability Systems and Reporting (ASR), CCSSO
SCASS Technical Issues in Large Scale Assessment (TILSA)
American Educational Research Association (AERA), Politics of Education
College of Education’s Board of Advocates, University of Oklahoma
Executive Secretary, Oklahoma Curriculum Improvement Commission, Oklahoma (1995 to 2003)
Oklahoma Commission on Educational Administration, Inc.

(continued next page)
Kathryn Dunlap (page 2 of 2)

Publications


Articles Referred


Proceedings

Sample Workshops/Presentations
2006 Moderator, Multiple Languages in Large Scale Assessments: A Theoretical and a Practical Perspective, CCSSO Large Scale Assessment Conference, San Francisco

2004 Selected presenter with Kris Kaase, Mary Crovo and Lou Fabrizio, Between a Rock and Hard Place: Keeping Accountability Systems Running Smoothly When Assessments Change, CCSSO Large Scale Assessment Conference, Boston


2003 Moderator, Educational Accountability: Performance Indexes and the NCLB Act, CCSSO Large Scale Assessment Conference, San Antonio, TX

2003 Resource Person, Department of Interior, Bureau of Indian Affairs, NCLB Rule Making

2003 Invited Presenter, No Child Left Behind, Oklahoma State Department of Education’s Annual Leadership Conference, Oklahoma City

2001 Selected presenter with Gayla Hudson and Lela Odom, Improving PK-12 Classroom Experience through Curriculum Standards and Professional Development: Enacting NCTAF Recommendations, American Association of Colleges for Teachers Education (AECTE) 53rd Annual Meeting, Dallas

1998 International Leadership Team, Wellington, New Zealand

1996 Selected Presenter, with Dr. Betty Crocker, Collaborative Planning for School Change, International Consortium for Research in Science and Mathematics Education, Belize City, Belize


© 2012 CTB/McGraw-Hill LLC (Unpublished)
Karla L. Egan

**Academic Background**

Ph.D.  Sociology with emphasis in Psychometric Methods, University of Massachusetts, Amherst, MA
M.A.  Sociology, University of Massachusetts, Amherst, MA
B.A.  Sociology, Truman State University, Kirksville, MO

**Professional Experience**

**CTB/McGraw-Hill, Monterey, CA: 1999 to present**

**Research Manager (2011 to present):** Manage team of Research Scientists and Research Associates. Ensure team deliverables and budgets are met. Provide guidance to team members on psychometric issues. Manage large-scale custom contracts.

**Research Scientist III (2006 to 2010):** Used testing and measurement expertise to design, implement, and maintain large-scale testing programs. Managed psychometric challenges to projects, including test scale maintenance; problematic anchor items; test form equivalency; and vertical articulation of cut scores. Communicated psychometric issues to internal and external customers. Created framework to develop achievement-level descriptors (Egan, Schneider, & Ferrara, in press). Awarded CTB Research & Development Grants for research on anchor item evaluation and for research on cheating-detection methods.

**Research Manager (2001 to 2006):** Managed team of Research Scientists and Research Associates. Mentored, motivated, and trained team members. Ensured timelines and budgets were met by team members. Ensured CTB’s quality assurance procedures were implemented for all team projects. Managed research portion of two custom contracts. Managed and implemented standard settings. Conceived of and led the development effort to create Bookmark Pro, CTB’s proprietary software that significantly reduced turnaround time for producing standard setting materials and increased the quality and efficiency of the feedback reports produced during the standard setting workshop. Streamlined processes used for standard setting, which decreased critical-path, turn-around time by 50 percent. Awarded Employee of the Month (2005) for leadership of Standard Setting Team.

**Associate Research Scientist/Research Scientist I (1999 to 2001):** Responsible for conducting research and technical studies related to CTB’s products and contracts. This included designing and implementing specific studies (such as studies in equating, test reliability, or test validity), as well as carrying out subsequent data analysis under the supervision of Senior Research Scientists. Designed and implemented standard settings. Awarded Employee of the Month (2000) for work on standard setting.

**Educational Testing Services, Princeton, NJ: 1999**

**ETS Summer Internship:** Worked as part of an ongoing project to look at item-level fit in IRT models. The work focused on two aspects of item fit: (a) fit of an item (dichotomous or polytomous) in a single population; and (b) fit of a single item response function in multiple populations (the problem of DIF and/or item parameter drift).

**Research and Evaluation Methods Program, School of Education, University of Massachusetts: 1998 to 1999**

**Psychometric Fellowship Recipient:** Conducted research to evaluate the impact of categorization on the estimation of the latent trait with polytomous data. Simulated linear and non-linear data to replicate the conditions of a 20-item test. The non-linear data were generated using both the GRM and...
Karla L. Egan (page 2 of 4)

Professional Experience (continued)

GPCM. The linear data were simulated using a linear factor analytic model. The number of categories ranged from two categories per item to nine categories per item.

MCAS Project, Research and Evaluation Methods Program, School of Education, University of Massachusetts: 1998 to 1999

Research Assistant: Worked with H. Swaminathan, R. K. Hambleton, S. G. Sireci, K. C. Meara, and G. Rodriguez on a study commissioned by the Massachusetts Department of Education on the evaluation of the Massachusetts Comprehensive Assessment System (MCAS). Responsible for leading the research on the Grade 8 MCAS. Duties included data management, item analyses, DIF analyses, dimensionality analyses, and IRT analyses. Prepared and presented the final report for Grade 8.

HLM Project, Research and Evaluation Methods Program, School of Education, University of Massachusetts: 1997 to 1999

Research Assistant: Worked with H. Swaminathan and H. J. Rogers on developing hierarchical procedures for the analysis of polytomous data.

Research and Evaluation Methods Program, School of Education, University of Massachusetts: 1997 to 1999

Research Consultant: Assisted faculty and students in research design and analyses.

Department of Mental Retardation, Boston, MA: 1998

Statistics Consultant: Evaluated the sampling methodology used by the Department of Mental Retardation in a high-stakes certification process. The sampling design was investigated through the use of power analyses and bootstrapping techniques.

American Institute for Certified Public Accountants, Jersey City, NJ: 1997

Psychometrics Consultant: Investigated the dimensionality of the Uniform CPA Examination through the use of principal components analysis, maximum likelihood factor analysis, and multidimensional scaling.

Social and Demographic Research Institute, Department of Sociology University of Massachusetts: 1995 to 1997

Research Assistant: Worked with D. L. Anderton to investigate the residential location of industrial and hazardous sites using segregation indices, logistic regression, and multidimensional scaling.

MRC Group, Las Vegas, NV: 1996 to 1997

Statistics Consultant: Developed and analyzed surveys to determine the market needs for various clients.

Selected Publications


Karla L. Egan (page 3 of 4)

Selected Publications (continued)


Technical Reports (partial list)


Selected Presentations


Karla L. Egan (page 4 of 4)

Selected Presentations (continued)


Teaching Experience

School of Education, University of Massachusetts: Spring 1999
Instructor, Research in Education: Taught undergraduate course in research methods for education. Responsible for organizing and planning the course content, preparing and delivering lectures and classroom exercises, marking student papers, and assigning semester grades.

Division of Continuing Education, University of Massachusetts: Fall 1998
Instructor, Data Collection and Analysis: Taught undergraduate course in research methods. Responsible for organizing and planning the course content, preparing and delivering lectures and classroom exercises, marking student papers, and assigning semester grades.

Division of Continuing Education, University of Massachusetts: Summer 1998
Instructor, Elementary Statistics: Taught undergraduate course in statistics. Responsible for organizing and planning the course content, preparing and delivering lectures and classroom exercises, marking student papers, and assigning semester grades.

Professional Affiliations
American Educational Research Association, Member
National Council of Measurement in Education, Member
Northeastern Educational Research Association, Member
Heather Farina

Academic Background
M.S.  Elementary Education, Indiana University, Bloomington, IN
B.S.  Elementary Education, Indiana University, Bloomington, IN

Professional Experience

CTB/McGraw-Hill, Indianapolis, IN: 2010 to present

Content Development Lead (2010 to present): Oversee all content areas (English/Language Arts, Mathematics, Science, etc.) to ensure accurate and timely completion of content area responsibilities. Provide and/or organize needed trainings to enable completion of item and form development procedures by the content development team. Determine schedule and track progress in relation to the schedule for all content areas. Coordinate communication with internal and external customers for all content areas. Facilitate item content review meetings, rangefinding meetings, and facilitates standard setting committees using CTB’s Benchmark procedures.

Mathematics Content Lead Editor (2009 to 2010): Oversaw mathematics content editors to ensure accurate and timely completion of mathematics content editor responsibilities. Provided and/or organized needed trainings to enable completion of item and form development tasks by the mathematics content team. Conducted reviews of items and forms developed by mathematics content editors and provided feedback to ensure quality products were developed. Tracked progress in relation to the schedule. Coordinated communication with internal and external customers. Facilitated item content review meetings, rangefinding meetings, and facilitates standard setting committees using CTB’s Benchmark procedures.

Mathematics Content Editor (2004 to 2009): Responsible for preparing manuscripts and content editing of mathematics items according to CTB processes including the use of CTB general specifications, state approved standards and assessment limits. Critically inspected items to ensure quality of mathematical content. Used Microsoft Excel to track items, generate content review forms, and created form maps to ensure content coverage. Used proprietary item development systems to develop new items and select existing items for parallel operational test forms; responded to queries and reviewed pages through production cycles using these proprietary systems and processes; responsible for materials integration reviews (MIR) and proprietary data-base processes and procedures for on-going tracking and documentation of items and test forms. Facilitated item content review meetings, rangefinding meetings, and facilitated standard setting committees using CTB’s Benchmark procedures.

Indianapolis Public Schools, Indianapolis, IN: 1999 to 2004

Classroom Teacher: Grades 1, 2, 5, and 4/5 split: Planned and implemented curricula and assessments for all students, including those with special needs, in all academic areas in accordance with district and state standards and policies. Used provided materials and developed supplemental materials and strategies as needed for instruction and assessment. This included the use of Open Court Reading/Language Arts Program and the Waterford Early Literacy Program to enable all students to succeed and meet state standards. In the area of Mathematics, developed strategies for teaching problem solving, mathematical communication, measurement, algebraic concepts, number sense, geometry, and computation skills included the implementation of activities and lesson plans acquired during

(continued next page)
Heather Farina (page 2 of 2)

Professional Experience (continued)
participation in Fostering Algebraic Thinking workshops, the NCTM National Conference and staff meetings to ensure that state standards and adequate yearly progress was met. Maintained communication with parents of students.

Carmel Clay Schools, Carmel, IN: 1998 to 1999
Substitute Teacher: Implemented classroom instruction according to provided lesson plans. Covered additional duties, such as supervising lunch, as needed. Maintained classroom discipline.

Metropolitan School District of Pike Township, Indianapolis, IN: 1998 to 1999
Computer Lab Supervisor: Incorporated computer literacy skills while reinforcing academic skills in all content areas for students in Kindergarten through fifth grade. Collaborated with classroom teachers to create projects.

Classroom Teacher: Grade 1: Planned and implemented curricula and assessments for students in all academic areas in accordance with district and state standards and policies. Used provided materials and developed supplemental materials as needed for instruction and assessment. Maintained communication with students’ parents. Attended staff meetings and workshops.
Kelly E. Godfrey – College Board

Academic Background
B.S. University of North Carolina at Chapel Hill, Psychology, 2002

Professional Experience
The College Board, New York, New York and Newtown, Pennsylvania: 2007 to present
Associate Research Scientist, Research & Development
Educational Research Methodology, UNC Greensboro: 2007
Instructor, Statistical Methods in Education
Educational Research Methodology, UNC Greensboro, NC: 2006
Teaching Assistant, Multivariate Analysis
2005-current
Trainer, QSR software: N6, NVivo 2, NVivo 7, NVivo 8, NVivo 9

Graduate Research Assistant:
- Evaluation team of North Carolina Partnership for Improving Math in Schools (NC-PIMS), funded by National Science Foundation and U.S. Department of Education.
- Data analysis, including scanning, cleaning, and reporting.
- Qualitative and quantitative data collection and analysis, including survey and assessment development.

Research & Development Graduate Student Summer Intern: Compared Kernel and Item Response Theory true score methods of test equating.

ATLAST, Horizon Research, Inc: 2004 to 2006
Contract Data Analyst: Test data analysis on pilots tests using Classical Test Theory, IRT, and Multidimensional IRT.

Measured Progress, Dover, NH: Summer 2005
Psychometric Intern: Applied kernel equating methods to alternate assessment data.

Horizon Research, Inc, Chapel Hill, NC: 2002
Research Data Analyst (Intern):
- Formatted tables for publication using Microsoft Word, Excel, and SPSS.
- Data entry and output in SPSS and Excel.
- Assisted on various aspects of longitudinal studies involving math and science teachers across the United States.
Kelly E. Godfrey (page 2 of 3)

Selected Bibliography


Program Experience

SPSS SAS
Fortran 90
N6, NVivo 2, NVivo 7, NVivo 8
Systat
Program Experience (continued)

JMP
Lisrel
HLM
BILOG MG
PARSCALE
SIBTEST
NOHARM
DETECT
DIMTEST
KE (Kernel Equating)
Microsoft Office: Word, Excel (including Solver), PowerPoint, Access, etc.
Some experience with C programming
Angelica Gordon

**Academic Background**
A.A. Liberal Arts, Los Angeles Harbor College, Wilmington, CA

**Professional Experience**

*CTB/McGraw-Hill, Monterey, CA: June 2003 to present*

**Program Office Coordinator:** Responsible for providing program office support and direction for one or more programs that vary in size from $100,000 to $10 million. Focus is on the planning, management, and delivery of programs and associated services. Program types include custom contract programs, base (shelf) contract delivery programs, CTB product development programs, and CTB systems development. Acts as the “right hand” of the program manager, supporting the program manager in the fulfillment of program responsibilities. Principal duties and responsibilities include program operations support, solutions management support, sales support, financial management support, leadership and people management, team building, and maintaining internal values and relationships.
Darin Homer

Academic Background
BS Business Administration/Marketing, California State University, Chico, California
Internship, Advertising Sales, Chico, California, Chico News & Review, Newspaper

Professional Experience

CTB/McGraw-Hill, Monterey, CA: 1997 to Present
Project Manager: Works with customers at the state department of education level and program management to develop solutions for custom assessment programs. Manage and oversee project tasks, scope, schedules, and budgets and coordinate with development leads and staff in the development of high quality testing materials.
Development Supervisor: As Development Supervisor, responsibilities included planning, researching, writing, and editing test questions for small and large-scale assessments. Supervised teams of editors working on various state test development programs.
Content Editor/Writer: Developed and edited mathematics test questions for large-scale assessments according to state contract specifications. Provided editorial support to content editors and assisted with test selections, manuscripts, and test layouts.
Scoring Representative: Received incoming calls from a variety of state school districts. Assisted in selection of testing products, test score report package options, and interpretation of various test score reports.

Merchandising Coordinator: Developed and managed all aspects of new product merchandising programs and new store layout plans. Managed and maintained new and existing product merchandising presentations, displays, and concepts in retail stores nationwide.
Assistant Store Manager: Oversaw daily operations of the store including inventory management, merchandising, and customer service. Provided training for other store associates on products, customer service, and point-of-sale transactions. Managed various internal and external sales promotions and special events.
Academic Background
Ph.D.  Assessment and Measurement, James Madison University, 2009
M.A.  Psychological Sciences (Concentration: Assessment, Measurement, & Statistics), James Madison University, 2006
B.A.  Psychology (Minors: Statistics, Religion), Appalachian State University, 2003

Professional Experience
The College Board: 2009 to present
Assistant Research Scientist:  Involved with many phases of multi-year project utilizing evidence-centered assessment design in the Advanced Placement program. Develop form assembly specifications using an evidence-centered assessment design framework. Conduct research studies (design of study, data analysis, present results, write manuscript of results) using multilevel modeling to better understand test scores from College Board assessments, and relationships between these test scores and other variables of interest. Conduct, manage and/or review research and psychometric processes (e.g., use of IRT, reliability and validity studies, generalizability theory studies, equating methods on mixed format exams) in support of College Board testing programs. Plan data collections for pilot tests, evaluate results of item analyses.

Paper Publications and Book Chapters

Selected Presentations

(continued next page)
Pamela K. Kaliski (page 2 of 3)

Selected Presentations (continued):


(continued next page)
Pamela K. Kaliski (page 3 of 3)

Selected Presentations (continued):


Related Professional Activities and Awards
Outstanding Scholarship Award, Spring 2008, presented by the JMU Department of Graduate Psychology
Outstanding Scholarship Award, Spring 2006, presented by the JMU Department of Graduate Psychology
Travel award from Motivation in Education Special Interest Group, AERA Spring 2006
Best Graduate Student Poster Award, Virginia Psychological Association, Spring 2005
Selected to fulfill competitive JMU Teaching Assistantship, Fall 2005-Spring 2006
Selected to fulfill competitive JMU Teaching Assistantship, May 2005-June 2005
Selected to fulfill competitive JMU Graduate Assistantship, Fall 2004-Spring 2005, Fall 2006- Spring 2007, Fall 2007-Spring 2008

Professional Affiliations
American Educational Research Association (AERA)
National Council on Measurement in Education (NCME)
Northeastern Educational Research Association (NERA)
American Psychological Association (APA) Division 5
Association of Psychological Sciences (APS)
Dan Macomber

Academic Background
M.A. English (Teaching of Writing), Humboldt State University, Arcata, CA
B.A. English Literature, Humboldt State University, Arcata, CA
Minor Teaching English to Speakers of Other Languages, Humboldt State University, Arcata, CA

Professional Experience
CTB/McGraw-Hill, Monterey, CA: 2003 to present
English Language Arts Assessment Editor: Assessment Editor work includes developing item and passage specifications; facilitating passage and item reviews, according to specifications and quality standards; providing content correlation and content process support; and developing assessment progression across grade levels with a single content area (ELA/RLA).

Additional skills and abilities include broad-based knowledge of evidence design (test parameters, blueprint, test design), framework specifications, research statistics, and design models (as well as extensive experience facilitating rooms at standard setting meetings). Highly proficient in Microsoft Word, Microsoft Excel, Adobe Acrobat, and Monarch (CTB-internal item database and development tool).

Item Writer: Wrote hundreds of Reading/Language Arts items for a variety of custom and shelf products.

Professional Affiliations
Returned Peace Corps Volunteer (RPCV), 2000-present.
Toastmasters International: Held office of President and VP of Public Relations for CTB's Toastmasters club, 2006-present
Ricardo Mercado

Academic Background
B.S. Psychology, University of California, Davis, CA
Certificate Project Management Professional (PMP)

Professional Experience
CTB/McGraw-Hill, Monterey, CA: 2001 to present
Senior Standard Setting Specialist (2008 to present): Responsible for the planning, implementation, and documentation of all CTB standard setting activities. Since 2001, participated in the successful implementation of over 70 standard setting activities for large-scale assessments worldwide.

Research Project Manager (2006 to 2008): Responsible for the successful planning and implementation of Research activities for the California English Language Development Test; and of all CTB standard setting activities, including those in Arizona, Colorado, the District of Columbia, Georgia, and West Virginia. Held responsibility for budgetary analysis, resource planning, and task coordination among geographically diverse team members.

Standard Setting Specialist (2003 to 2006): Responsible for the implementation of standard setting activities, including associated analyses, technical studies, and training in consultation with research scientists; and providing technical support for research related to assessments under contract. Facilitated standard setting activities for numerous large-scale assessments, including those in California, Maryland, New York City, Puerto Rico, and Qatar.

Research Associate (2001 to 2003): Responsible for the preparation of materials for standard setting activities, including associated analyses, under the direction of research scientists; and providing technical support for research related to assessments under contract.

Publications
Book Chapter

Presentations


(continued next page)
Ricardo Mercado (page 2 of 2)

Presentations (continued)


Professional Organizations
Project Management Institute
M. Christina Schneider

**Academic Background**

Ph.D.  Music Education, University of South Carolina, Columbia, SC  
M.Ed.  Educational Research, University of South Carolina, Columbia, SC  
M.ME.  Music Education, University of South Carolina, Columbia, SC  
BME.  Music Education, University of South Carolina, Columbia, SC

**Professional Experience**

**CTB/McGraw-Hill, Columbia, SC: 2006 to present**

**Research Scientist:**  Conduct field test analyses, monitor form selection, and conduct field test and operational item response theory scaling and equating with experience using 1PL, 3PL, and partial credit models. Lead psychometric work for contracts using pre-equated designs with post-equated verifications and contracts using post-equated only designs. Monitor technical quality of horizontal and vertical scales. Design and oversee sampling of students for calibration samples and field test analyses. Design and conduct research studies for statewide testing programs. Produce technical reports. Design and conduct peer review approved standard settings for multiple statewide assessments in reading, mathematics, science, social studies, and English language proficiency. Design model building and implementation studies of custom and shelf automated essay scoring engines. Develop professional development programs in classroom assessment.

**South Carolina Department of Education, Columbia, SC: 2002 to 2006**

**Psychometric and Data Analysis Group Coordinator:**  Managed the psychometric and data analysis personnel within the Office of Assessment. Monitored contractor technical work for the South Carolina statewide testing program and the Office of School Quality Benchmark Assessments. Directed and co-designed in-house research studies for the statewide testing program and produced corresponding reports. Directed and edited contractor production of technical manuals. Facilitated development of performance descriptors for English language arts, mathematics, and science for the Palmetto Achievement Challenge Tests. Developed technical portions of peer review submission for English language arts and mathematics. Implemented professional development sessions on the use of item maps and development of performance descriptors for the Office of School Quality Benchmark Assessments.

**National Assessment of Educational Progress (NAEP) Coordinator:**  Conducted item tryouts and item response theory scaling (1PL) for the Office of School Quality Benchmark Assessments. Facilitated development of performance descriptors for mathematics and English language arts for the Palmetto Achievement Challenge Tests. Monitored hand scoring by contractors of constructed response items. Designed and implemented assessment-related professional development for South Carolina public schools (e.g., rubric development, standardized test interpretation, and item and test statistic use and interpretation). Produced Assessment Informational Overview, an instructional television program in South Carolina designed to update public school teachers and administrators about the large scale assessments in South Carolina. Produced and co-authored Assessing Standards in the Classroom, an instructional television series and companion document in South Carolina designed to provide professional development to public school teachers and administrators about creating standards-based, classroom assessments. Coordinated the administration of NAEP in South Carolina.
M. Christina Schneider (page 2 of 4)

Professional Experience (continued)

University of South Carolina, Columbia, SC: 2001 to 2002
Research Assistant: Coordinated The South Carolina Arts Assessment Program within the Office of Program Evaluation in the School of Education. Duties included coordinating item writing production, bias and sensitivity reviews, test book production and scoring, database management, and grant writing.

Band Director: Taught instrumental music to students in Grades 6–8.

University of South Carolina, Columbia, SC: 1998 to 1999
Research Assistant: Duties included experimental design, data collection, and grant writing for projects related to the School of Music Children’s Music Development Center.

University of Southern Mississippi, Hattiesburg, MS: 1997 to 1998
Graduate Assistant: Supervised School of Music student teachers.

Spartanburg County School District Number Seven, Spartanburg, SC: 1993 to 1997
Band Director: Taught instrumental music to students in Grades 5–9.

Professional Organizations
American Educational Research Association
National Council on Measurement in Education

Selected Presentations


(continued next page)
M. Christina Schneider (page 3 of 4)

Selected Presentations (continued)


M. Christina Schneider (page 4 of 4)

Selected Presentations (continued)


Selected Publications


Gretchen R. Schultz

Academic Background
M.L.A  History of Ideas and English, John Hopkins University, Baltimore, MD
B.A.  Humanities, Washington, College, Chestertown, MD

Professional Experience
CTB/McGraw-Hill, Monterey, CA: 9/2001 to 11/2003; 2/2005 to present Principal Assessment Editor, ELA: Contribute to the conceptualization, planning, and implementation of complex assessment contracts/programs/products in a specific content area; develop test specifications and guide the integration of industry-leading assessment practices in all phases of item development, item selection, and test construction in a specific content area; provide key product support to customers, both internal and external within CTB and across McGraw-Hill Education, interacting with customers and traveling to customer sites as needed; produce high quality exemplary materials within the established timelines; serve as subject matter expert in the response to requests for proposals, representing concerns, best practices, and thought leadership. Led English and Social Studies Publishing Group training sessions on item and assessment development and processes, content-related topics (e.g., assessing theme, assessing vocabulary, passage mapping).

Development Manager: Responsible for conceptualizing, planning, supervising and organizing the overall development of major and sometimes, multiple medium-to-large scale custom contracts or shelf products. Responsible for managing projects and staff, and for developing resources.

Maryland State Department of Education, Baltimore, MD: 2003 to 2005
Technical Assessment Specialist: Associate manager of operation of state assessment programs.

English Assessment Specialist 1998 to 2001: Development manager of statewide high-stakes English assessment; Technical Assistance Facilitator: Instructional and assessment facilitator of staff development to local school systems

Queen Anne’s County Board of Education, Centerville, MD: 1981 to 1998
English Teacher: Classroom teacher, Grades 7–12; Instructional leader; Gifted and Talented Program lead; Local assessment lead; County representative to state assessment teams; Curriculum writer.

Prince George’s Community College, Largo, MD: 1992 to 1996
English Adjunct Professor: Evening and summer school instructor.

Anne Arundel County Board of Education, Annapolis, MD: 1972 to 1979
English teacher: Classroom teacher, Grades 10–12.

Professional Presentations
District of Columbia Public Schools, October 2006: “Scoring Student Writing”
NCTE Convention: workshop coordinator and presenter, November 2001; “Is It Really Teaching to the Test?”
CCSSO Convention: presenter, June 2001; “Advantages and Disadvantages of Staff Development via State Assessments”

(continued next page)
Gretchen R. Schultz (page 2 of 2)

Professional Presentations (continued)
ASCD Convention: presenter, March 2000; “Goals, Instruction, and Assessment: Putting It All Together”
MSTA Convention: presenter, October 1999; “High School Reading Skills and High-Stakes Assessments”

Professional Affiliations
National Council of Teachers of English
Association for Supervision and Curriculum Development
Phi Delta Kappa International
American Association of University Women

Achievements
Recipient of CTB/McGraw-Hill Silver Award for Lead by Example, 2nd quarter 2011
Member of CTB Employee Advisory Committee
Kevin P. Sweeney – College Board

Academic Background
Ph.D.  Psychometrics, Fordham University, New York, NY, 1996
M.A.  Psychometrics, Fordham University, New York, NY, 1989
B.A.  Psychology, St. Bonaventure University, St. Bonaventure, NY, 1985

Professional Experience
The College Board, New York, NY
Executive Director of Psychometrics, Research and Development (2009-present): Since joining the College Board in 2009, manages a staff of over a dozen research scientists and psychometricians, providing supervision and leadership to Psychometric staff within the Research and Development Division. Responsibilities have included implementation and oversight of all internal psychometric procedures, serving as an internal consultant and resource to both psychometric and non-psychometric staff, and representing College Board at national and regional conferences. Work with staff and external consultants to design and implement standard setting studies for the Advanced Placement and Accuplacer programs. Oversee shadow equating and other psychometric analyses of the SAT and all College Board programs. Work closely with external advisory committees on implementing recommendations for the improvement of College Board assessment programs.

Measured Progress, Dover, NH
Vice President, Research and Analysis (since 2005). Prior position was Director of Measurement, Design, and analysis (1998-2005): Responsible for the management of psychometric activities to support corporate initiatives and state-level contracts, including score reporting, scaling, equating, data analysis, and standard setting. Regularly met with clients and prospective clients to design solutions that best meet the needs of their testing programs. Responsible for providing supervision to psychometric and data analytic staff. Oversaw and directed the development of corporate psychometric and data analytic systems. Directly oversaw and participated in the planning and conduct of hundreds of standard setting panels, including those for general k-12 statewide assessments, alternate assessments, and English language proficiency assessments. Recognized as a corporate leader who provides measurement solutions to real world client problems. Proven excellent communication skills, routinely making presentations to convey complex psychometric and measurement topics to audiences ranging from lay people to nationally recognized measurement experts. Represented Measured Progress at Technical Advisory Committee meetings. Demonstrated effective management history as evidenced by low departmental turnover rate and consistently achieving departmental growth rate lower than corporate growth rate without sacrificing quality or timeliness of work. Provided technical support and guidance to all major corporate initiatives including developing a local market assessment tool, serving as technical advisor to federally funded enhanced assessment grants, and investigating various techniques for supplying diagnostic feedback to students based on assessment results. Instrumental in establishing the national reputation of Measured Progress as a leader in the measurement field through various means including making presentations at national conferences, encouraging staff to participate in national conferences, submitting publications to peer reviewed journals, and participating in professional organizations. As a senior technical advisor, represent Measured Progress to clients and potential clients.

(continued next page)
Kevin P. Sweeney (page 2 of 3)

Professional Experience (continued)

American Institute of Certified Public Accountants, New York, NY

Assistant Director, Psychometrics (1996-1998). Prior position was psychometrician (1992-1996): Responsible for overseeing all operational aspects of the psychometric work at AICPA. Initiated preliminary efforts to develop and implement a computerized version of the Uniform CPA Examination. Ensured that all phases of data analysis and reporting were completed accurately and on schedule. Established the psychometric research agenda and authored several reports, technical manuals, and other documents related to the psychometric work done to support the Uniform CPA Examination, the Accredited Business Valuation examination, and the International CPA Qualification Examination. Planned and executed standard-setting studies, reliability studies, validity studies, equating analyses, and item and test analyses. Presented and explained psychometric information to examination policy-making bodies as well as psychometric and lay audiences. As psychometrician, performed all operational psychometric work for the licensure and credentialing examinations produced by the AICPA and designed and implemented psychometric research studies for the development of those examinations. Operational work included analyzing standard-setting data, performing equating analyses, and performing item and test analyses. Responsible for writing and maintaining the computer programs used to conduct these analyses.

Research

Specific areas of research include uses and development of standard setting methodologies, applications of item response theory, comparisons of equating models, and efficacy of differential item functioning.

Selected Presentations and Publications


(continued next page)
Kevin P. Sweeney (page 3 of 3)

Selected Presentations and Publications (continued)


Professional Affiliations

- American Educational Research Association
- National Council for Measurement in Education
- Northeastern Educational Research Association
Kristal Taylor

Academic Background
B.A.  English, Spanish minor, California Polytechnic State University, San Luis Obispo, CA

Professional Experience
CTB/McGraw-Hill, Monterey, CA: 1999 to present

Program Schedule Analyst: Responsible for supporting program and project managers in creating, maintaining, and sustaining resource-loaded schedules for multiple simultaneous programs at CTB. Includes product development programs, system development programs, custom contract programs, shelf contract programs, and internal cross-company initiatives. Reports to the program schedule manager and is matrixed to multiple programs to provide program administration support. Reporting to the program managers, and in partnership with project managers, program office coordinators, and other program staff, is responsible for schedule creation, quality, maintenance, and security. As a Program Management team member, is also responsible for participating in program activities including policy and procedure creation, strategic and operational planning, and related activities.
Natasha Vasavada - College Board

**Academic Background**
M.Ed  Education Policy and Leadership, Teachers College, Columbia University, New York, NY 2010
M.A.  Educational Leadership, Fairleigh Dickinson University, Madison, NJ 2003
B.A.  History and German, Davidson College, Davidson, NC 1999

**Professional Experience**


**Executive Director, Standards and Curriculum Alignment Services:** Provides standards, curriculum, and alignment expertise across the organization related to curriculum and assessment. Develops alignment and assessment strategies based on the Common Core State Standards, PARCC and SBAC assessment consortia, and Next Generation Science Standards. Provides guidance and technical support to states and districts regarding alignment and implementation of the Common Core State Standards and state standards. Provides strategic product management for the design and implementation of the College Board Standards for College Success in English Language Arts, Mathematics and Statistics, and Science. Provides face-to-face status updates and reports to the senior leadership team at the College Board. Develops strategic solutions for aligning College Board programs and services to allow them to function as a cohesive college readiness system. Supervises content, curriculum, assessment, standards specialists within Research and Development. Develops and delivers presentations to state and national audiences on College Board programs and services. Builds and manages fiscal budgets and project schedules.

**Senior Director, College Board Standards for College Success:** Provided strategic management and leadership for the College Board Standards for College Success and developed alignment tools, capabilities, and solutions. Provided standards alignment expertise across the organization as it relates to curriculum and assessment. Developed strategic state and national partnerships by providing standards guidance and alignment tools. Provided consultation and guidance on the AP History Course and Exam Review Process. Delivered presentations to state and national audiences on College Board programs and services. Built and managed fiscal budgets and project schedules.

**Livingston Public Schools, Livingston, NJ 2005-2007**

**K-12 Social Studies/History District Supervisor:** Provided overall program and teacher supervision, curriculum development, and professional development for K-12 Social Studies/History program (supervising 70+ teachers across 9 schools, including middle and high schools). Conducted teacher performance reviews using the clinical and coaching models; executed over 100 classroom observations and 30+ performance reviews yearly. Designed and conducted K-12 program evaluation and curricular review of Social Studies program. Led K-12 curriculum review, revision, and writing process with an emphasis on case-studies in history, document-based analysis using primary sources, strengthening the humanities and geography components of the curriculum, and increasing the depth of conceptual study over broad survey courses. Created and or redesigned the following curricula: Middle Eastern Studies, Historical Research Seminar, World History, Geography, US History, and AP Government and Politics.

**Scotch Plains Fanwood High School, Scotch Plains, NJ 2000-2005**

**Teacher of Social Studies/History:** Taught and developed courses for grades 9-12 at the College Prep., Honors, and AP levels. Designed interdisciplinary Humanities course aligned to Social Studies and
Natasha Vasavada (page 2 of 3)

Professional Experience (continued)

English standards. Embedded technology, collaborative and project-based learning, and student portfolios into instruction.

National Standards and Assessment Advisory Activities

New York State Education Department

Project Director: Providing standards, alignment, and assessment technical guidance to the New York State Education Department as they shift their assessments to be more aligned to the Common Core State Standards.

Achieve – GED Social Studies Assessment

Chief Content Advisor: Providing content, assessment, and standards expertise to Achieve and the GED program as they seek to align the assessment to college and career readiness knowledge and skills, including the Common Core State Standards for Literacy in Social Studies/History, Science, and Technical Subjects.

Common Core State Standards

Advisory Committee and Content Teams: Served on advisory panel to guide policy decisions related to the development of the Common Core State Standards in English Language Arts and Mathematics. Served on the content development teams and feedback groups.

Advanced Placement European History National Commission

Content Expert: Served on 12-member national commission to examine and revise the content domain of the AP European History test as part of the Course and Exam Review Process

Selected Publications and Presentations


Developing Common Core State Standards (Cincinnati, 2009) Presentation to the National State Boards of Education Association.


Importance of Standards: Colorado’s Alignment to College Readiness Standards (Denver, 2008) Presentation to state Department of Education officials, Colorado district educators, and higher education representatives.


© 2012 CTB/McGraw-Hill LLC (Unpublished)
Natasha Vasavada (page 3 of 3)

Licensures and Certifications

Public School District Administrator (New Jersey, 2003)
Principal/Supervisor (New Jersey, 2003)
Teacher of Social Studies (New Jersey, 2000)
Teacher of German (New Jersey, 2000)
Lindy L. Wienand

Academic Background
B.S. Human Environmental Sciences, University of Alabama, Tuscaloosa, AL

Professional Experience

CTB/McGraw-Hill, Monterey, CA: 2004 to present

Program Manager (2008-present): Experience with a number of statewide assessment programs, various shelf programs such as the Department of Defense Education Activity contract. Responsible for the management and oversight of all program activities including schedule creation, risk identification and mitigation, decision documentation and record keeping, monitoring project implementation progress, establishing priorities, acting as a team leader for representatives of all departments involved, and interfacing with and ensuring the satisfaction of the customer. Collaborates with and manages a staff of cross functional professionals to execute contract requirements, plans and controls all contract events to meet required timelines, monitors and controls costs to complete projects within budget, while ensuring deliverable quality.

Program Management Specialist (2007-2008): Performed Program Management support for custom contracts. Duties included, but not limited to, budget responsibility/accountability, ETC/EAC analysis, COS management, sell-up targets, ROS and gross margin maintenance; contract management, change management, customer interface, and project team management. Prepared documentation and reported contract status to senior management through monthly program status meetings.

Program Coordinator (2004-2007): Provided support to the program manager in the fulfillment of custom and shelf contract management responsibilities. Planned, established and maintained program communications with internal and external customers. Identified, gathered, organized and maintained program information, specifications, and deliverables. Collected status data used for preparation and presentation for internal and external audiences. Planned, coordinated, and facilitated customer meetings and committee workshops. Review of program portals to reflect current programs documents.


Branch Manager: Managed a team of four, three staffing coordinators and one payroll specialist. Established and build rapport with area businesses for Westaff placements. Provided service to over seventy five accounts for client satisfaction, attended to current business trends, and identified peak points of activity. Interacted with executive level management to discuss expectations of the branch and strategized ways to obtain goals set to meet and exceed revenue targets. Facilitated monthly development sessions with branch personnel. Handled all A/P, A/R and payroll related tasks.

Hyatt Regency Monterey, Monterey, CA: 2000 to 2000

Group Sales Manager: Responsible for a $1,200,000.00 room sales quota per year for a 575 room resort/conference center property. Responsibilities included managing group size of 80 rooms or less

(continue next page)
Professional Experience (continued)

per night, soliciting corporate business located in Silicon Valley, and attended industry trade shows and special events. Negotiated group contracts requesting rooms and Food and Beverage. Supervised and developed Administrative Assistants. Achieved 102 percent of quota for last quota period worked.


Meeting Connections Manager: Responsible for a $750,000 room sales quota and $165,000 food and beverage quota per year for a 365 room property with 15,000 square feet of meeting space. Markets served included corporate, religious, government, equine, social, medical, educational and national/state associations. Handled all groups requiring 10 to 50 sleeping rooms and provided convention service and catering planning for each. Finished at 100 percent of quota for last quota period worked.


Administrative Assistant: Acted as Administrative Assistant to the Director of Sales and Sales Managers, generated contracts, proposals, VIP reservations and general correspondence. Created and distributed quarterly department Newsletter to over 300 clients. Consistent quality performance with the coordinating efforts in a variety of market segments including Group Sales, Meeting Connections, and Individual Travel Sales Markets.
Andrew Wiley – College Board

**Academic Background**

Ph.D.  Psychometrics, Fordham University, NY
M.A.    Psychometrics, Fordham University, NY
B.A.     Psychology, LaSalle University, PA

**Professional Experience**

**The College Board, New York, NY: 2002 to present**

**Executive Director, Research & Development (since January of 2007, previous positions included Senior Director, Associate Research Scientist):** Responsible for the management of a team of nine full-time test development staff responsible for final sign off of all SAT and PSAT/NMSQT test forms. Staff also works with other College Board programs, such as ACCUPLACER, CLEP, and SpringBoard, to design and complete a variety of new assessment products. Responsible for the management of a team of four full-time staff members charged with providing content expertise associated with curriculum, alignment and standards. Staff members are content experts in English Language Arts, and Mathematics. Provide assessment design and psychometric expertise in the development of a series of interim assessments to be piloted with the AP Biology program in fall of 2011. Provide assessment design and psychometric expertise for College Board programs as they design strategic plans and explore and evaluate the feasibility of creating new and revised assessments and products. Provide College Board program support by responding to media inquiries, preparing reports for external constituents, and representing the College Board at professional conferences. Previous responsibilities have include the management of a team of eight full time staff members and three part-time interns responsible for the production of research investigating the reliability, validity, and fairness of the SAT and PSAT/NMSQT examination programs. Managed a team of project managers and business analysts who set schedules, determined business requirements and allocated internal resources for all projects related to aggregate reporting for the College Board. Managed a team of staff members that determined the content for aggregate reports that summarize SAT, PSAT, and AP data for schools, districts, and states. This involves formal market research, such as focus groups and on-line surveys, as well as informal review of reports with internal College Board staff and external consumers of the reports.


**Adjunct Professor, Department of Psychology:** Teach a Research Methods course in the graduate program in forensic psychology.


**Adjunct Professor, Department of Psychology:** Taught an Introduction to Testing course in the graduate program in education.

**American Board of Internal Medicine, ABIM, Philadelphia, PA: 1999 to 2002**

**Psychometrician, Research and Psychometrics:** Coordinated all activities related to the scoring, standard setting, and equating for four medical specialty examinations. Worked with several ABIM committees and ABIM staff to develop and maintain multimedia take-home examinations designed to evaluate the clinical skills of practicing physicians. Designed and conducted research investigating the validity of the medical subspecialty examinations. Answered questions brought to the Board by examinees related to the scoring and standard setting for the medical subspecialty examinations.
Andrew Wiley (page 2 of 4)

Professional Experience (continued)

American Institute of Certified Public Accountants, Jersey City, NJ 1996 to 1999

Psychometrician, Research: Managed all work associated with the standard setting and equating of the Uniform CPA Examination. Coordinated the activities of research assistants who assisted in operational and research work performed by the Examinations team staff. Worked with several AICPA committees to develop examinations to license chartered accountants in Canada and Australia to practice in the United States and to certify CPAs as experts in Personal Finance and Business Valuation. Produced research designed to investigate the integrity of the current Uniform CPA Examination and to investigate the feasibility of alternative designs for the future examination.

Association of American Medical Colleges, Washington DC 1995 to 1996

Research Associate, Examination Department: Performed analyses and generated reports for the Medical College Admission Test (MCAT) Predictive Validity Studies. Coordinated organizational activities devoted to the Writing Sample Delivery System, a system designed to provide medical school administrators with direct access to examinee’s MCAT Writing Sample. Developed and performed a study designed to analyze the differential predictive validity of the MCAT for under-represented minorities and women.


Research Associate, Examinations Program: Developed a sampling design for item tryout for the General Equivalency Diploma (GED) examination. Performed a dimensionality analysis of GED Test Battery. Performed reliability and dimensionality analyses of multiple-choice and performance based tests. Determined appropriate weights to assign constructed response and multiple-choice sections of the GED Writing Skills Test using both IRT and CTT models.

Port Authority of NY and NJ, New York, NY 1993 to 1994

Psychometric Assistant, Examinations Program: Assisted in the development and grading of written and oral examinations for promotion to Police Lieutenant and Police Sergeant. Examined the utility of the Differential Aptitude Test and the SRA Achievement Tests for use in the selection and placement of General Maintainer employees.

Publications and Presentations


© 2012 CTB/McGraw-Hill LLC (Unpublished)
Andrew Wiley (page 3 of 4)

Publications and Presentations (continued)


Andrew Wiley (page 4 of 4)

Publications and Presentations (continued)


Professional Affiliations

American Educational Research Association

American Psychological Association, Division 5 - Evaluation, Measurement, and Statistics

Association of Test Publishers – Elected to Board of Trustees, effective February, 2010

National Council on Measurement in Education

Personnel Testing Council of Metropolitan Washington

© 2012 CTB/McGraw-Hill LLC (Unpublished)