NCSC Summative Assessment Overview

March 2014
NCSC Theory of Action

Long-term goal:
To ensure that students with significant cognitive disabilities achieve increasingly higher academic outcomes and leave high school ready for post-secondary options.

A well-designed summative assessment alone is insufficient.

To achieve this goal, an AA-AAS system also requires:

- Curricular and instructional frameworks
- Teacher resources and professional development
NCSC Overall Timeline January 2011-October 2015

Year 1 (2011): **Content Model Phase**: Define model of domain learning in math/ELA for these students, identify prioritized content for assessment

Year 2 (2012): **Principled Design Phase**: Design Patterns, Task Templates, C/I/PD design and pilot; Technology architecture design

Year 3 (2013): **Item and Test Development Phase**: Task Template Tryouts, Item Specs/item development/item reviews, Student Interaction Studies (SIS), Draft grade level PLDs, finalize pilot/field design, Tech build

Year 4 (2014): **Pilot, Field, Research Phase**:
  – Pilot Phase I: National Sample, generate item statistics Winter/Spring 2014, Finalize blueprints, revise items, assemble forms
  – Pilot Phase II: Field Test Forms Fall 2014, finalize administration training and supports

Year 5 (2015): **Operational administration of NCSC assessments**
  – Summer 2015: Set Standards
  – Fall 2015: Technical reporting complete
Curricular and Instructional Resources

• Provide guidance on how to “unpack” the instructional and assessed content;

• Promote strategies and resources for teaching challenging academic content through professional development opportunities; and

• Align challenging and attainable content that is observable and measurable for use in instruction and a thorough system of assessments.
Quality Indicators for Instructional Resources

• Promote Common Core State Standards;
• Set high expectations for all students;
• Apply principles of Universal Design for Learning (UDL); and
• Apply evidence-based teaching practices for students with the most significant cognitive disabilities.
Quality Indicators for Instructional Resources

• Use general curriculum resources and general education content experts’ review;
• Offer options for ALL students in the 1%;
• Provide a teacher-friendly resource that promotes effective instruction.
General Description of Assessment System

• Summative math and ELA tests for 3-8, 11 administered in a two-month window in spring
  – Up to 30 items, approximately 1.5-2 hours per test

• Technology delivery with teacher test facilitator/administrator

• universal design features and accommodations guidelines derived from Design Pattern/Task Templates Tryouts and Student Interaction Studies
Assessment Administration

• Presented via computer with flexibility for presentation on devices/platforms (e.g. tablets) available for the operational test.
• It is expected that most students will interact with an examiner during the administration. Other students may respond to the test items directly via interaction with computer presentation.
• Prior access to summative content will be provided to support examiners preparation for accommodations/adaptations.
• For most students, it is expected that testing time will be no more than approximately 1.5 to 2 hours per content area, divided between at least two sessions with flexibility to stop and resume.
  – Some students will qualify to take a shorter assessment based on evidence collected before and during the assessment.
Item Types

• Most of the items will be machine-scored, multiple choice.

• A small number of the items will require human scoring using a scoring rubric. In most cases, these items are open-response activities with correct answers; each student would complete one writing prompt for which application of a rubric is necessary.
  – All of the human scored items will be evaluated by the examiner during the assessment and scores entered.
  – A 20% audit of human scored items is expected
Assessment Outcomes

• **Total score and performance level** for each of mathematics and ELA (reading and writing).
  – Comparable within year and across years.

• **Writing is still being decided** (raw score and/or narrative description of student performance)
Assessment Design Illustration

**Pre-Session**
Prior to testing, educators input learner characteristic and/or performance data into the system.

**Preliminary Items:**
One or two sample items immediately prior to the assessment to promote familiarity and engagement.

**Session IA**
Start with a small number of items at a low level of complexity.

**End Rule:** If student is not responsive and evidence from both pre-session and preliminary items indicates meaningful interaction with the assessment is unlikely, the early stopping rule is invoked.

**Session 1 Continued**
Broad representation of content across full range of complexity and difficulty.

**Session 2A (OPTION)**
Broad distribution of content, but heavier emphasis on items with lower levels of complexity and difficulty.

**Session 2B (OPTION)**
Broad distribution of content, but heavier emphasis on items with higher levels of complexity and difficulty.

OR
# Blueprint Illustration - Grade 3 Math

<table>
<thead>
<tr>
<th>Emphasis</th>
<th>Domain</th>
<th>CCC(s)</th>
<th>Target by Item Type (Levels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>Geometry</td>
<td>3.GM.1i1</td>
<td>SR: 2 (1-3)    Multi-SR: 1 (4) CR: 1 (4)</td>
</tr>
<tr>
<td>20%</td>
<td>Measurement and Data</td>
<td>3.ME.1d2</td>
<td>3 (1-4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.DPS.1g1</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 (2-4)</td>
</tr>
<tr>
<td>20%</td>
<td>Number and Operations Fractions</td>
<td>3.SE.1g1</td>
<td>3 (1-4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.NO.1l3</td>
<td>3 (1-4)</td>
</tr>
<tr>
<td>20%</td>
<td>Numbers and Operations Base Ten</td>
<td>3.NO.1j3</td>
<td>3 (1-4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.NO.2c1</td>
<td>1-2 (1-2) 1-2 (3-4)</td>
</tr>
<tr>
<td>30%</td>
<td>Operations and Algebraic Thinking</td>
<td>3.PRF.2d1</td>
<td>3 (1-4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.NO.2d3</td>
<td>3 (1-4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.NO.2e1</td>
<td>3 (1-4)</td>
</tr>
</tbody>
</table>
# Item Development

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number of Tasks</th>
<th>Number of Levels/Tiers</th>
<th>Number of Items per Level/Tier for each Task</th>
<th>Total ELA Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>HS</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number of Tasks</th>
<th>Number of Levels/Tiers</th>
<th>Number of Items per Level/Tier for each Task</th>
<th>Total Math Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
<tr>
<td>HS</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>160</td>
</tr>
</tbody>
</table>

**Grand Total** | **2240**
Student Interaction Studies

• The project is conducting a series of focused studies to inform ongoing development.

• Research questions will include:
  – To what extent do item levels represent a progression of difficulty/complexity?
  – To what extent does performance vary by skill/context?
  – What item/presentation features promote access?
Pilot Test – Spring/ Fall 2014

• **Phase I:** Items will be administered broadly in a ‘matrix’ design for the primary purpose of evaluating item performance and developing bank calibrations.

• **Phase II:** Drawing on the item calibrations from Phase I, intact form/sessions will be constructed and administered. By so doing, these forms will be built to meet psychometric targets and a decision rule governing progression from session one to two can be applied.
Census Testing – Spring 2015

• Full operational assessment in the spring of 2015
• Standard setting will be conducted based on results in summer 2015
• Reports and technical documentation will be produced
Summative Assessment Products
Following Grant

At the end of the project’s grant funding by fall 2015, states will have:

• A minimum of two forms per grade and content area suitable for future operational use.
  – Information necessary to direct presentation of content in such a manner as to honor the test specifications, blueprints, and psychometric targets for the assessment.
  – Algorithms, rules, and/or tables necessary to produce overall scale scores and performance levels.
  – Rubrics for all human scored items as well as protocols for training scorers and implementing the scoring process.
• Design specifications for all static reports produced to include individual student reports.
• Ancillary material to support administration of summative assessment to include test administration manuals, resources to support training of test examiners, and score interpretation guide.
• Detailed technical documentation of process, procedures, and results from all test development activities.
• All test items developed for the NCSC assessments (i.e. the item bank) will be made available in a format that meets industry standards for interoperability.
• Specifications for certification and implementation of the technology system and training resources.