The Future of Assessment in Connecticut

The Connecticut Assessment Forum

Shelbi K. Cole, Ph.D.
Gail Pagano

Overview

- SMARTER Balanced Assessment Consortium (SBAC) and Connecticut’s Role
- Computer Adaptive Testing (CAT)
- Innovative Item Types
- Artificial Intelligence (IA) Scoring
- Preparing Students to be College and Career Ready
In 2009, the U.S. Department of Education announced a Race to the Top Assessment Program to provide funding to consortia of States to develop assessments that are valid, support and inform instruction, provide accurate information about what students know and can do, and measure student achievement against standards designed to ensure that all students gain the knowledge and skills needed to succeed in college and the workplace.
History and Connecticut’s Role

In June 2010, 31 states, including Connecticut, joined together to form the SMARTER Balanced Assessment Consortium and submitted an application for the Race to the Top Assessment competition. The Department awarded two Comprehensive Assessment Systems grants in September 2010; one to SBAC and the other to the Partnership for Assessment of Readiness for College and Careers (PARCC).

Connecticut SDE

Why Did CT Choose SBAC?

The original PARCC system included through-course assessments that require a prescriptive approach to curriculum and instruction.
The PARCC System

The Partnership for the Assessment of Readiness for College and Careers (PARCC) Design

**English Language Arts and Mathematics, Grades 3-11**

**PARCC RESOURCE CENTER**: A digital library of released items, summative assessments, model content frameworks, instructional and formative tools and resources, student and educator tutorials, and performance-based modules.

**Thematic Resource Center**: Professional development materials, and an interactive report generation system.

**Summative Assessment Components**:
- **ELA Writing**: Tests the ability to construct arguments and narratives, and to apply research strategies.
- **Math**: Tests the ability to solve multi-step problems and to apply mathematical reasoning in real-world contexts.
- **Science**: Tests the ability to conduct experiments, analyze data, and interpret results.
- **Social Studies**: Tests the ability to understand and evaluate historical events and theories.

**End-of-Year Comprehensive Assessment**: The final summative assessment for the year, it includes all content areas.

**Smarter Balanced**: The test used by PARCC, it includes both summative and formative assessments.

**PARCC** is the Partnership for the Assessment of Readiness for College and Careers, a national initiative to develop high-quality assessments that align with common core standards for English Language Arts and Mathematics.

**End of Year Comprehensive Assessment**: All students will take the summative assessment, it will be scored by computer, and results will be available within two weeks.

**PARCC Score Reports**: Available online, they provide detailed information about student performance.

**Development**:
- For detailed information on PARCC, go to [http://parcc.org](http://parcc.org).
Connecticut and SBAC

Connecticut is a governing state in SBAC and is taking an active role in the consortium through representation on several workgroups.

Work Groups

1. Transition to Common Core State Standards
2. Technology Approach
3. Assessment Design: Item Development
4. Assessment Design: Performance Tasks
5. Assessment Design: Test Design
6. Assessment Design: Test Administration
7. Reporting
8. Formative Processes and Tools/Professional Development
9. Accessibility and Accommodations
10. Research and Evaluation
The System

Digital Clearinghouse of formative tools, processes and exemplars; released items and tasks; model curriculum units; educator training; professional development tools and resources; scorer training modules; and teacher collaboration tools.

Scope, sequence, number, and timing of interim assessments locally determined

Re-take option

* Time windows may be adjusted based on results from the research agenda and final implementation decisions.

Source: http://www.ets.org

Assessment System Components

Assessment system that balances summative, interim, and formative components for ELA and mathematics:

**Summative Assessment (Computer Adaptive)**
- Mandatory comprehensive assessment in grades 3–8 and 11 (testing window within the last 12 weeks of the instructional year) that supports accountability and measures growth
- Selected response, short constructed response, extended constructed response, technology enhanced, and performance tasks

**Interim Assessment (Computer Adaptive)**
- Optional comprehensive and content-cluster assessment
- Learning progressions
- Available for administration throughout the year
- Selected response, short constructed response, extended constructed response, technology enhanced, and performance tasks

**Formative Processes and Tools**
- Optional resources for improving instructional learning
- Assessment literacy
Connecticut SDE

Assessment: Knowing Your Purpose

- Summative – Snapshot; Accountability; High-level analyses of trends (e.g., district or state-level)?

- Interim – Actionable feedback; progress monitoring; district- or school-level comparisons?

- Formative – Instructionally useful information; class- or student-level; qualitative or quantitative?

Formative Assessment

- Formative assessment is a *process* used by teachers and students *during* instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievement of intended instructional outcome.

- In other words, there is no such thing as “a formative test” or “a formative assessment.”
Connecticut SDE

Formative Assessment

- A second important part of the definition is its unequivocal requirement that the formative assessment process involve both teachers *and* students.
- The students must be actively involved in the systematic process intended to improve their learning.

Source: *Attributes of Effective Formative Assessment*, Council of Chief State School Officers (CCSSO) Formative Assessment for Students and Teachers (FAST) Collaborative
Connecticut SDE

Content Specifications

- The SBAC Content Specifications are designed to provide clear and rigorous prioritized assessment targets that translate the grade-level standards into content/curricular frameworks along a learning continuum.
- They are designed to clarify the connections between instructional processes and assessment outcomes.
- [http://www.k12.wa.us/smarter/](http://www.k12.wa.us/smarter/)

Bureau of Student Assessment

Computer Adaptive Testing
What is an Adaptive Test?

An Adaptive Test is a test that dynamically adjusts to the trait level of each examinee as the test is being administered.

Pop Quiz

The first adaptive test, which has been in continuous use (with some revision) since it became available, was released in

A. 1905
B. 1970
C. 1984
D. 1990
E. 2007
### Pop Quiz

The first adaptive test, which has been in continuous use (with some revision) since it became available, was released in:

A. 1905 (Correct!)
B. 1970 (First adaptive test on computer)
C. 1984 (CAT software released for public use)
D. 1990 (ASVAB administered by computer)
E. 2007 (First CAT on the internet)

### The First Adaptive Test

**Developed in France around 1905, Alfred Binet’s IQ Test**

- Was individually administered by a trained psychologist,
- Was later published as the Stanford-BinetIQ Test,
- Is still in use in schools and clinics today,
- Is the standard against which IQ tests are compared, and
- Incorporates all the elements of an adaptive test.
Schematic Structure of Binet’s Item Bank

**Connecticut SDE**

What Have We Learned from Binet’s Test?

- It is based on a calibrated item bank.
- A different starting point can be used for each examinee.
- It uses an adaptive item-selection procedure.
- It uses a scoring method that allows a common score to be obtained from different subsets of items.
- Length can vary across examinees with the use of a variable termination rule.

**Bureau of Student Assessment**
A More Efficient Version of Binet’s Test

History

- Adaptive testing disappeared when large-scale group-based testing took over (World War I and II)
- First computerized adaptive tests in early 1990s
  - Armed Services Vocational Aptitude Battery (ASVAB)
  - Graduate Record Examination (GRE)
  - Graduate Management Admission Test (GMAT)
## What is CAT?

Administered by computer, a **Computerized Adaptive Test (CAT)** dynamically adjusts to the trait level of each examinee as the test is being administered.

## Components of a CAT

1. A pre-calibrated item bank
2. A starting rule for selecting the first item
3. A procedure for scoring item responses and estimating trait level
4. A method of selecting the next item
5. A rule for ending the test
Connecticut SDE

Understanding the Range of Item Complexity

Mike is using cubes that measure $\frac{1}{2}$ inch on each side to fill a box that has a height of 5 $\frac{1}{2}$ inches, width of 3 inches, and length of 2 $\frac{1}{2}$ inches. How many $\frac{1}{2}$ inch cubes will Mike need to fill the box?

Mike is using cubes that measure $\frac{1}{4}$ inch on each side to fill a box that has a height of 5 $\frac{1}{4}$ inches, width of 3 inches, and length of 2 $\frac{1}{2}$ inches. How many $\frac{1}{4}$ inch cubes will Mike need to fill the box?

Mike is using cubes that measure $\frac{1}{2}$ inch on each side to fill a box that has a height of 5 $\frac{1}{4}$ inches, width of 3 inches, and length of 2 $\frac{1}{2}$ inches. How many $\frac{1}{2}$ inch cubes will Mike need to fill the box?

Content Standard: 6.G.2
Practice Standards: MP.1; MP.5; MP.7

---

Connecticut SDE

Advantages of CAT versus Conventional Tests

- **Efficiency:** CATs are more efficient than conventional tests—they generally reduce test length by 50% or more.

- **Control of measurement precision:** A properly designed CAT can measure or classify all examinees with the same degree of precision.
Connecticut SDE

Advantages of CAT versus Conventional Tests

- **Reporting:** Results can be made available more quickly

- **Test Security/Item Exposure:** All students are not presented with the same items

- **More Flexibility for Computer Capacity:** Students do not need to be assessed on the same schedule
### Psychological Issues

- CAT equalizes the psychological environment of the test across ability levels.
  - High-ability students will get about 50% correct.
  - Low-ability students will get about 50% correct.

---

### “Public Relations” Issues

- Students cannot change their answer to an item once they have submitted it.
  - They should be told in advance of starting the test.
- Because CAT is dynamic, it can recover from an occasional student error in answering an item.
  - Literature shows little or no gain from answer changing.
Innovative Item Types

Technology Enhanced Items employ technology to:

- Present information to the student (e.g., through animations, simulations, on-line access to information, video or audio stimulus, moveable models)
- Elicit a response from the student (e.g., selecting one or more points on a graphic, dragging and dropping a graphic from one location to another, manipulating a graph)
- TE Items employ technology to assess content, cognitive complexity and depth of knowledge not assessable otherwise. Because of the cost in development, scoring, and ongoing calibration, SBAC will employ TE in situations in which static SR and static CR are inadequate.
Taxonomy of Innovative Item Types (University of Oregon)

http://pages.uoregon.edu/kscalise/taxonomy/taxonomy.html
Artificial Intelligence (AI) Scoring

- All constructed response items in the CAT will be AI scored
- Test items that cannot be accurately scored with AI will be delivered outside of the CAT engine (e.g., some elements of performance tasks)
- SBAC will require 10 - 20% read behind to ensure accuracy
- AI scoring is nearly 100% reliable
**Connecticut SDE**

**SBAC Assessment System**

**Components**

- Summative assessments benchmarked to college and career readiness
- Teachers can access formative processes and tools to improve instruction
- All students leave high school college and career ready
- Interim assessments that are flexible, open, and provide actionable feedback

**Connecticut SDE**

**Preparing Students to be College and Career Ready**

**Common Core State Standards:** Importance of focusing time, energy and resources on implementing the CCSS beginning this school year

- Teachers must read the standards
  - Mathematics
  - ELA/Literacy in History/Social Studies, Science and Technical subjects
- Transition Guidelines
- CCSS Resources on CSDE Web site and other states’ Web sites

**Bureau of Student Assessment**
### Connecticut SDE

#### Transition Plan for Mathematics

(See document for full plan)

<table>
<thead>
<tr>
<th>Grade 8 Domains</th>
<th>Grade 8 Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011-2012</td>
<td>• Expressions and Equations</td>
</tr>
<tr>
<td></td>
<td>• Unit 1: Real Numbers</td>
</tr>
<tr>
<td></td>
<td>• Unit 3: Linear Relationships</td>
</tr>
<tr>
<td></td>
<td>• Unit 4: Systems of Linear Relationships</td>
</tr>
<tr>
<td>2012-2013</td>
<td>• Expressions and Equations</td>
</tr>
<tr>
<td></td>
<td>• Functions</td>
</tr>
<tr>
<td></td>
<td>• The Number System</td>
</tr>
<tr>
<td></td>
<td>• Unit 1: Real Numbers</td>
</tr>
<tr>
<td></td>
<td>• Unit 3: Linear Relationships</td>
</tr>
<tr>
<td></td>
<td>• Unit 4: Systems of Linear Relationships</td>
</tr>
<tr>
<td>2013-2014</td>
<td>• Expressions and Equations</td>
</tr>
<tr>
<td></td>
<td>• Functions</td>
</tr>
<tr>
<td></td>
<td>• The Number System</td>
</tr>
<tr>
<td></td>
<td>• Geometry</td>
</tr>
<tr>
<td></td>
<td>• Unit 1: Real Numbers</td>
</tr>
<tr>
<td></td>
<td>• Unit 2: Pythagorean Theorem</td>
</tr>
<tr>
<td></td>
<td>• Unit 3: Linear Relationships</td>
</tr>
<tr>
<td></td>
<td>• Unit 4: Systems of Linear Relationships</td>
</tr>
<tr>
<td></td>
<td>• Unit 5: Congruence and Similarity</td>
</tr>
<tr>
<td></td>
<td>• Unit 6: Volume</td>
</tr>
</tbody>
</table>

Displaced Grade-Level Concepts (Former CT Grade 8 content that is no longer in Grade 8 under the CCSS):
- Recursive and explicit formulas
- Equivalent forms of fractions, mixed numbers, decimals and percent
- Computation with numbers and operations (Should be done in the context of Expressions and Equations in Grade 8)
- Percent
- Exponential growth and decay
- Surface area
- Data representations (limited to scatterplots in Grade 8 CCSS)
- Sampling for statistical analyses
- Permutations and combinations

### Connecticut SDE

#### Understanding a Standard: A Standard is Part of a Learning Progression

<table>
<thead>
<tr>
<th>Grade 5</th>
<th>Grade 6</th>
<th>Grade 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.NBT.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluently multiply multi-digit whole numbers using the standard algorithm.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.NBT.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find whole-number quotients of whole numbers with up to four-digit dividends and two-digit divisors, using strategies based on place value, the properties of operations, and/or the relationship between multiplication and division.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.NBT.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add, subtract, multiply, and divide decimals to hundredths, using concrete models or drawings and strategies based on place value, properties of operations, and/or the relationship between addition and subtraction; relate the strategy to a written method and explain the reasoning used.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.NS.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluently divide multi-digit numbers using the standard algorithm.</td>
<td></td>
<td>All standards in Grade 7 in the Number System domain examine rational numbers. The Grade 6 standard should therefore bring closure to the study of division of positive integers (although its applications will extend into other domains and grades).</td>
</tr>
</tbody>
</table>
Connecticut SDE
ELA/Literacy 6 Shifts in Instruction
1. Balance of literature and literary non-fiction (K-5)
2. Literacy as part of science and social studies/history; informational text as part of ELA (6-12)
3. Appropriately complex text
4. Questions regarding text are text-dependent
5. Writing to inform or argue using evidence
6. Academic Vocabulary

Connecticut SDE
Mathematics Shifts in Instruction
- FOCUS, FOCUS, FOCUS – Deeper understanding of fewer concepts (Less focus on data, statistics, probability in elementary math, more focus on number, operations and algebraic thinking)
- COHERENCE – One year builds to the next
- FLUENCY – These standards expect speed and accuracy
- DEEP UNDERSTANDING – fewer standards allow for this
- APPLICATION – deep understanding will provide students the ability to apply what they know
## Key Fluencies

<table>
<thead>
<tr>
<th>Grade</th>
<th>Required Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>Add/subtract within 5</td>
</tr>
<tr>
<td>1</td>
<td>Add/subtract within 10</td>
</tr>
<tr>
<td>2</td>
<td>Add/subtract within 20</td>
</tr>
<tr>
<td></td>
<td>Add/subtract within 100 (pencil and paper)</td>
</tr>
<tr>
<td>3</td>
<td>Multiply/divide within 100</td>
</tr>
<tr>
<td>4</td>
<td>Add/subtract within 1,000</td>
</tr>
<tr>
<td>5</td>
<td>Multi-digit multiplication</td>
</tr>
<tr>
<td>6</td>
<td>Multi-digit division</td>
</tr>
<tr>
<td>7</td>
<td>Multi-digit decimal operations</td>
</tr>
<tr>
<td>8</td>
<td>Solve px + q = r, p(x + q) = r</td>
</tr>
<tr>
<td></td>
<td>Solve simple 2×2 systems by inspection</td>
</tr>
</tbody>
</table>

## Priorities in Mathematics

<table>
<thead>
<tr>
<th>Grade</th>
<th>Priorities in Support of Rich Instruction and Expectations of Fluency and Conceptual Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>K–2</td>
<td>Addition and subtraction, measurement using whole number quantities</td>
</tr>
<tr>
<td>3–5</td>
<td>Multiplication and division of whole numbers and fractions</td>
</tr>
<tr>
<td>6</td>
<td>Ratios and proportional reasoning; early expressions and equations</td>
</tr>
<tr>
<td>7</td>
<td>Ratios and proportional reasoning; arithmetic of rational numbers</td>
</tr>
<tr>
<td>8</td>
<td>Linear algebra</td>
</tr>
</tbody>
</table>
### Connecticut SDE

#### Moving Forward

**MATERIALS:**
- ELA reading lists are balanced
- Math materials are focused

**TEACHERS:**
- 80% of questions are text-dependent
- Have a deep understanding of the Mathematics content and Mathematical Practices

**STUDENTS:**
- Demonstrate close encounters with text through writing
- Demonstrate fluency and deep understanding of the Mathematics content and can demonstrate the Mathematical Practices

---

**Technology:** Start building capacity for technology
- SBAC Technology Readiness Tool

---

**Bureau of Student Assessment**
Importance for our Nation’s Children

This is a commitment to ensuring that our students are college and career ready upon graduation from high school

- Where we are now will not get us there
- We need to reduce the number of students who require remedial courses in college
- Results on NAEP, TIMSS, PISA indicate that we need to better prepare our students to succeed in a global economy

Summary

- It is important to have resources that show how the critical pieces work together: standards, curriculum, instruction, assessment.
- We are often guilty of asking a single assessment to have too many goals; therefore, we must define our purpose before creating or using an assessment.
- The floor and the ceiling are important for instruction and assessment – and we are moving in a direction for assessment that acknowledges the need for these parts to be moving – instruction must follow suit.
## Resources

  - Student Assessment Link – SBAC
  - Teaching and Learning Link – CCSS

- **SBAC Web site** - [www.smarterbalanced.org](http://www.smarterbalanced.org)

- **Center for K-12 Assessment and Performance Management at ETS** – Updated guide to the assessment consortia
  - [http://www.k12center.org/publications.html](http://www.k12center.org/publications.html)

### Connecticut SDE Resources

<table>
<thead>
<tr>
<th>Formative Assessment - <em>Attributes of Effective Formative Assessment</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.dpi.state.nc.us/docs/accountability/educators/fastatributes04081.pdf">http://www.dpi.state.nc.us/docs/accountability/educators/fastatributes04081.pdf</a></td>
</tr>
</tbody>
</table>

- **Computer Adaptive Testing**
  - CAT Central [http://www.psych.umn.edu/psylabs/catcentral](http://www.psych.umn.edu/psylabs/catcentral)
  - Oregon Assessment of Knowledge and Skills [http://www.oaks.k12.or.us/](http://www.oaks.k12.or.us/)

- **Examples of Innovative Item Types**
  - [http://pages.uoregon.edu/kscalise/taxonomy/taxonomy.html](http://pages.uoregon.edu/kscalise/taxonomy/taxonomy.html)
Connecticut SDE

Contact Information

Shelbi K. Cole Ph.D.
shelbi.cole@ct.gov
860.713.6878

Gail Pagano
gail.pagano@ct.gov
860.713.6821

Bureau of Student Assessment

Questions?

Bureau of Student Assessment

August 15 and 16, 2011