Assessment Systems Supporting Accountability and Learning: Examples from the United States

Joan L. Herman

British Columbia Assessment Forum

October 4, 2013
Vancouver, British Columbia
Overview

• Accountability orientation: Historical context
• Recent developments: Why and what
• System orientation: Multiple users, multiple purposes
• Evidence centered design and validation for deeper learning
• Lessons learned (again!)
Historical Context: Accountability Testing
Performance-Driven Educational Policy

• 1989 Governors’ Summit: Goals 2000

• 1994 IASA required states to establish standards in core academic subjects and assess competency at each school level

• 2001 NCLB Act required states to do:
  
  ✓ Annual testing grades 3-8 and one grade high school
  
  ✓ Accountability for meeting annual performance goals for student proficiency, on target to all students proficient in 2014
  
  ✓ Transparency/reporting to parents and public
Student Outcomes

Instruction + Assessment

On Track to All Students Proficient

Standards/Learning Goals

Classroom Instruction

Individual State Defined
School Report Cards:

• Interest in school performance and accountability
• Increasing availability of data and data management tools
• Parents’ right to know/ school choice
• By 2001: federal legislation requires states to report on district and school progress in achieving No Child Left Behind goals:
  • Academic proficiency for all students, every subgroup
  • Student attendance, graduation and drop out rates
  • Teacher qualifications
  • AYP status
  • Status on state accountability index
  • Other data
State Accountability Indices

- Consider status and progress
- Consider performance relative to demographically similar schools
- Highlight equity goals
Effects of NCLB: Mixed Bag

- Curriculum narrowing
- Educator motivation/school climate
- Pressure → gaming, cheating
- Watering down of standards in some states
- Performance gains -- some
- System alignment, Opportunity-to-learn
- Attention to equity
Standards-Based Reform: The Reality

- Standards/Learning Goals
- Classroom Instruction
- Student Outcomes Assessment

- Ill defined goal
- Instruction + Assessment
- Student proficiency classification

National Center for Research on Evaluation, Standards, & Student Testing
Current United States Developments
Concerns

• Global competitiveness

• College and career readiness (CCR)

• Preparation for 21st century success (vs. what’s tested)

• Disparity in states’ standards and assessments
Solution 1: More Rigorous Standards

• State, not nationally, developed

• Multiply benchmarked
  ✓ Leading countries on international assessments
  ✓ Backward designed from competencies thought essential for college success

• Incorporate content and deeper learning, 21st century competence

• Vertically aligned, progression based
Solution 2: Standards-Based, Test Driven Reform
If WYTIWYG, then Testing is Both Problem and Solution

- **Needed:**
  - ✓ *Tests that send the right signal: Tests worth teaching to*
  - ✓ *Tests that incorporate 21st century competency*

- **Solution**
  - ✓ *$400 million federal investment*
  - ✓ *Two consortia of STATES: PARCC and Smarter Balanced*
Goal: Coordinated Systems of Assessment *Of* and *For* Learning
### RTT Consortia Expectations

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Assessment Type</th>
<th>Primary Users</th>
<th>Use</th>
</tr>
</thead>
</table>
| Accountability/Evaluation| On-demand annual                        | State, District, Schools, Teachers, Parents, Students, Public | ▪ School/Teacher/Principal effectiveness and capacity building  
  ▪ Status/Growth re: college career readiness  
  ▪ General Feedback curriculum strengths and weaknesses, student strengths and weaknesses?  
  ▪ Recognize and build on excellence |
| Accountability/Evaluation| Performance assessment Curriculum embedded | Same as above                                      | ▪ Same as above                                                      |
| Monitoring/Supervision   | Interim Assessment, Curriculum embedded | District /School Administrators, Teachers, Parents, Students | ▪ Identify struggling students, teachers schools  
  ▪ Diagnose/adjust?  
  ▪ Identifying promising practices |
| Formative                | Curriculum-embedded, student work, class discourse, discussion | Teachers, Students | ▪ Inform immediate teaching and learning |
Coherent Systems

END OF YEAR ASSESSMENT
ON TRACK TO COLLEGE READINESS
Current Status of Consortia
Summative Tests: Are They an Improvement?
How Do We Know?

- Evidence-Centered Design (ECD)
- Depth of Knowledge (DOK) expectations
- Focus on summative assessment
  - *Technology enhanced end of year test*
  - *Performance assessment*
One Way to Look for College and Career Readiness (CCR): Norman Webb’s Depth of Knowledge Framework

- DOK1: Recall
- DOK2: Simple Application (some mental processing, 2 steps)
- DOK3: Reasoning, inference
- DOK4: Extended planning and investigation
Example: DOK1 Math

Look at the length of nails A and B.

A

B

inches

How much longer is nail A than nail B?

○ A. $\frac{1}{2}$ inch

○ B. $1 \frac{1}{2}$ inches

○ C. $3 \frac{1}{2}$ inches
Example: DOK2 Math

These cards are placed in a bag.

- 7 + 8
- 6 + 7
- 5 + 8
- 6 + 9
- 8 + 8
- 9 + 5

What is the probability Lauren will pick a card with a sum greater than 15?

A \(\frac{1}{6}\)  
B \(\frac{1}{5}\)  
C \(\frac{3}{6}\)  
D \(\frac{2}{4}\)
DOK Example: Grade 8 Math

Kayla asked 10 students in her class whether they owned a dog or a cat or both.

Drag one number into each box to complete the table, given this information:

- 40% of the students own a dog.
- 30% of the students own a cat.
- 10% of the students own both a dog and a cat.
Example: DOK4 Math

- Your class is going on a field trip. There are three possible choices for the field trip: an aquarium, a science museum, or a zoo. Your teacher asked students to write down their first and second choices. In this task, you will determine where the class should go on the field trip based on the survey results and the cost per student.

  ✓ This is a map of your school and the three different field trip locations.
  ✓ Here’s how students voted, first and second choice
  ✓ Here are costs in time and transportation costs
Students carefully consider two literary texts worthy of close study.

They are asked to answer several evidence-based reading and technology enhanced questions about each text to demonstrate their ability to do close analytic reading and to compare and synthesize ideas.

Students write a literary analysis about the two texts.
Daedalus and Icarus Performance Assessment

• Use what you have learned from reading “Daedalus and Icarus” by Ovid and “To a Friend Whose Work Has Come to Triumph” by Anne Sexton to write an essay that provides an analysis of how Sexton transforms Daedalus and Icarus.

• As a starting point, you may want to consider what is emphasized, absent, or different in the two texts, but feel free to develop your own focus for analysis.

• Develop your essay by providing textual evidence from both texts. Be sure to follow the conventions of standard English.
How are Current State Tests Doing in Addressing Higher Levels of DOK*?

• Mathematics: Great preponderance at DOK1 and DOK2, even with constructed response items

• Reading: Great preponderance DOK1 and DOK2

  ✓ Even for constructed response items, only about a third at DOK3 and less than 10% at DOK4

• This is NOT sufficient to support college and career readiness

* Sources: Webb et al., 2013; Yuan & Le, 2012
Evidence Centered Design

CCSS reorganized as
 CLAIMS AND ASSESSMENT TARGETS (MAJOR AND MINOR) which are the basis for developing
 ITEM AND TASK SPECIFICATIONS FOR GENERATING EVIDENCE AT SPECIFIED DOK OR COGNITIVE COMPLEXITY LEVELS FOR EACH ASSESSMENT TARGET which are the foundation for
 ITEM AND TASK DEVELOPMENT which are subjected to
 END-OF-YEAR TESTS TO MEASURE STUDENT ACCOMPLISHMENT OF CCSS AND ON-TRACKNESS TO CR
 FIELD TESTING AND VALIDITY STUDIES to establish/refine psychometric quality and validity of assessments
 REVIEW AND PILOT TESTING to establish/refine item quality
 TEST BLUEPRINTS which are used to allocate items and tasks to test forms
Do PARCC and Smarter Balanced ELA Claims Represent Deeper Learning?

• Reading: Students can independently read and closely analyze a range of increasingly complex texts.

• Writing: Students can produce well grounded and effective writing for a variety of purposes and audiences.

• Research: Students build and present knowledge through research and the integration, comparison, and synthesis of ideas.

• Speaking and listening: Students employs effective speaking and listening skills for a variety of audiences and purposes.
**DOK in Specs and Blueprints**

<table>
<thead>
<tr>
<th></th>
<th>ELA</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOK1</td>
<td>25%</td>
<td>24%</td>
</tr>
<tr>
<td>DOK2</td>
<td>38%</td>
<td>40%</td>
</tr>
<tr>
<td>DOK3</td>
<td>26%</td>
<td>25%</td>
</tr>
<tr>
<td>DOK4</td>
<td>11%</td>
<td>11%</td>
</tr>
</tbody>
</table>

*Example based on Smarter Balanced Specifications - Elementary Grades*
Lessons Learned
Lessons Learned: Cautions

• Accountability leverage only as good as underlying measures

• Assessment rather than intended standards/goals will take precedence

• Campbell’s law:

The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor
Expanded Validity Criteria*

- Assessment of higher order applications
- High fidelity assessment of critical CCR skills
- Instructional sensitivity and educational value
- Validity, reliability and fairness

*See Darling-Hammond et al., 2013
Context Matters: Design Constraints

• Commitment to individual student scores
• Emphasis on reliability vs. validity
• Paradox of trust in teachers
• Stakeholder priorities
• Cost and feasibility
Wisdom from the past

Don’t let the perfect be the enemy of the good

Voltaire
Balance is the Key to Life
National Center for Research on Evaluation, Standards, & Student Testing

UCLA | Graduate School of Education & Information Studies

VISIT US ON THE WEB
cresst.org

herman@cse.ucla.edu