Next-Generation Assessments: Speed or Substance?

Lorrie A. Shepard
University of Colorado at Boulder
Lorrie.Shepard@Colorado.edu

Washington Educational Research Association
March 24, 2010
Overview

- Negative effects of high-stakes testing on student learning and motivation
- Reforming assessments with conceptually rich tasks
- Coherence between large-scale & classroom assessments
- Cognitive science and learning theory underpinning effective formative assessment
- Learning progressions and curriculum embedded assessments versus multiple-choice-only interim assessments
To be effective in furthering student learning, assessment must be exemplary in two fundamentally important ways.

➤ First, its form and content must fully represent important thinking and competencies in each of the disciplines. (authentic assessment)

➤ Second, classroom assessment processes and purposes must support learning and a learning orientation. (formative assessment)
Negative Effects of Teaching the Test

Research on test-driven reforms based on standardized tests has consistently documented:

- test score inflation and
- curriculum distortion

“Although embarrassing to some State policymakers, the Lake Wobegon report illustrated the potential mischief caused by high-stakes testing: higher test scores without more learning.”

*Testing in American Schools, 1992*
Assessment reformers used the terms “Performance,” (Resnick & Resnick, 1992) “Direct,” (Frederiksen & Collins, 1989) and “Authentic” Assessment (Wiggins, 1989) to convey the idea that assessments must capture real learning if they are to avoid distorting instruction.
A campground has a large lawn with a soccer field that measures $100 \times 50$ meters (Figure 1). The park manager decides to keep the field open at night. Therefore, a decision needs to be made about where to place some light posts. Standard lamp posts are 13 meters high and light a circular region with a radius of 50 meters (see Figure 2).

1. The diagram below (Figure 3) shows the lighting of the field when lights are placed at points D and B. What is the area of the soccer field that is **NOT** lit when these two light posts are used. Show your work.

```
Dutch examination item by the courtesy of the Freudenthal Institute.
```
Coherence between classroom & large-scale assessments

- Requires that both share the same underlying model of learning,
- i.e., both must share the same conception of developing competence in a domain.
Current Learning Theory Brings Together Cognitive and Social Development

- Children develop their abilities to think and reason in the same way that they learn language, gestures, interpersonal behaviors, manners, and tastes through their social interactions with family and community.

- Mathematical discourse is a critical feature of math education reforms, for example, because it creates the opportunity for students to develop problem-solving and thinking abilities.
Formative Assessment // Instructional Scaffolding

Sadler’s formative assessment model

- Where are you trying to go?
- Where are you now?
- How can you get there?

A shared understanding and ownership of the learning goal, support in reaching the goal, internalization, and taking over of responsibility by the learner are closely parallel features of formative assessment and scaffolding.
As Processes, Good Formative Assessment Practices

- Elicit prior knowledge, both prerequisite and background knowledge, and develop classroom routines to help students become aware of knowledge use.

- Provide feedback that enables students to see how to improve performance over time.

- Develop classroom expectations that call for application and generalization, rather than repetitive practice.

- Use self-assessment as a means to support internalization of criteria and personal ownership of the learning process.
To Be Effective, Formative Student Assessments Must

- “Embody” learning goals.
- Be curriculum-embedded (both in timing and substance). Tasks are instructional tasks so no instructional time is lost; occurs “midstream”* to inform instruction not as a unit summative test.
- Enable the supportive learning processes invoked in the formative assessment literature.
- Instructional insights should lead to coherent, theoretically-sound improvements in teaching.

*see Stuart Kahl, Ed Week, 9.21.05
Beyond Learning Theory and Formative Assessment Principles…

Strategic Help:

Curriculum-embedded assessment tools support teachers learning about formative assessment by providing:

- Better tasks
- Typical misconceptions
- Strategies for eliciting student thinking
- Learning progressions are the Cadillac of curriculum-embedded assessment tools.
Assessments Can/Should Be Coherently Linked to Curriculum by Shared Learning Progressions.

- Learning progressions “describe skills, understandings, and knowledge in the sequence in which they typically develop: a picture of what it means to ‘improve’ in an area of learning.” (Masters & Forster, 1996).

- Learning progressions or Progress Maps provide an underlying model of learning to coherently link classroom and large-scale assessments.

- A criterion-referenced growth model.

See Knowing What Students Know (2001)
BEAR Assessment System
(Wilson & Sloane, 2000)

Best example in the US, illustrates KWSK concept of *coherence*.

- Coherence between classroom and external tests requires a shared understanding of the construct at the level of the progress dimension and at the level of specific assessment tasks and scoring guides. (Wilson & Draney, 2004)

- Progress variables help to align learning goals, instruction, and assessment. Embedded assessments help to compare students’ thinking with curricular expectations. Student performance gains on classroom assessments can be linked to expected gains on an external large-scale assessment without “teaching to the test.” (Kennedy, Brown, Draney, & Wilson)
Beyond Assessments that Mirror Instruction, Learning Progressions Answer Instructional Questions (Qualitative as well as Quantitative)

- An understanding of learning progressions or learning trajectories is important for monitoring and supporting learning and development over time.

- How learning typically unfolds helps teachers know “what next” and how to “back-up” (though we must also be aware of natural variations and departures from the typical pattern).
North Carolina Children’s Writing Continuum
## Strategies in Children’s Spelling at Different Stages

<table>
<thead>
<tr>
<th>Prominent Strategy</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prephonemic</td>
<td>&quot;C&quot; for &quot;hat&quot;</td>
</tr>
<tr>
<td>Early phonemic</td>
<td>&quot;DR&quot; for &quot;Dear&quot;</td>
</tr>
<tr>
<td>Phonetic</td>
<td>&quot;wns&quot; for &quot;once&quot;</td>
</tr>
<tr>
<td>Simple Associations</td>
<td>&quot;bid&quot; for &quot;bird&quot;</td>
</tr>
<tr>
<td>Strategic Extensions</td>
<td>&quot;bote&quot; for &quot;boat&quot;</td>
</tr>
<tr>
<td>Conventional</td>
<td></td>
</tr>
</tbody>
</table>

Hiebert & Raphael (1998)
The instructional sequence of learning trajectories is composed of key tasks designed to promote learning at a particular conceptual level or benchmark in the developmental progression. What tasks will prompt students to construct the concepts and skills that characterize the succeeding level? (Clements & Samara, 2004)

Learning Trajectories are different from previous efforts to define learning sequences or learning hierarchies. They focus on meaningful learning, are guided by constructivist learning theory, and are based on instructional design experiments. (Baroody, Cibulskis, Lai, & Li)
“This demonstrates natural selection because after the draught it shows birds beaks depth getting larger. Since the environment is changing the birds have to adapt to their new surroundings.”

What happens if you cut the nail in half?

Otero, Jalovec, & HerManyHorses. (2008).
Cautions

- Beware of dishonest or misleading learning progressions. Disordered item banks, scaled by difficulty, do not provide adequate learning progressions.

- Today, most benchmark tests identify which kids need more attention and they provide subtest scores intended to be “diagnostic.” They do not provide instructional insights about what each student understands or does not understand.
Normative Models of Growth
Cautions

- Unlike Australia, England, New Zealand, The Netherlands, Norway, Scotland, Switzerland, and Wales where more instructionally ground examples of formative assessment are being implemented, the U.S. context is hampered by lack of substantively rich curricula adopted at the same level as accountability tests and data systems.
Substantive insights are rare.

- **Blanc et al.**:
  “Instructional Communities”
  In one school, principal was clearly the moral and organizational star.
  Coherent emphasis on core curriculum helped to link Benchmark results to instructional strengths and weaknesses.

- **Olah et al.**:
  “Analyzing Benchmark Data”
  Two of 25 teachers connected item and standards data aback to Everyday Mathematics curriculum.
  Diagnostic insights from items, e.g. problems with regrouping, were rare.
Item-by-item teaching

- Bulkley et al.: “Role of the District”
  Assumed that teachers would be able to figure out what to do by looking at the data, talking with other teachers, and drawing inferences about their own needs for professional development.

- Blanc et al.,
  High-stakes state testing talk pervaded grade-level meetings, focus on “bubble kids,” reteach usually with the same instructional strategy, alternative strategies not informed by assessment results.

- Olah et al.,
  Teachers invented their own thresholds to determine priorities for reteaching. Analyses of items were most often procedural or about item validity. Reteaching was focused procedurally on reteaching items missed, step by step.
Substantively, Good Formative Assessment Tools

- Can never be *all* multiple-choice.
- Should provide *qualitative* insights about understandings and misconceptions not just a numeric score.
- Should have immediate implications for what to do besides re-teaching every missed item. (the 1000 mini-lessons problem).
Thoughts for the future

- States and districts with curricular authority could invest in development of curriculum embedded assessments and accompanying learning progressions.

- Districts could invest in professional development for teachers emphasizing the use of formative assessment specifically tied to the district’s curriculum (examining student work, identifying typical misconceptions, developing learning progressions, sharing related instructional practices, etc.).
Silicon Valley Mathematics Initiative (SVMI)

- Continuous loop of focusing on standards, assessing student work, analyzing students’ understandings from assessments, developing effective educational strategies based on findings, and tailoring instruction to enhance student learning.

- Districts administer the MARS summative performance assessment to measure students’ problem solving and ability to explain and justify solutions. Results are used to improve instructional program. MARS-like tasks are used as part of PD throughout the year.
In order for formative assessment not to be just another hijacked reform – implementation efforts must attend to both the substance of what is assessed and the instructional processes that support learning. Getting scores back quickly should not be how we think about instructionally relevant assessment.