Student Assessment

Session 2

Why student assessment is important

February 6, 2013
Room I2-220
Washington, DC
A theory of change

Standards

Assessment

Accountability

Clear expectations for students and schools
Motivation to work hard

Professional development
Improved teaching

Higher levels of learning
Session overview

• Key evidence and good practices for student assessment
• Different assessment models
• What matters most for student assessment systems?
  – Introduction to SABER-Student Assessment framework
Definitions, key evidence, good practices
What is student assessment?

...the process of gathering and evaluating information on what students know, understand, and can do in order to make an informed decision about next steps in the educational process.
What is a student assessment system?

... a group of policies, structures, practices, and tools for generating and using information on student learning and achievement.
What counts as key evidence and best practices for effective assessment?
Classroom assessment

- Strong link between high-quality assessment and better learning outcomes

- Gains of half to full standard deviation on standardized tests (Black & Wiliam, 1998; Rodriguez, 2004)

- More work needed to define and isolate characteristics that lead to improved outcomes (Bennett, 2011)
Examinations

Link between exit exam policies and higher performance levels on international assessments (Bishop, Mane, & Bishop, 2001)

Link between specific characteristics of the test and student learning outcomes (Au, 2007; Hill, 2010)

Can have negative impact on students (Greaney & Kellaghan, 1995; Madaus & Clarke, 2001)
Overall weak, positive link between accountability uses of data and better learning outcomes (Carnoy & Loeb, 2002)

Simply reporting information about average school scores can lead to increased student performance (Hanushek & Raymond, 2003)

Much to learn about optimal design for accountability models that will produce best outcomes with fewest negative side effects
Different countries... different assessment models
Finland

• Strong emphasis on classroom-based formative assessment activities
• Limited use of high-stakes examinations
• Regular schedule of low-stakes, large-scale assessment activities for system monitoring
China

• Strong emphasis on high-stakes examinations
• Emergent focus on classroom assessment and large-scale assessment
• Very strong emphasis on large-scale assessment
• Strong emphasis on classroom assessment and examinations
<Name of country in which you work>?

- Classroom assessment
- Examinations
- Large-scale assessment
Activity: Choosing among models

• Identify your preferred assessment model and discuss the reasons behind your preference.
  – Which is your preferred model?
    1. More emphasis on classroom assessment
    2. More emphasis on external examinations
    3. More emphasis on large-scale assessments
  – Why?
  – Under what conditions would you recommend it to a client?
What matters most for student assessment systems?
Assessment types/purposes

• **Classroom assessment**
  • For improving teaching and learning

• **Examinations**
  • For making high-stakes decisions about individual students

• **Large-scale assessment**
  • For determining system learning levels and related factors
Quality drivers

• **Enabling context**
  - Leadership, policies, institutional arrangements, human/fiscal resources

• **System alignment**
  - Learning goals, curricula, opportunities

• **Assessment quality**
  - Design, administration, analysis
  - Reporting and use
Framework

1. Indicators

2. Development Levels

Classroom assessment

Examinations

Large-scale assessment

Enabling context

System alignment

Assessment quality
Indicators

• Enabling context
  – Policies, leadership, institutional arrangements, budget, human resources

• System alignment
  – Learning standards and goals, curricula

• Assessment quality
  – Technical features, dissemination, use, quality assurance
Development levels

- Latent
- Emerging
- Established
- Advanced
Activity: Applying the framework to your regional context

• Using the framework as an organizing structure, discuss the relative strengths and needs of countries in the region where you work in the area of assessment.
• What data did you draw on to come to your conclusions?
• What other data would you need before taking any action?
Indicators

• Enabling context
  – Policies, leadership, institutional arrangements, budget, human resources

• System alignment
  – Learning standards and goals, curricula

• Assessment quality
  – Technical features, dissemination, use, quality assurance
Extra slides
Classroom Assessment Development Levels

- Angola
- Mozambique
- Zambia
- Bangladesh
- Nepal
- Pakistan
- Sri Lanka
- Serbia

- Ethiopia
- Kyrgyz Republic

- Weak system-wide institutional capacity.
- Some/sufficient system-wide institutional capacity.
- Strong system-wide institutional capacity to support and ensure quality of classroom assessment practices.

Levels:
- Latent
- Emerging
- Established
- Advanced
Examinations
Development Levels

No standardized exam.

Partially stable standardized exam, typically of poor quality.

Stable standardized exam of acceptable quality; limited monitoring mechanisms.

Stable standardized exam of high quality; institutional capacity and strong monitoring mechanisms.

Latent
Emerging
Established
Advanced

Ethiopia
Zambia
Bangladesh
Nepal
Sri Lanka
Armenia
Korea

Angola
Mozambique
Pakistan
Kazakhstan
Kyrgyz Republic
Serbia

Finland
National Large-Scale Assessment Development Levels

- Ethiopia
- Mozambique
- Zambia
- Bangladesh
- Nepal
- Pakistan
- Sri Lanka
- Armenia
- Kazakhstan
- Kyrgyz Republic
- Serbia
- Angola
- Finland
- Korea

Latent
Emerging
Established
Advanced

- No NLSA in place.
- Unstable NLSA in place; assessment quality and impact weak.
- Stable NLSA of moderate quality in place; information disseminated, but not used effectively.
- Stable NLSA of high quality in place; information effectively used.
International Large-Scale Assessment Development Levels

No history of participation in ILSA.

More or less stable participation in ILSA.

Stable participation in ILSA. Information effectively used to improve education.

Angola
Ethiopia
Bangladesh
Nepal
Pakistan
Sri Lanka
Mozambique
Zambia
Kyrgyz Republic
Armenia
Kazakhstan
Serbia
Korea

Latent
Emerging
Established
Advanced
### Sub-Saharan Africa

<table>
<thead>
<tr>
<th></th>
<th>Classroom Assessment</th>
<th>Examinations</th>
<th>Large-scale Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>National</td>
<td>International</td>
<td></td>
</tr>
<tr>
<td>Angola</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# South Asia

<table>
<thead>
<tr>
<th>Country</th>
<th>Classroom Assessment</th>
<th>Examinations</th>
<th>Large-scale Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nepal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pakistan (Punjab &amp; Sindh)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sri Lanka</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Eastern Europe, Central Asia

<table>
<thead>
<tr>
<th></th>
<th>Classroom Assessment</th>
<th>Examinations</th>
<th>Large-scale Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>National</td>
<td>International</td>
<td></td>
</tr>
<tr>
<td>Armenia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kazakhstan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serbia</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Key trends and issues

• Blended cognitive and psychometric models
• Blurring of lines between classroom and external assessment events
• Measuring 21st century skills, cognitive, non-cognitive
• Key role of technology
• Key role of teachers
• Accountability uses of assessment data
  – Value-added models
  – NGOs and community mobilization
21st century skills

WAYS OF THINKING
- Creativity and innovation
- Critical thinking, problem solving, decision making
- Learning to learn

TOOLS FOR WORKING
- Information literacy
- ICT literacy

WAYS OF WORKING
- Communication
- Collaboration

WAYS OF LIVING IN THE WORLD
- Citizenship
- Life and career
- Personal and social responsibility

EDUCATION STAFF DEVELOPMENT PROGRAM
Sample item

How do different payload masses affect the altitude of a helium balloon?

Design Experiment

Run Experiment

Interpret results

<table>
<thead>
<tr>
<th>Payload Mass (pounds)</th>
<th>Amount of Helium (cubic feet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2275</td>
</tr>
</tbody>
</table>