What Matters Most for Student Assessment Systems: A Framework

Marguerite Clarke
Definitions, theories, key evidence
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.
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
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)
Large-scale ‘survey’ assessment

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

Overall weak, positive link between accountability uses of data and better learning outcomes (Carnoy & Loeb, 2002)

Much to learn about optimal design for accountability models that will produce best outcomes with fewest negative side effects
Mexico: Programa de Atención Específica para la Mejora del Logro Educativo (PAE)

Average School Performance during 2006-2012, Colima

Before the PAE

After the PAE

Average School Performance

Spring term of the school year

2006 2007 2008 2009 2010 2011 2012

PAE Schools
Non-PAE Schools
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
USA

- Very strong emphasis on large-scale assessment
- Less strong emphasis on classroom assessment and examinations
What matters most for student assessment systems?
Key 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

<table>
<thead>
<tr>
<th>Enabling context</th>
<th>Classroom assessment</th>
<th>Examinations</th>
<th>Large-scale assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>System alignment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment quality</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Questionnaires and Rubrics**
Applying the framework to countries
Key trends and issues
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
Gracias
Thank You

For more information and to download materials, please visit:

www.worldbank.org/education/saber

www.worldbank.org/readtf