SABER - Student Assessment

Benchmarking Results for READ Countries

2009 Baseline

Marguerite Clarke
The World Bank
Overview

• Introduction to benchmarking exercise
• Summary baseline (2009) report for READ countries
  – Classroom assessment
  – Examinations
  – National large-scale assessment
  – International large-scale assessment
• Next steps
Overview

• **Introduction to benchmarking exercise**

• Summary baseline (2009) report for READ countries
  – Classroom assessment
  – Examinations
  – National large-scale assessment
  – International large-scale assessment

• Next steps
Purpose

What?
• SABER - Student Assessment
  – Evidence-based tool for benchmarking student assessment policies and systems around the world

Why?
• Apply to READ countries to determine where their assessment systems are strong, and where they are in need of improvement
• Create baseline (2009) for tracking progress under READ
# Benchmarking 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>

- 1. Indicators
- 2. Development Levels
Indicator Data Collection

Survey of Student Assessment Systems
National Large-Scale Assessment

2011 QUESTIONNAIRE

1. Please provide the name of the country or education system.

2. Please provide the data of data collection.

3. SYSTEM ASSESSMENT AND BENCHMARKING FOR EDUCATION RESULTS

   The World Bank
   Human Development Network

Please answer all remaining questions with respect to the assessment indicated in question 4.

5. Does the country/system have a policy that authorizes the large-scale assessment program?
   a. ( ) Yes, a formal policy
   b. ( ) No, informal or draft policy
   c. ( ) No → Go to question 6

Comments:

6. Please provide the following information on the policy that authorizes the large-scale assessment program:

   a. [Official document citation: ____________________________]
   b. [Authoring body: ________________________________]
   c. [Year of authorization: ______]

Please provide a link or attach a copy of the policy with your submission of the completed questionnaire.

Comments:

7. Is the policy identified in question 6 available to the public?
   a. ( ) Yes
   b. ( ) No

Comments:

8. Does the country/system’s government have a large-scale assessment plan for the coming year or future assessment rounds?
   a. ( ) Yes
   b. ( ) No → Go to question 10

Please provide the link or attach a copy of the plan with your submission of the completed questionnaire.

Comments:
<table>
<thead>
<tr>
<th>LATENT</th>
<th>EMERGING</th>
<th>ESTABLISHED</th>
<th>CUTTING-EDGE</th>
<th>JUSTIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence of, or deviation from, the attribute</td>
<td>On way to meeting minimum standard</td>
<td>Acceptable minimum standard</td>
<td>Best practice</td>
<td></td>
</tr>
</tbody>
</table>

### KYRGYZ REPUBLIC

**National Large-Scale Assessment (NLSA)**

**Rubric**

<table>
<thead>
<tr>
<th>Enabling Context (EC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall framework of policies, leadership, organizational structures, fiscal and human resources in which NLSA activity takes place in a country or system and the extent to which that framework is directly conducive to, or supportive of, the NLSA activity.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EC1 -- Setting clear policies for NLSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Q3.1) No NLSA exercise has taken place.</td>
</tr>
<tr>
<td>(Q3.2) The NLSA has been operating on an irregular basis.</td>
</tr>
<tr>
<td>(Q3.3) The NLSA is a stable program that has been operating regularly.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Kyrgyz Republic has been conducting NLSAs on an irregular basis and those that were conducted were driven by donor agencies (MLA with UNICEF funding, and two rounds of NSBA with WB funding).</td>
</tr>
</tbody>
</table>
Overall Score, Development Level for Each Assessment Type

1. Latent
   - 1.5
   - 1.75 ± 0.25

2. Emerging
   - 2.0
   - 2.50 ± 0.25

3. Established
   - 2.25
   - 2.75
   - 3.0
   - 3.0 ± 0.25

4. Cutting Edge
   - 3.5
   - 3.25 ± 0.25

Cutoff points with fixed intervals
Margin of discretion around score to assign development level
Validated Country Summary Report

<table>
<thead>
<tr>
<th>Validated Country Summary Report</th>
</tr>
</thead>
</table>

**KYRGYZ REPUBLIC 2009 Baseline**

**Level of Development**

<table>
<thead>
<tr>
<th>ASSESSMENT TYPE</th>
<th>DESCRIPTION</th>
<th>LEVEL OF DEVELOPMENT</th>
<th>POSSIBLE NEXT STEPS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLASSROOM ASSESSMENT</strong></td>
<td><em>There is a recognition among key stakeholders of the importance of classroom assessment and the urgent need to strengthen existing practices (which are very weak) in order to better support student learning.</em></td>
<td>LATE-NIGHT</td>
<td><em>Making available system-wide resources for teachers to engage in better-quality instruction, assessment, and moderation.</em></td>
</tr>
<tr>
<td>EXAMINATIONS</td>
<td><em>The examination system has been operating on a regular basis under the provisions approved by the government of the Kyrgyz Republic.</em></td>
<td>EMERGING</td>
<td><em>Challenging the credibility of the new examination system by making quality-assurance a high-priority technical report on their design and conduct, and introducing systematic quality-control mechanisms.</em></td>
</tr>
<tr>
<td><strong>NATIONAL LARGE-SCALE ASSESSMENT (LSA)</strong></td>
<td><em>There is a strong effort to ensure the quality of LSA through various mechanisms, including monitoring as well as training staff who are conducting statistical analysis.</em></td>
<td>EMERGING</td>
<td><em>Introducing regular government funding to cover costs and establishing a permanent LSA office that would be responsible for LSA reporting.</em></td>
</tr>
<tr>
<td><strong>INTERNATIONAL LARGE-SCALE ASSESSMENT (ILSA)</strong></td>
<td><em>The Kyrgyz Republic was invited to join the Progress in International Reading Literacy Study (PIRLS) and, at the time of preparing this report, had not yet participated.</em></td>
<td>EMERGING</td>
<td><em>Building capacity to ensure the quality of future LSA activities through, for example, offering training opportunities and providing funding for attending international workshops or training on international assessments.</em></td>
</tr>
</tbody>
</table>
Overview

• Introduction to benchmarking exercise
• **Summary baseline (2009) report for READ countries**
  – Classroom assessment
  – Examinations
  – National large-scale assessment
  – International large-scale assessment
• Next steps
## READ Baseline (2009) Ratings

<table>
<thead>
<tr>
<th></th>
<th>Classroom</th>
<th>Examinations</th>
<th>NLSA</th>
<th>ILSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tajikistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## READ Baseline (2009) Ratings

### Classroom Assessment

<table>
<thead>
<tr>
<th>Country</th>
<th>Classroom</th>
<th>Examinations</th>
<th>NLSA</th>
<th>ILSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tajikistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Classroom Assessment Development Levels

- Strong system-wide institutional capacity to support and ensure quality of classroom assessment practices.
- Some/sufficient system-wide institutional capacity.
- Weak system-wide institutional capacity.
- No system-wide institutional capacity.

Countries:
- Angola
- Mozambique
- Tajikistan
- Vietnam
- Zambia
- Ethiopia
- Kyrgyz Republic
Classroom Assessment Development Levels

- No system-wide institutional capacity.
  - No formal guidelines at system (or sub-system) level
  - Limited use of classroom assessment information
  - No system-wide resources for teachers, monitoring of practices, training

Ethiopia
Kyrgyz Republic
Classroom Assessment Development Levels

- No system-wide institutional capacity.
- 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.

- Angola
- Mozambique
- Tajikistan
- Vietnam
- Zambia

• System-level guidelines
• Some system-wide resources, training
• Teacher, curriculum quality issues
• Limited oversight, uses
<table>
<thead>
<tr>
<th></th>
<th>Classroom</th>
<th>Examinations</th>
<th>NLSA</th>
<th>ILSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tajikistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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.

Countries:
- Angola
- Kyrgyz Republic
- Mozambique
- Tajikistan
- Vietnam
- Ethiopia
- Zambia
Examinations
Development Levels

Angola
Kyrgyz Republic
Mozambique
Tajikistan
Vietnam

- Partially stable standardized exam, typically of poor quality.

- Limited national policy framework and funding
- Separate school leaving, university entrance exams
- Credibility and security issues
- Inappropriate uses of results
- Limited options for those not doing well
Examinations
Development Levels

Ethiopia
Zambia

Stable standardized exam of acceptable quality; limited monitoring mechanisms.

• National policy framework, funding
• Unified exam aligned with national curriculum
• Technical quality reporting
• Credible
• Options for those not doing well
<table>
<thead>
<tr>
<th>Country</th>
<th>Classroom</th>
<th>Examinations</th>
<th>NLSA</th>
<th>ILSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tajikistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
National Large-Scale Assessment Development Levels

- Angola
- Tajikistan
- No NLSA in place.

- Ethiopia
- Kyrgyz Republic
- Mozambique
- Vietnam
- Zambia
- 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.
National Large-Scale Assessment Development Levels

No NLSA in place.

- Unstable NLSA in place; need to develop institutional capacity; assessment quality and impact weak.
- Stable NLSA of moderate quality in place; institutional capacity and some information dissemination, but not used effectively.
- Stable NLSA of high quality in place; institutional capacity and strong monitoring mechanisms; information used effectively.

Angola
Tajikistan

• No NLSA (some provincial activity), but acknowledgment that it is necessary.
National Large-Scale Assessment Development Levels

Ethiopia
Kyrgyz Republic
Mozambique
Vietnam
Zambia

- Unstable NLSA in place; assessment quality and impact weak.
- NLSA operating with donor support, no policy framework (except Zambia)
- Need for staff training/learning, quality assurance
- Results not well disseminated or used
### READ Baseline (2009) Ratings

#### International Large-Scale Assessment

<table>
<thead>
<tr>
<th></th>
<th>Classroom</th>
<th>Examinations</th>
<th>NLSA</th>
<th>ILSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td></td>
<td></td>
<td></td>
<td>Red</td>
</tr>
<tr>
<td>Ethiopia</td>
<td></td>
<td></td>
<td></td>
<td>Red</td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td></td>
<td></td>
<td></td>
<td>Yellow</td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td></td>
<td></td>
<td>Yellow</td>
</tr>
<tr>
<td>Tajikistan</td>
<td></td>
<td></td>
<td></td>
<td>Red</td>
</tr>
<tr>
<td>Vietnam</td>
<td></td>
<td></td>
<td></td>
<td>Red</td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
<td></td>
<td></td>
<td>Yellow</td>
</tr>
</tbody>
</table>
No history of participation in ILSA.

Participation in ILSA initiated.

More or less stable participation in ILSA.

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

Angola
Ethiopia
Tajikistan
Vietnam

Kyrgyz Republic
Mozambique
Zambia
International Large-Scale Assessment Development Levels

- Angola: No history of participation in an ILSA.
- Ethiopia: Interest in participating in ILSA in the next few years, but no concrete plans.
- Tajikistan: No previous activity.
- Vietnam: Stable participation in ILSA and institutional capacity to run it. Information from ILSA effectively used to improve education.
- Kyrgyz Republic: Participation in an ILSA has been initiated, but still need to develop institutional capacity to run it.
- Mozambique: More or less stable participation in an ILSA. Institutional capacity is there. Information from ILSA is disseminated, but not used effectively.
- Zambia: Stabilized participation in ILSA and institutional capacity to run it. Information from ILSA effectively used to improve education.
International Large-Scale Assessment Development Levels

Kyrgyz Republic
Mozambique
Zambia

Participation in ILSA initiated.

- PISA 2006, 2009
- SACMEQ I, II, III plus plans for future
- Often donor funded
- Limited use of results to inform decision making

No history of participation in an ILSA.

Participation in ILSA initiated.

More or less stable participation in an ILSA.

Institutional capacity to run it. Information from ILSA effectively used to improve education.

Angola
Ethiopia
Tajikistan
Vietnam

• PISA 2006, 2009
• SACMEQ I, II, III plus plans for future
• Often donor funded
• Limited use of results to inform decision making
Overview

• Introduction to benchmarking exercise

• Summary baseline (2009) report for READ countries
  – Classroom assessment
  – Examinations
  – National large-scale assessment
  – International large-scale assessment

• Next steps
### READ Baseline (2009) Ratings

**Country Focus under READ**

<table>
<thead>
<tr>
<th></th>
<th>Classroom</th>
<th>Examinations</th>
<th>NLSA</th>
<th>ILSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angola</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tajikistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusion

• Countries mainly at **Emerging** stage, but variation within and across countries
  – Examinations the most developed assessment type, ILSA the least
  – Countries can learn from one another

• Next steps driven by
  – Benchmarking results
  – Country needs and priorities
Thank you!

SABER – Student Assessment
Benchmarking Results for READ Countries

SABER – Student Assessment Core Team
Marguerite Clarke, mclarke2@worldbank.org
Maria-Jose Ramirez
Julia Liberman

in collaboration with READ Country Teams and TTLs