Why is Student Learning Assessment Important?

Learning matters to growth.

- Solid evidence shows that learning -- the acquisition of knowledge, skills, and competencies -- drives economic returns.
- More than years of schooling, it is learning that produces gains in individual productivity and income. Several studies show a strong correlation between higher test scores and higher wages.
- A one standard deviation increase in scores on international assessments of reading and mathematics achievement is associated with a 2 percent increase in annual growth rates of GDP per capita.

Assessment is fundamental to improving learning.

- Assessment of student learning and achievement can be a powerful agent for change in schools by informing policymakers about the learning outcomes of the nation's students, helping teachers understand how to improve classroom instruction, and influencing societies as they think about education quality and learning goals.
- Effective classroom assessment practices by teachers have been shown to yield gains of a half to a full standard deviation on external standardized achievement tests, with the largest gains for low achievers.
- Assessment can contribute to an effective education system by providing systemic information on levels of student achievement and related factors; supporting educators and students with real-time information to improve teaching and learning; and offering a means to hold stakeholders accountable for results.

Assessment provides good 'bang for the buck'.

• Studies show testing to be among the least expensive innovations in education reform, costing far less than increasing teacher salaries or reducing class size, and having a higher ratio of benefits to costs.

The Global View

Compared to 10 years ago, far fewer children in developing countries are out of school, but more schooling has not resulted in greater learning.

• Youth are leaving school and entering the workforce without the knowledge, skills, or competencies necessary to adapt to a competitive and increasingly globalized economy.

Large numbers of students are not learning.

 Recent studies in Latin America and South Asia found half of Grade 5 children (India) and Grade 2 children (Peru) unable to read, and half of Grade 3 students (Pakistan) unable to do basic multiplication.

Korea and Finland were the highest performing countries in the 2009 OECD Programme for International Student Assessment (PISA).

- In 18 of 65 participating countries (31 of which were non-OECD), the highest level of reading proficiency achieved by most students was the baseline Level 2.
- Countries with the highest overall levels of reading performance (Finland and Korea) also had the lowest variation in student scores; the PISA report (2010) notes that "the best performing school systems manage to provide high-quality education to *all* students".
- Girls outperformed boys in reading skills in every participating country.
- Countries of similar prosperity produced very different results.

In many countries, there is little knowledge of whether learning is taking place.

- There is no systematic measurement of learning in many countries and thus no roadmap for analyzing where problems lie or how to address them.
- Data on student learning and achievement are often sporadic and limited in scope, and rarely cover the poor.

However, more countries are starting to measure the reading and math competencies of their students through national large-scale assessment exercises, and more countries are starting to participate in regional and international large-scale assessment exercises.

- In 2009, 65 countries participated in PISA, compared to 43 in 2000.
- In 2011, 65 countries participated in TIMSS, compared to 45 in 1995.

Countries that measure student learning and achievement can use the information from these assessments to influence education reform.

- A 1990 assessment in Uruguay triggered improved assistance for teachers, schools, and poor students.
- A national assessment exercise in Sri Lanka in 2003 was the basis for the allocation of increased public resources to poor schools.
- Low performance on the 2009 PISA exercise in the United States spurred intense scrutiny of U.S. education policies and practices.

How Can Countries Improve Their Approaches to Assessing Student Learning and Achievement?

Establishing a student assessment <u>system</u> is key.

• Ad hoc assessment can offer a snapshot of learning and a useful wake-up call, but a sustained, systematic approach to assessment is key to monitoring, analyzing, and improving student learning and achievement levels.

SABER-Student Assessment is a World Bank initiative to review and help improve assessment policies and systems in developed and developing countries around the world.

 The goal of SABER-Student Assessment is to promote stronger assessment systems that contribute to improved education quality and learning for all. It is one of the subsystems within the World Bank's Systems Approach for Better Education Results (SABER) program, which is an evidence-based program to help countries systematically examine and strengthen the performance of their education systems. SABER-Student Assessment focuses on three main kinds of activities in assessment systems, corresponding to three main information needs or purposes. These are:

- *classroom assessments* for providing real-time information to support teaching and learning in individual classrooms
- examinations for making decisions about an individual student's progress through the education system (for example, certification or selection decisions), including the allocation of 'scarce' educational opportunities
- *large-scale, system-level assessments* for monitoring and providing policymakerrelevant information on overall performance levels in the system, changes in those levels, and related or contributing factors

SABER–Student Assessment places emphasis on three main drivers of quality in an assessment system. Countries must address these performance drivers in order to strengthen the quality and utility of the information produced by the various assessment activities in their assessment system. The drivers are:

- the enabling context, which refers to the broader context in which an assessment activity takes place and the extent to which that context is conducive to, or supportive of, the assessment (includes the legislative or policy framework; leadership; institutional arrangements; availability of sufficient and stable funding and competent staff)
- system alignment, which refers to the extent to which the assessment is aligned or coherent with other components of the education system (includes the degree of congruence between assessment activities and system learning goals, standards, curriculum)
- assessment quality, which refers to the psychometric quality of the instruments, processes, and procedures used for the assessment activity (includes the design of assessment activities; analysis of student responses to those activities; and the appropriateness of how results are reported and used)

	Classroom assessment	Examinations	Large-scale, system-level assessment
Enabling context			assessment
System alignment			
Assessment quality			

Table 1: Framework for Building a More Effective Student Assessment System

Source: Clarke (2012)

Case studies on countries' experiences in strengthening their student assessment systems reveal that it often takes a decade or more for a set of reforms and inputs to really take hold and produce tangible results.

 Therefore, countries must plan from the outset to have a long-term commitment to, and investment in, the policies, inputs, and actions that will be required to transform their assessment system. The payoff will be an assessment system that can support better decision making and contribute to higher levels of education quality and learning for all.

Recent World Bank Assistance

Between 1998 and 2009, just over half of all World Bank education projects supported student assessment activities, covering 90 countries (166 out of a total of 306 education projects, or 54%).

- AFR accounted for a quarter of these assessment-related projects (41 out of 166 projects), slightly above ECA, LCR, and SAR.
- Projects supporting assessment activities were more prevalent in LICs than MICs (90 versus 76).
- An overwhelming majority of education projects in ECA supported assessment activities.
- Support for assessment was mainly focused at the primary level, except for ECA which focused more on secondary.
- Large-scale assessments of national achievement levels were by far the most common type of assessment activity to be supported, in contrast to a historically (pre 1998) greater emphasis on assistance for examination reform.

Figure 1: Distribution of World Bank Education Projects Supporting Student Assessment Activities, 1998-2009



Source: Liberman & Clarke (2012) **Table 2: Regional Emphases in the World Bank's Support for Student Assessment Activities. 1998-2009**

Region	Education Level	Assessment	Quality driver	Support for participation in		
		type		international assessment		
AFR	Primary	LSA, Ex	EC	None		
EAP	Primary	LSA	EC, SA, AQ	Some		
ECA	Secondary,	LSA, Ex, CA	EC	Substantial		
	Primary					
LCR	Primary	LSA	EC	Some		
MNA	Primary,	CA, LSA	SA	Some		
	Secondary					
SAR	Primary	LSA, CA	SA, EC, AQ	None		

LSA-Large-scale assessment; Ex-Examinations; CA-Classroom assessment EC-Enabling context; SA-System alignment; AQ-Assessment quality

Source: Based on Liberman & Clarke (2012).

Lessons from Implementation Completion Reports for World Bank Projects

Implementation Completion Reports (ICRs) suggest that it is especially important to identify key stakeholders for assessment activities upfront and ensure that they are directly involved in project preparation.

- It is essential to build stakeholder consensus on assessment objectives and to include these objectives in the project's results framework.
- Further, it is important to evaluate existing infrastructure (physical space, IT management) and institutional constraints, identify risks and back-up strategies, and ensure that there is the minimum capacity required for implementation of the proposed assessment activities.

A multi-faceted and flexible approach in project implementation can contribute to developing assessment capacity, especially where local staff will need to sustain the activities.

• The ICRs conclude that it is especially important to deepen and broaden the capacity of local mid-level managers and to build teacher capacity in using newly-introduced assessment approaches.

Well-designed study tours were found to be useful in exposing country practitioners to good practice.

• Partnering local and international universities also was found to help build local human capital.

Ensuring sustainability after project completion is important and requires ensuring adequate fiscal and human resources and clearly defined roles and responsibilities.

• Disseminating assessment results to key stakeholders in a timely manner also contributed to better use of the assessment results and greater sustainability of associated efforts.

Support to Countries under the Russia Education Aid for Development (READ) Trust Fund

The READ program is an important part of Bank support for strengthening student assessment systems in selected countries.

- The \$32 million trust fund was established in October 2008 and provides both direct support to 8 countries (Angola, Armenia, Ethiopia, Kyrgyz Republic, Mozambique, Tajikistan, Vietnam, Zambia) and support at the global level for knowledge generation and sharing on assessment.
- It is the only aid program entirely dedicated to strengthening countries' assessment systems.

All 8 READ countries have completed "Self-Diagnosis" exercises to take stock of their assessment activities and identify gaps, formulated Action Plans based on the findings, and built consensus among key stakeholders for moving forward.

 Other important results include training to build local capacity, improved understanding of what makes for an effective assessment system, and concrete steps toward increased and stable domestic funding of assessment.

At the global level, READ funds SABER–Student Assessment, under which knowledge and tools are being developed based on the latest evidence for effective assessment policies and systems.

• Global products developed to date include questionnaires for collecting data on key characteristics of student assessment systems, rubrics to judge the stage of development of a country's assessment system, country reports summarizing findings from the application of the questionnaires and rubrics to a particular country's assessment system, country case studies describing the reform paths of a diverse mix of countries, and databases and papers on World Bank project assistance for student assessment.

Annual Global Conferences organized under READ have contributed to knowledge sharing and Russian visibility as a key emerging donor committed to improving education quality.

- Lessons learned from case studies on countries' experiences presented at these conferences include the importance of
 - o training of teachers in effective assessment practices (South Africa)
 - strong ministry support for assessment as key to the effective use of results to inform education policy (Singapore)
 - strengthening linkages across all assessment activities in an education system to maximize their impact (Poland)
 - defining national learning standards before creating an assessment system (Brazil)
 - carrying out in-depth analysis of assessment results to identify factors that affect learning outcomes (Russia)