



Key Policy Areas for Student Assessment

Status

1. Classroom Assessment

There is recognition among key stakeholders of the importance of classroom assessment and the urgent need to strengthen existing classroom assessment practices (which are very weak) in order to better support student learning. There also is recognition of the need to introduce systematic mechanisms to ensure the quality of classroom assessment practices. While such mechanisms have been piloted in some regions of the country with the support of donor-funded projects, there is a need to institutionalize them at the system level.



2. Examinations

The examination system has been operating on a regular basis under the provisions approved by the Government of the Kyrgyz Republic. As the current examination system is based on Soviet principles, it is in the process of being reformed to better align with the demands of the free-market economy. Efforts to improve the examination system have been welcomed by various stakeholders, including the leadership in charge of the examinations.



3. National Large-Scale Assessment (NLSA)

The Kyrgyz Republic has been conducting NLSAs on an irregular basis with the support of donor agencies. These assessments measure performance against national curriculum/learning standards. Various mechanisms are being used to better ensure the quality of these NLSAs, including pretesting items and providing technical training for staff. Policy makers have used NLSA results to modify aspects of the education system, including curriculum and assessment.



4. International Large-Scale Assessment (ILSA)

The Kyrgyz Republic has participated in two rounds of PISA (2006 and 2009). In both instances, it met all technical standards required to have its data presented in the main displays of the international report. Within the Kyrgyz Republic, results were presented in a consolidated national report and distributed to key decision makers and schools. The Kyrgyz Republic has not, however, taken concrete steps to participate in any other ILSAs and there is no policy document that addresses participation in future ILSAs.



Introduction

In 2009, the Kyrgyz Republic joined the Russia Education Aid for Development (READ) Trust Fund program, the goal of which is to help countries improve their capacity to design, carry out, analyze, and use assessments for improved student learning. As part of the READ Trust Fund program, and in order to gain a better understanding of the strengths and weaknesses of its existing assessment system, the Kyrgyz Republic participated in a formal exercise to benchmark its student assessment system under The World Bank's Systems Approach for Better Education Results (SABER) program. SABER is an evidence-based program to help countries systematically examine and strengthen the performance of different aspects of their education systems.

What is SABER-Student Assessment?

SABER-Student Assessment is a component of the SABER program that focuses specifically on benchmarking student assessment policies and systems. The goal of SABER-Student Assessment is to promote stronger assessment systems that contribute to improved education quality and learning for all.

National governments and international agencies are increasingly recognizing the key role that assessment of student learning plays in an effective education system. The importance of assessment is linked to its role in:

- (i) providing information on levels of student learning and achievement in the system;
- (ii) monitoring trends in education quality over time;
- (iii) supporting educators and students with real-time information to improve teaching and learning; and
- (iv) holding stakeholders accountable for results.

SABER-Student Assessment methodology

The SABER-Student Assessment framework is built on the available evidence base for what an effective assessment system looks like. The framework provides guidance on how countries can build more effective student assessment systems. The framework is structured around two main dimensions of assessment systems: the types/purposes of assessment activities and the quality of those activities.

Assessment types and purposes

Assessment systems tend to be comprised of three main types of assessment activities, each of which serves a different purpose and addresses different information needs. These three main types are: classroom assessment, examinations, and large-scale, system level assessments.

Classroom assessment provides real-time information to support ongoing teaching and learning in individual classrooms. Classroom assessments use a variety of formats, including observation, questioning, and paper-and-pencil tests, to evaluate student learning, generally on a daily basis.

Examinations provide a basis for selecting or certifying students as they move from one level of the education system to the next (or into the workforce). All eligible students are tested on an annual basis (or more often if the system allows for repeat testing). Examinations cover the main subject areas in the curriculum and usually involve essays and multiple-choice questions.

Large-scale, system-level assessments provide feedback on the overall performance of the education system at particular grades or age levels. These assessments typically cover a few subjects on a regular basis (such as every 3 to 5 years), are often sample based, and use multiple-choice and short-answer formats. They may be national or international in scope.

Appendix 1 summarizes the key features of these main types of assessment activities.

Quality drivers of an assessment system

The key considerations when evaluating a student assessment system are the individual and combined quality of assessment activities in terms of the adequacy of the information generated to support decision making. There are three main drivers of information quality in an assessment system: enabling context, system alignment, and assessment quality.

Enabling context refers to the broader context in which the assessment activity takes place and the extent to which that context is conducive to, or supportive of, the assessment. It covers such issues as the legislative or policy framework for assessment activities; institutional and organizational structures for designing, carrying out, or using results from the assessment; the availability of sufficient and stable sources of funding; and the presence of trained assessment staff.

System alignment refers to the extent to which the assessment is aligned with the rest of the education system. This includes the degree of congruence between assessment activities and system learning goals, standards, curriculum, and pre- and in-service teacher training.

Assessment quality refers to the psychometric quality of the instruments, processes, and procedures for the assessment activity. It covers such issues as design and implementation of assessment activities, analysis and interpretation of student responses to those activities, and the appropriateness of how assessment results are reported and used.

Crossing the quality drivers with the different assessment types/purposes provides the framework and broad indicator areas shown in Table 1. This framework is a starting point for identifying indicators that can be used to review assessment systems and plan for their improvement.

Table 1: Framework for building an effective assessment system, with indicator areas

	Assessment types/purposes		
	Classroom assessment	Examinations	Large-scale, system-level assessment
Enabling context	Policies Leadership and public engagement Funding Institutional arrangements Human resources		
System alignment	Learning/quality goals Curriculum Pre- and in-service teacher training opportunities		
Assessment quality	Ensuring quality (design, administration, analysis) Ensuring effective uses		

Source: World Bank.

The indicators are identified based on a combination of criteria, including:

- professional standards for assessment;
- empirical research on the characteristics of effective assessment systems, including analysis of the characteristics that differentiate between the assessment systems of low- versus high-performing nations; and
- theory—that is, general consensus among experts that it contributes to effective assessment.

Levels of development

The World Bank has developed a set of standardized questionnaires and rubrics for collecting and evaluating data on the three assessment types and related quality drivers.

The questionnaires are used to collect data on the characteristics of the assessment system in a particular country. The information from the questionnaires is then applied to the rubrics in order to judge the development level of the country's assessment system in different areas.

The basic structure of the rubrics for evaluating data collected using the standardized questionnaires is summarized in Appendix 2. The goal of the rubrics is to provide a country with some sense of the development level of its assessment activities compared to best or recommended practice in each area. For each indicator, the rubric displays four development levels—*Latent*,

Emerging, Established, and Advanced. These levels are artificially constructed categories chosen to represent key stages on the underlying continuum for each indicator. Each level is accompanied by a description of what performance on the indicator looks like at that level.

- *Latent* is the lowest level of performance; it represents absence of, or deviation from, the desired attribute.
- *Emerging* is the next level; it represents partial presence of the attribute.
- *Established* represents the acceptable minimum standard.
- *Advanced* represents the ideal or current best practice.

A summary of the development levels for each assessment type is presented in Appendix 3.

In reality, assessment systems are likely to be at different levels of development in different areas. For example, a system may be *Established* in the area of examinations, but *Emerging* in the area of large-scale, system-level assessment, and vice versa. While intuition suggests that it is probably better to be further along in as many areas as possible, the evidence is unclear as to whether it is necessary to be functioning at *Advanced* levels in all areas. Therefore, one might view the *Established* level as a desirable minimum outcome to achieve in all areas, but only aspire beyond that in those areas that most contribute to the national vision or priorities for education. In line with these considerations, the ratings generated by the rubrics are not meant to be additive across assessment types (that is, they are not meant to be added to create an overall rating for an assessment system; they are only meant to produce an overall rating for each assessment type). The methodology for assigning development levels is summarized in Appendix 4.

Education in the Kyrgyz Republic

The Kyrgyz Republic is a low income country in Central Asia. GDP per capita (current US\$, 2009) is \$871, with annual growth of approximately 1.7 percent.

At the primary level, net school enrollment is 86 percent; at the secondary level, it is 79 percent. Net enrollment for boys and girls is equal at both levels. The government has a policy of compulsory education for all children through grade 9. Grades 10-12 are non-compulsory. The primary completion rate is 98 percent.

Although nearly all students complete basic education, results from the 2006 and 2009 administration of the Program for International Student Assessment (PISA), which assesses the competencies of 15-year-olds in reading, mathematics, and science, demonstrate dramatically low levels of performance of Kyrgyz students as compared to the average scores of students from other participating countries. While there were improvements in the Kyrgyz Republic's scores on PISA between 2006 and 2009, analysis of the PISA results indicates that a lack of quality in teaching is a major determinant of low results on these assessments.

Government priorities for education reform include the revision and improvement of the national curriculum, introduction of improved teaching strategies and an incentive structure to enhance teacher performance, and increased monitoring of school fees.

Detailed information on the Kyrgyz Republic's student assessment system was obtained from the country's self-diagnosis report that was prepared as part of the READ Trust Fund program in 2009. At the time of the review, SABER-Student Assessment questionnaires had not been developed. Self-diagnosis information was analyzed using the SABER-Student Assessment rubrics. It is important to remember that the SABER-Student Assessment rubrics primarily focus on benchmarking a country's policies and arrangements for assessment activities at the system or macro level. Additional data would need to be collected to determine actual, on-the-ground practices in the Kyrgyz Republic, particularly by teachers and students in schools. The following sections discuss the findings by each assessment type, accompanied by suggested policy options. The suggested policy options were determined in collaboration with key local stakeholders based on the Kyrgyz Republic's immediate interests and needs. Detailed, completed rubrics for each assessment type in the Kyrgyz Republic are provided in Appendix 5.

Classroom Assessment

Level of development



In the Kyrgyz Republic, an informal system-level document provides guidelines for classroom assessment. The availability of this document is restricted. The one formal document in existence, does not apply to the entire system, but only to regions where classroom assessment projects have been piloted.

There are no system-wide resources for teachers to engage in classroom assessment activities. Although some manuals were developed under a World Bank-financed project, and subsequently distributed to all schools, these materials are nonetheless scarce.

There is an official curriculum or standards document that outlines what students are expected to learn, but the document is not specific, especially in identifying the required levels of student performance. There also are no system-level mechanisms to ensure that teachers develop the necessary skills and expertise in classroom assessment.

Classroom assessment practices generally appear to be weak, and there are no systematic mechanisms in place to monitor their quality. While some monitoring has been done in regions where classroom assessment projects have been piloted, there is as yet no monitoring in place at the system level.

Classroom assessment information is not required to be disseminated to key stakeholders. There also are no required uses of classroom assessment to support student learning.

Overall, there is recognition among key stakeholders of the importance of classroom assessment and the urgent need to strengthen existing practices in order to better support student learning.

Suggested policy options:

- 1. Make available system-wide resources for teachers to engage in better-quality classroom assessment as well as teacher training opportunities on effective classroom assessment practices.
- 2. Formalize, and make widely available, a system-level document that provides evidence-based guidelines for classroom assessment practices.

Examinations

Level of development



In the Kyrgyz Republic, the examination system has been operating on a regular basis under provisions approved by the government in the Soviet era. The Ministry of Education and Science annually issues orders on the regulations for the final state examinations of grade 9 and grade 11 graduates, and the university entrance examinations.

Final examinations are administered by teachers at schools. Essay topics for Kyrgyz, Russian, and Uzbek languages and literature are prepared by the Kyrgyz Academy of Education (KAE) and are distributed to schools by the Ministry of Education and Science two months before the examinations. KAE staff use a lottery system to select questions for the math examination from an existing database of questions (all of which have previously been made available to the public by KAE). The Center for Educational Assessment and Teaching Methods (CEATM) administers the university entrance examination.

Some opportunities that prepare individuals for work on these examinations are offered in the country. Some KAE and CEATM staff members acquire additional skills in assessment by attending training courses and seminars under international projects.

While the secondary school leaving examination receives consistent funding from the government, funding for the university entrance examination is irregular and provided by donors and through student fees. Funding covers design and peer review of tests, translation of test questions (into Kyrgyz or Russian), replication of test materials, organization and administration of tests, processing and distribution of test results, statistical analysis, and submission of findings to education authorities. Funding does not cover research and development.

While most stakeholders groups support the examinations, independent attempts have been made to improve the examinations by some groups (primarily donors). These efforts are generally welcomed by the leadership in charge of the examinations.

There is a clear understanding of what the examinations measure, and there is comprehensive material (such as questions likely to be on the test) available to prepare for the exams that is accessible to most students. While there are no up-to-date courses or workshops on the examinations available to teachers, teachers are involved in some examination-related tasks (for example, administering and grading the tests).

There is some technical documentation on the examinations, but it is not in a formal report format. There also are limited systematic mechanisms in place—such as item validity, and reliability checks—to better ensure the quality of the examinations.

Inappropriate behavior surrounding the examination process is moderate and includes leakage of the content of a paper prior to the examination; impersonation; copying from other candidates; using unauthorized materials, such as prepared answers and notes; collusion among candidates via mobile phones, passing of papers, or equivalent; and obtaining external assistance via, for example, the examination supervisor or mobile phones. There are no mechanisms in place to monitor the consequences of the examinations.

A small proportion of students (less than 10 percent) are unable to take the examinations. For example, there are no testing accommodations (e.g., braille tests, extended time) that would allow students with disabilities to take the school leaving exam. In addition, the cost of the university entrance exam and the distance to the testing center is prohibitive for some students.

Suggested policy options:

1. Enhance the credibility of the examinations by making publicly available a high-quality technical report on their design and conduct, and introducing systematic quality control mechanisms, such as internal and external review or audit.
2. Monitor the quality and equity consequences of the examinations by funding independent research on their impact or creating a permanent oversight committee.

National Large-Scale Assessment (NLSA)

Level of development



The Kyrgyz Republic has been conducting NLSAs on an irregular basis, and those that were conducted were driven by donor agencies. The Monitoring Learning Achievements assessment (MLA) of fourth grade students in mathematics, grammar, and life skills was conducted in 2001 and 2005, with UNICEF providing funding for the 2005 exercise. A National Sample Based Assessment (NSBA) of fourth and eighth grade students in mathematics, natural sciences, and reading comprehension was conducted in 2007 and 2009 using World Bank funding. There are no government policy documents regulating or authorizing NLSAs. While some Ministerial orders have been issued, these are not policy documents.

There is irregular funding allocated to NLSAs. The funding that has been provided (mainly by donors) has covered the design and peer review of assessment material, translation of test questions into Kyrgyz or Russian, replication of test materials, organization and administration of the assessments, processing and distribution of assessment results, statistical analysis of assessment results, and submission of findings to education authorities. Funding has not been sufficient to cover research and development activities.

To date, the NLSA program has been managed by a temporary group of people accountable to the Ministry of Education and Science. The NGO that conducted the NSBA was not appointed by the Government as a permanent agency to perform NSBA activities, but instead was contracted by the World Bank to carry out NSBA activities. In general, the people who have been contracted to carry out NLSA activities in the Kyrgyz Republic are not adequately trained to carry out these activities. The Kyrgyz Republic also does not currently offer opportunities that prepare individuals for work on NLSAs. Occasional courses or workshops on NLSA are made available to teachers; however, these are offered primarily through donor support.

While the NLSAs are supposed to measure student performance against curriculum or learning standards, there are no mechanisms in place to ensure that they are actually measuring this. Additionally, what these NLSAs measure is generally, questioned by stakeholder groups.

At the same time, there are some mechanisms in place to help ensure the quality of the NLSAs. For example, pretesting takes place, as does some preliminary training of interviewers and staff conducting statistical analysis. Quality control of NLSAs is conducted using the findings of psychometric analysis.

There is some documentation about the technical aspects of the NLSAs, such as test specifications that provide the basis for the compilation of test questions, but it is not in a formal report format.

NLSA results have not been disseminated in an effective way. However, NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessments. For example, policy makers have used NLSA results to modify areas of the education system, including curriculum and student assessment.

Currently, there are no mechanisms in place to monitor the consequences of NLSAs in the Kyrgyz Republic.

Suggested policy options:

1. Institutionalize the NLSA program by creating a policy framework, ensuring regular government funding to cover core activities, and creating a permanent NLSA office/unit that would be responsible for NLSA activities.
2. Build capacity to ensure the quality of future NLSA activities by, for example, offering training opportunities to prepare individuals to work on the NLSA.

International Large-Scale Assessment (ILSA)

Level of development



The Kyrgyz Republic has participated in PISA 2006 and PISA 2009. The Kyrgyz Republic has not put in place a policy that addresses participation in ILSAs and has not taken concrete steps to participate in an ILSA in the next five years. Participation in the 2006 and 2009 rounds of PISA was financed under a World Bank project. The financing included funding of core ILSA activities. Funding did not cover research and development activities.

While a national coordinator and a team from the Center for Educational Assessment and Teaching Methods (CEATM) carried out the ILSA activities, CEATM staff lacked the capacity needed to carry out the assessment activities effectively.

Opportunities to learn about ILSA, including training in basic assessment techniques, are available only to the Kyrgyz Republic's ILSA team members. The ILSA team was able to attend some of the international workshops and meetings for the 2006 and 2009 PISA exercises.

For both the 2006 and 2009 PISA exercises, the Kyrgyz Republic met all of the technical standards required to have its data presented in the main displays of the international report. However the country has not

contributed new knowledge on ILSA to the global evidence base.

There is limited media coverage of ILSA results in the Kyrgyz Republic. Country-specific results and information are disseminated irregularly throughout the country, and products to provide feedback to schools and educators are only sometimes made available. PISA 2009 results were presented in a consolidated report, and distributed to decision makers and schools across the country.

It is not clear whether and to what extent decisions based on ILSA results have had a positive impact on students' achievement levels in the Kyrgyz Republic.

Suggested policy options:

1. Build capacity in the Kyrgyz Republic to carry out ILSA activities by offering workshops on using international assessment databases, conducting university courses on the topic of international assessments, and providing funding for attending international workshops or training on international assessments.
2. Build policymaker interest in, and support for, participation in future ILSAs by informing them and other key stakeholders (media, educators) of ILSA results and analysis, and their policy and practical implications.

Appendix 1: Assessment Types and Their Key Differences

	Classroom	Large-scale assessment Surveys		Examinations	
		National	International	Exit	Entrance
Purpose	To provide immediate feedback to inform classroom instruction	To provide feedback on overall health of the system at particular grade/age level(s), and to monitor trends in learning	To provide feedback on the comparative performance of the education system at particular grade/age level(s)	To certify students as they move from one level of the education system to the next (or into the workforce)	To select students for further educational opportunities
Frequency	Daily	For individual subjects offered on a regular basis (such as every 3-5 years)	For individual subjects offered on a regular basis (such as every 3-5 years)	Annually and more often where the system allows for repeats	Annually and more often where the system allows for repeats
Who is tested?	All students	Sample or census of students at a particular grade or age level(s)	A sample of students at a particular grade or age level(s)	All eligible students	All eligible students
Format	Varies from observation to questioning to paper-and-pencil tests to student performances	Usually multiple choice and short answer	Usually multiple choice and short answer	Usually essay and multiple choice	Usually essay and multiple choice
Coverage of curriculum	All subject areas	Generally confined to a few subjects	Generally confined to one or two subjects	Covers main subject areas	Covers main subject areas
Additional information collected from students?	Yes, as part of the teaching process	Frequently	Yes	Seldom	Seldom
Scoring	Usually informal and simple	Varies from simple to more statistically sophisticated techniques	Usually involves statistically sophisticated techniques	Varies from simple to more statistically sophisticated techniques	Varies from simple to more statistically sophisticated techniques

Appendix 2: Basic Structure of Rubrics for Evaluating Data Collected on a Student Assessment System

Dimension	Development Level				
	LATENT (Absence of, or deviation from, attribute)	EMERGING (On way to meeting minimum standard)	ESTABLISHED (Acceptable minimum standard)	ADVANCED (Best practice)	Justification
EC—ENABLING CONTEXT					
EC1—Policies					
EC2—Leadership, public engagement					
EC3—Funding					
EC4—Institutional arrangements					
EC5—Human resources					
SA—SYSTEM ALIGNMENT					
SA1—Learning/quality goals					
SA2—Curriculum					
SA3—Pre-, in-service teacher training					
AQ—ASSESSMENT QUALITY					
AQ1—Ensuring quality (design, administration, analysis)					
AQ2—Ensuring effective uses					

Appendix 3: Summary of the Development Levels for Each Assessment Type

Assessment Type	LATENT	EMERGING	ESTABLISHED	ADVANCED
	<i>Absence of, or deviation from, the attribute</i>	<i>On way to meeting minimum standard</i>	<i>Acceptable minimum standard</i>	<i>Best practice</i>
CLASSROOM ASSESSMENT	There is no system-wide institutional capacity to support and ensure the quality of classroom assessment practices.	There is weak system-wide institutional capacity to support and ensure the quality of classroom assessment practices.	There is sufficient system-wide institutional capacity to support and ensure the quality of classroom assessment practices.	There is strong system-wide institutional capacity to support and ensure the quality of classroom assessment practices.
EXAMINATIONS	There is no standardized examination in place for key decisions.	There is a partially stable standardized examination in place, and a need to develop institutional capacity to run the examination. The examination typically is of poor quality and is perceived as unfair or corrupt.	There is a stable standardized examination in place. There is institutional capacity and some limited mechanisms to monitor it. The examination is of acceptable quality and is perceived as fair for most students and free from corruption.	There is a stable standardized examination in place and institutional capacity and strong mechanisms to monitor it. The examination is of high quality and is perceived as fair and free from corruption.
NATIONAL (OR SYSTEM-LEVEL) LARGE-SCALE ASSESSMENT	There is no NLSA in place.	There is an unstable NLSA in place and a need to develop institutional capacity to run the NLSA. Assessment quality and impact are weak.	There is a stable NLSA in place. There is institutional capacity and some limited mechanisms to monitor it. The NLSA is of moderate quality and its information is disseminated, but not always used in effective ways.	There is a stable NLSA in place and institutional capacity and strong mechanisms to monitor it. The NLSA is of high quality and its information is effectively used to improve education.
INTERNATIONAL LARGE-SCALE ASSESSMENT	There is no history of participation in an ILSA nor plans to participate in one.	Participation in an ILSA has been initiated, but there still is need to develop institutional capacity to carry out the ILSA.	There is more or less stable participation in an ILSA. There is institutional capacity to carry out the ILSA. The information from the ILSA is disseminated, but not always used in effective ways.	There is stable participation in an ILSA and institutional capacity to run the ILSA. The information from the ILSA is effectively used to improve education.

Appendix 4: Methodology for Assigning Development Levels

1. The country team or consultant collects information about the assessment system in the country.

2. Based on the collected information, a level of development and score is assigned to each dimension in the rubrics:

- Latent = 1 score point
- Emerging = 2 score points
- Established = 3 score points
- Advanced = 4 score points

3. The score for each quality driver is computed by aggregating the scores for each of its constituent dimensions. For example:

The quality driver, 'Enabling Context,' in the case of ILSA, has 3 dimensions on which a hypothetical country receives the following scores: Dimension A = 2 points; Dimension B = 2 points; Dimension C = 3 points. The hypothetical country's overall score for this quality driver would be: $(2+2+3)/3 = 2.33$

4. A preliminary level of development is assigned to each quality driver.

5. The preliminary development level is validated using expert judgment in cooperation with the country team and The World Bank Task Team Leader.

For scores that allow a margin of discretion (i.e., to choose between two levels of development), a final decision has to be made based on expert judgment. For example, the aforementioned hypothetical country has an 'Enabling Context' score of 2.33, corresponding to a preliminary level of development of 'Emerging or Established.' Based on qualitative information not captured in the rubric, along with expert judgment, the country team chooses 'Emerging' as the most appropriate level.

6. Scores for certain key dimensions under 'Enabling Context' (in the case of EXAM, NLSA, and ILSA) and under 'System Alignment' (in the case of CLASS) were set as ceiling scores, i.e., the overall mean score for the particular assessment type cannot be greater than the score for these key dimensions. These key variables include formal policy, regular funding, having a permanent assessment unit, and the quality of assessment practices.





Appendix 5: SABER-Student Assessment Rubrics for the Kyrgyz Republic

This appendix provides the completed SABER-Student Assessment rubrics for each type of assessment activity in the Kyrgyz Republic. In each row of the rubric, the relevant selection is indicated by a thick border and an asterisk. The selection may include a superscript number that refers to the justification or explanation for the selection (as indicated by a thick border and an asterisk). The explanation or justification text can be located in the “Development level rating justifications” section at the end of each rubric. If a row includes a superscript, but not a thick border and an asterisk, this means that insufficient information was available to determine the relevant selection in the row.





KYRGYZ REPUBLIC *Classroom Assessment*

ENABLING CONTEXT AND SYSTEM ALIGNMENT

Overall policy and resource framework within which classroom assessment activity takes place in a country or system, and the degree to which classroom assessment activity is coherent with other components of the education system.

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT AND SYSTEM ALIGNMENT 1: <i>Setting clear guidelines for classroom assessment</i>			
There is no system-level document that provides guidelines for classroom assessment.	There is an informal system-level document that provides guidelines for classroom assessment. ¹ *	There is a formal system-level document that provides guidelines for classroom assessment.	This option does not apply to this dimension.
This option does not apply to this dimension.	This option does not apply to this dimension. ² *	The availability of the document is restricted.	The document is widely available.
ENABLING CONTEXT AND SYSTEM ALIGNMENT 2: <i>Aligning classroom assessment with system learning goals</i>			
There are no system-wide resources for teachers for classroom assessment. ³ *	There are scarce system-wide resources for teachers for classroom assessment.	There are some system-wide resources for teachers for classroom assessment.	There are a variety of system-wide resources available for teachers for classroom assessment.
There is no official curriculum or standards document.	There is an official curriculum or standards document, but it is not clear what students are expected to learn or to what level of performance.	There is an official curriculum or standards document that specifies what students are expected to learn, but the level of performance required is not clear. ⁴ *	There is an official curriculum or standards document that specifies what students are expected to learn and to what level of performance.
ENABLING CONTEXT AND SYSTEM ALIGNMENT 3: <i>Having effective human resources to carry out classroom assessment activities</i>			
There are no system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment. *	This option does not apply to this dimension.	There are some system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.	There are a variety of system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.

ASSESSMENT QUALITY*Quality of classroom assessment design, administration, analysis, and use.*

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ASSESSMENT QUALITY 1: <i>Ensuring the quality of classroom assessment</i>			
Classroom assessment practices suffer from widespread weaknesses or there is no information available on classroom assessment practices.	Classroom assessment practices are known to be weak. ⁵ *	Classroom assessment practices are known to be of moderate quality.	Classroom assessment practices are known to be generally of high quality.
There are no mechanisms to monitor the quality of classroom assessment practices. ⁶ *	There are ad hoc mechanisms to monitor the quality of classroom assessment practices.	There are limited systematic mechanisms to monitor the quality of classroom assessment practices.	There are varied and systematic mechanisms in place to monitor the quality of classroom assessment practices.
ASSESSMENT QUALITY 2: <i>Ensuring effective uses of classroom assessment</i>			
Classroom assessment information is not required to be disseminated to key stakeholders. ⁷ *	This option does not apply to this dimension.	Classroom assessment information is required to be disseminated to some key stakeholders.	Classroom assessment information is required to be disseminated to all key stakeholders.
There are no required uses of classroom assessment to support student learning. *	There are limited required uses of classroom assessment to support student learning.	There are adequate required uses of classroom assessment to support student learning, excluding its use as an input for external examination results.	There are adequate required uses of classroom assessment to support student learning, including its use as an input for external examination results.





Classroom Assessment: Development level rating justifications

1. The formal document that exists does not apply to the entire system. The existing documents only relate to pilot regions where classroom assessment has been piloted.
2. The availability of the informal system-level document that provides guidelines for classroom assessment activities is restricted.
3. There are no system-wide resources for teachers for classroom assessment Even after the provision of manuals developed under a World-Bank-financed project, which were distributed to all schools, these materials are very scarce.
4. The document is not specific, particularly in identifying the level of required performance.
5. It appears that classroom assessment practices are weak.
6. There is some monitoring done in pilot regions where classroom assessment has been introduced, but there is nothing at the system level.
7. There is no requirement to report the results of classroom assessment on a regular basis to anyone.

KYRGYZ REPUBLIC *Examinations*

ENABLING CONTEXT

Overall framework of policies, leadership, organizational structures, fiscal and human resources in which assessment activity takes place in a country or system and the extent to which that framework is conducive to, or supportive of, the assessment activity.

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT 1: <i>Setting clear policies</i>			
No standardized examination has taken place.	The standardized examination has been operating on an irregular basis.	The examination is a stable program that has been operating regularly. ¹ *	This option does not apply to this dimension
There is no policy document that authorizes the examination.	There is an informal or draft policy document that authorizes the examination.	There is a formal policy document that authorizes the examination. ² *	This option does not apply to this dimension.
This option does not apply to this dimension.	The policy document is not available to the public	The policy document is available to the public. ³ *	This option does not apply to this dimension.
This option does not apply to this dimension.	This option does not apply to this dimension.	The policy document addresses some key aspects of the examination. ⁴ *	The policy document addresses all key aspects of the examination.
ENABLING CONTEXT 2: <i>Having strong leadership</i>			
All stakeholder groups strongly oppose the examination or are indifferent to it.	Most stakeholder groups oppose the examination.	Most stakeholders groups support the examination. *	All stakeholder groups support the examination.
There are no attempts to improve the examination by stakeholder groups.	This option does not apply to this dimension.	There are independent attempts to improve the examination by stakeholder groups. ⁵ *	There are coordinated attempts to improve the examination by stakeholder groups.
Efforts to improve the examination are not welcomed by the leadership in charge of the examination	This option does not apply to this dimension.	Efforts to improve the examination are generally welcomed by the leadership in charge of the examination. *	This option does not apply to this dimension.

(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 3: <i>Having regular funding</i>			
There is no funding allocated for the examination.	There is irregular funding allocated for the examination. ⁶ *	There is regular funding allocated for the examination.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding covers some core examination activities: design, administration, data processing or reporting. ⁷ *	Funding covers all core examination activities: design, administration, data processing and reporting.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development. *	This option does not apply to this dimension.	Funding covers research and development.
ENABLING CONTEXT 4: <i>Having strong organizational structures</i>			
The examination office does not exist or is newly established.	The examination office is newly established. ¹⁰	The examination office is a stable organization. *	This option does not apply to this dimension.
The examination office is not accountable to an external board or agency.	This option does not apply to this dimension.	The examination office is accountable to an external board or agency. ⁸ *	This option does not apply to this dimension.
Examination results are not recognized by any certification or selection system.	Examination results are recognized by certification or selection system in the country. ⁹ *	Examination results are recognized by one certification or selection system in another country.	Examination results are recognized by two or more certification or selection system in another country.
The examination office does not have the required facilities to carry out the examination.	The examination office has some of the required facilities to carry out the examination.	The examination office has all of the required facilities to carry out the examination. *	The examination office has state of the art facilities to carry out the examination.





(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 5: <i>Having effective human resources</i>			
There is no staff to carry out the examination.	The examination office is inadequately staffed to effectively carry out the examination, issues are pervasive. ¹⁰ *	The examination office is adequately staffed to carry out the examination effectively, with minimal issues.	The examination office is adequately staffed to carry out the assessment effectively, with no issues.
The country does not offer opportunities that prepare for work on the examination.	This option does not apply to this dimension.	The country offers some opportunities that prepare for work on the examination. ¹¹ *	The country offers a wide range of opportunities that prepare for work on the examination.





SYSTEM ALIGNMENT*Degree to which the assessment is coherent with other components of the education system.*

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
SYSTEM ALIGNMENT 1: <i>Aligning examinations with learning goals and opportunities to learn</i>			
It is not clear what the examination measures.	This option does not apply to this dimension.	There is a clear understanding of what the examination measures. *	This option does not apply to this dimension.
What the examination measures is questioned by some stakeholder groups. *	This option does not apply to this dimension.	What is measured by the examination is largely accepted by stakeholder groups.	This option does not apply to this dimension.
Material to prepare for the examination is minimal and it is only accessible to very few students.	There is some material to prepare for the examination that is accessible to some students.	There is comprehensive material to prepare for the examination that is accessible to most students. ¹² *	There is comprehensive material to prepare for the examination that is accessible to all students.
SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the examination</i>			
There are no courses or workshops on examinations available to teachers.	There are no up-to-date courses or workshops on examinations available to teachers. ¹³ *	There are up-to-date voluntary courses or workshops on examinations available to teachers.	There are up-to-date compulsory courses or workshops on examinations for teachers.
Teachers are excluded from all examination-related tasks.	Teachers are involved in very few examination-related tasks.	Teachers are involved in some examination-related tasks. ¹⁴ *	Teachers are involved in most examination-related tasks. ¹⁹

ASSESSMENT QUALITY*Degree to which the assessment meets quality standards, is fair, and is used in an effective way.*

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ASSESSMENT QUALITY 1: <i>Ensuring quality</i>			
There is no technical report or other documentation.	There is some documentation on the examination, but it is not in a formal report format. ¹⁵ *	There is a comprehensive technical report but with restricted circulation.	There is a comprehensive, high quality technical report available to the general public.
There are no mechanisms in place to ensure the quality of the examination.	This option does not apply to this dimension.	There are limited systematic mechanisms in place to ensure the quality of the examination. ¹⁶ *	There are varied and systematic mechanisms in place to ensure the quality of the examination.
ASSESSMENT QUALITY 2: <i>Ensuring fairness</i>			
Inappropriate behavior surrounding the examination process is high.	Inappropriate behavior surrounding the examination process is moderate. ¹⁷ *	Inappropriate behavior surrounding the examination process is low.	Inappropriate behavior surrounding the examination process is marginal.
The examination results lack credibility for all stakeholder groups.	The examination results are credible for some stakeholder groups. ¹⁸ *	The examination results are credible for all stakeholder groups.	This option does not apply to this dimension.
The majority of the students (over 50%) may not take the examination because of language, gender, or other equivalent barriers.	A significant proportion of students (10%-50%) may not take the examination because of language, gender, or other equivalent barriers.	A small proportion of students (less than 10%) may not take the examination because of language, gender, or other equivalent barriers. ¹⁹ *	All students can take the examination; there are no language, gender or other equivalent barriers.

(CONTINUED)

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ASSESSMENT QUALITY 3: <i>Using examination information in a fair way</i>			
Examination results are not used in a proper way by all stakeholder groups.	Examination results are used by some stakeholder groups in a proper way. *	Examination results are used by most stakeholder groups in a proper way.	Examination results are used by all stakeholder groups in a proper way.
Student names and results are public. ²⁰ *	This option does not apply to this dimension.	Students' results are confidential.	This option does not apply to this dimension.
ASSESSMENT QUALITY 4: <i>Ensuring positive consequences of the examination</i>			
There are no options for students who do not perform well on the examination, or students must leave the education system.	There are very limited options for students who do not perform well on the examination. ²⁴	There are some options for students who do not perform well on the examination. ²¹ *	There is a variety of options for students who do not perform well on the examination.
There are no mechanisms in place to monitor the consequences of the examination. *	This option does not apply to this dimension.	There are some mechanisms in place to monitor the consequences of the examination.	There is a variety of mechanisms in place to monitor the consequences of the examination.

Examinations: Development level rating justifications





1. The MOES issues annual orders based on the regulations on final state examination of graduates and the procedure for transfer of students to the next grade. The examination system has been operating on a regular basis since the Soviet Union.
2. The procedure of student testing is established under the provisions approved by the Government of the Kyrgyz Republic.
3. The public can have access to the policy document.
4. The policy document does not address some key aspects of the examination, such as all administration procedures and examiner responsibilities.
5. Primarily by donors.
6. While the secondary school leaving examination exercise is funded consistently without donor aid, funding for the university entrance examination is irregular and is funded by donors and students.
7. Funding includes design and peer review of tests, translation of test questions (into Kyrgyz or Russian), replication of test materials, organization and administration of tests, processing and distribution of test results, statistical analysis and submission of findings to education authorities.
8. Accountable to the Ministry of Education.
9. Examination results are recognized by the certification or selection system in the Kyrgyz Republic.
10. Staff is not adequately trained to carry out examinations.
11. Some staff acquire skills in the assessment of students by attending training courses and seminars under various international projects. The test is designed at the CEATM office by teachers from schools and universities in Kyrgyzstan, who have attended special training and already have sufficient experience in the development of test questions.
12. Likely test questions are widely available to most students on the market.
13. Training is primarily done on the job.
14. Teachers may administer and grade tests, depending on the subject area.
15. There is some documentation on the examination, but it is not in a formal report format.

16. There are limited systematic mechanisms, such as determining item design, validity, and reliability, in place to ensure the quality of the examination.
17. Inappropriate behavior includes: leakage of the content of an examination paper or part of a paper prior to the examination; impersonation; copying from other candidates; using unauthorized materials such as prepared answers and notes; collusion among candidates via mobile phones, passing of paper, or equivalent; and obtaining external assistance via the supervisor, mobile phone etc.
18. Some stakeholders question school leaving examination and university entrance examination results.
19. For the university entrance examination, the cost of the examination and the distance to the testing center may be prohibitive. For the school leaving examination, students with disabilities are not provided for.
20. Results are not confidential.
21. There are some options for students who do not perform well on the examination. Students can repeat the grade, retake the examination, and/or may attend preparatory courses in order to prepare to retake the examination.

KYRGYZ REPUBLIC
National (or System-Level) Large-Scale Assessment (NLSA)

ENABLING CONTEXT

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 conducive to, or supportive of, the NLSA activity.

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT 1: <i>Setting clear policies for NLSA</i>			
No NLSA exercise has taken place.	The NLSA has been operating on an irregular basis. ¹ <i>*</i>	The NLSA is a stable program that has been operating regularly.	This option does not apply to this dimension.
There is no policy document pertaining to NLSA. ² <i>*</i>	There is an informal or draft policy document that authorizes the NLSA.	There is a formal policy document that authorizes the NLSA.	This option does not apply to this dimension.
This option does not apply to this dimension.	The policy document is not available to the public. <i>*</i>	The policy document is available to the public.	This option does not apply to this dimension.
There is no plan for NLSA activity.	This option does not apply to this dimension.	There is a general understanding that the NLSA will take place. <i>*</i>	There is a written NLSA plan for the coming years.
ENABLING CONTEXT 2: <i>Having strong public engagement for NLSA</i>			
All stakeholder groups strongly oppose the NLSA or are indifferent to it.	Some stakeholder groups oppose the NLSA. <i>*</i>	Most stakeholders groups support the NLSA.	All stakeholder groups support the NLSA.

(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 3: <i>Having regular funding for NLSA</i>			
There is no funding allocated to the NLSA.	There is irregular funding allocated to the NLSA. *	There is regular funding allocated to the NLSA.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding covers some core NLSA activities: design, administration, analysis and reporting. ³ *	Funding covers all core NLSA activities: design, administration, analysis and reporting.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development activities. *	This option does not apply to this dimension.	Funding covers research and development activities.
ENABLING CONTEXT 4: <i>Having strong organizational structures for NLSA</i>			
There is no NLSA office, ad hoc unit or team.	The NLSA office is a temporary agency or group of people. ⁴ *	The NLSA office is a permanent agency, institution or unit.	This option does not apply to this dimension.
This option does not apply to this dimension.	Political considerations regularly hamper technical considerations. *	Political considerations sometimes hamper technical considerations.	Political considerations never hamper technical considerations.
This option does not apply to this dimension.	The NLSA office is not accountable to a clearly recognized body.	The NLSA office is accountable to a clearly recognized body. ⁵ *	This option does not apply to this dimension.





(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 5: <i>Having effective human resources for NLSA</i>			
There is no staff allocated for running an NLSA.	The NLSA office is inadequately staffed to effectively carry out the assessment. ⁶ *	The NLSA office is adequately staffed to carry out the NLSA effectively, with minimal issues.	The NLSA office is adequately staffed to carry out the NLSA effectively, with no issues.
The country does not offer opportunities that prepare individuals for work on NLSA. *	This option does not apply to this dimension.	The country offers some opportunities to prepare individuals for work on the NLSA.	The country offers a wide range of opportunities to prepare individuals for work on the NLSA.

SYSTEM ALIGNMENT*Degree to which the NLSA is coherent with other components of the education system.*

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
SYSTEM ALIGNMENT 1: <i>Aligning the NLSA with learning goals</i>			
It is not clear if the NLSA is based on curriculum or learning standards.	This option does not apply to this dimension.	The NLSA measures performance against curriculum or learning standards. *	This option does not apply to this dimension.
What the NLSA measures is generally questioned by stakeholder groups. *	This option does not apply to this dimension.	What the NLSA measures is questioned by some stakeholder groups.	What the NLSA measures is largely accepted by stakeholder groups.
There are no mechanisms in place to ensure that the NLSA accurately measures what it is supposed to measure. *	There are ad hoc reviews of the NLSA to ensure that it measures what it is intended to measure.	There are regular internal reviews of the NLSA to ensure that it measures what it is intended to measure.	This option does not apply to this dimension.
SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the NLSA</i>			
There are no courses or workshops on the NLSA.	There are occasional courses or workshops on the NLSA. ⁷ *	There are some courses or workshops on the NLSA offered on a regular basis.	There are widely available high quality courses or workshops on the NLSA offered on a regular basis.

ASSESSMENT QUALITY*Degree to which the NLSA meets technical standards, is fair, and is used in an effective way.*

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ASSESSMENT QUALITY 1: <i>Ensuring the quality of the NLSA</i>			
No options are offered to include all groups of students in the NLSA. *	This option does not apply to this dimension.	At least one option is offered to include all groups of students in the NLSA.	Different options are offered to include all groups of students in the NLSA.
There are no mechanisms in place to ensure the quality of the NLSA.	This option does not apply to this dimension.	There are some mechanisms in place to ensure the quality of the NLSA. ⁸ *	There are a variety of mechanisms in place to ensure the quality of the NLSA.
There is no technical report or other documentation about the NLSA.	There is some documentation about the technical aspects of the NLSA, but it is not in a formal report format. ⁹ *	There is a comprehensive technical report but with restricted circulation.	There is a comprehensive, high quality technical report available to the general public.
ASSESSMENT QUALITY 2: <i>Ensuring effective uses of the NLSA</i>			
NLSA results are not disseminated.	NLSA results are poorly disseminated. ¹⁰ *	NLSA results are disseminated in an effective way. ¹⁷	This option does not apply to this dimension.
NLSA information is not used or is used in ways inconsistent with the purposes or the technical characteristics of the assessment.	This option does not apply to this dimension.	NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment. ¹¹ *	NLSA information is used by all stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.
There are no mechanisms in place to monitor the consequences of the NLSA. *	This option does not apply to this dimension.	There are some mechanisms in place to monitor the consequences of the NLSA.	There are a variety of mechanisms in place to monitor the consequences of the NLSA.





National (or System-Level) Large Scale Assessment (NLSA): Development level rating justifications

1. 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).
2. There are no government policy documents (informal or formal) regulating/authorizing the NLSAs. Both the MLA & NSBA were conducted with donor funding and while some Ministerial orders have been issued, these were not policy documents.
3. Funding includes design and peer review of tests, translation of test questions (into Kyrgyz or Russian), replication of test materials, organization and administration of tests, processing and distribution of test results, statistical analysis and submission of findings to education authorities.
4. The NGO that was conducting the NLSA was contracted by the World Bank. It was not appointed by the Government as a permanent agency for this work. Additionally, there are other institutions that are coming into the picture and may be playing the role of the 'NLSA office'.
5. To the Ministry of Education.
6. Staff is not adequately trained to carry out assessment activities.
7. There are occasional courses or workshops on the NLSA, offered primarily through donor support.
8. There are some mechanisms in place to ensure the quality of the NLSA. Pretesting takes place; there has been some preliminary training of interviewers and staff conducting statistical analysis; testing quality control was conducted using the findings of psychometric analysis.
9. There are specifications that provide the basis for compilation of test questions.
10. Results have not been disseminated in an effective way.
11. Policy makers have used NLSA results to modify areas of the education system, including the areas of curriculum and student assessment.

KYRGYZ REPUBLIC
International Large-Scale Assessment (ILSA)

ENABLING CONTEXT

Overall framework of policies, leadership, organizational structures, fiscal and human resources in which ILSA takes place in a country or system and the extent to which that framework is conducive to, or supportive of, ILSA activity.

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT 1: <i>Setting clear policies for ILSA</i>			
The country/system has not participated in an ILSA in the last 10 years.	This option does not apply to this dimension.	The country/system has participated in at least one ILSA in the last 10 years.	The country/system has participated in two or more ILSA in the last 10 years. ¹ *
The country/system has not taken concrete steps to participate in an ILSA in the next 5 years. ² *	This option does not apply to this dimension.	The country/system has taken concrete steps to participate in at least one ILSA in the next 5 years.	This option does not apply to this dimension.
There is no policy document that addresses participation in ILSA. ³ *	There is an informal or draft policy document that addresses participation in ILSA.	There is a formal policy document that addresses participation in ILSA.	This option does not apply to this dimension.
This option does not apply to this dimension.	The policy document is not available to the public.	The policy document is available to the public.	This option does not apply to this dimension.
ENABLING CONTEXT 2: <i>Having regular funding for ILSA</i>			
There is no funding for participation in ILSA.	There is funding from loans or external donors. ⁴ *	There is regular funding allocated at discretion.	There is regular funding approved by law, decree or norm.
This option does not apply to this dimension.	Funding covers some core activities of the ILSA.	Funding covers all core activities of the ILSA. ⁵ *	This option does not apply to this dimension.
Funding does not cover research and development activities. ⁶ *	This option does not apply to this dimension.	This option does not apply to this dimension.	Funding covers research and development activities.





(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 3: <i>Having effective human resources for ILSA</i>			
There is no team or national/system coordinator to carry out the ILSA activities.	There is a team or national/system coordinator to carry out the ILSA activities. ⁷ *	There is a team and national/system coordinator to carry out the ILSA activities.	This option does not apply to this dimension.
This option does not apply to this dimension.	The national/system coordinator or other designated team member may not be fluent in the language of the assessment.	The national/system coordinator is fluent in the language of the assessment. *	This option does not apply to this dimension.
This option does not apply to this dimension.	The ILSA office is inadequately staffed or trained to carry out the assessment effectively. ⁸ *	The ILSA office is adequately staffed or trained to carry out the ILSA effectively, with minimal issues.	The ILSA office is adequately staffed and trained to carry out the ILSA effectively, with no issues.

SYSTEM ALIGNMENT*Degree to which the ILSA meets technical quality standards, is fair, and is used in an effective way.*

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
SYSTEM ALIGNMENT 1: <i>Providing opportunities to learn about ILSA</i>			
The ILSA team has not attended international workshops or meetings.	The ILSA team attended some international workshops or meetings. *	The ILSA team attended all international workshops or meetings.	This option does not apply to this dimension.
The country/system offers no opportunities to learn about ILSA. *	This option does not apply to this dimension.	The country/system offers some opportunities to learn about ILSA.	The country/system offers a wide range of opportunities to learn about ILSA.
This option does not apply to this dimension.	This option does not apply to this dimension.	Opportunities to learn about ILSA are available to the country's/system's ILSA team members only. ⁹ *	Opportunities to learn about ILSA are available to a wide audience, in addition to the country's/system's ILSA team members.

ASSESSMENT QUALITY*Degree to which the ILSA meets technical quality standards, is fair, and is used in an effective way.*

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ASSESSMENT QUALITY 1: <i>Ensuring the quality of ILSA</i>			
Data from the ILSA has not been published.	The country/system met sufficient standards to have its data presented beneath the main display of the international report or in an annex.	The country/system met all technical standards required to have its data presented in the main displays of the international report. *	The country/system met all technical standards required to have its data presented in the main displays of the international report.
The country/system has not contributed new knowledge on ILSA. *	This option does not apply to this dimension.	This option does not apply to this dimension.	The country/system has contributed new knowledge on ILSA.
ASSESSMENT QUALITY 2: <i>Ensuring effective uses of ILSA</i>			
If any, country/system-specific results and information are not disseminated in the country/system.	Country/system-specific results and information are disseminated irregularly in the country/system. ¹⁰ *	Country/system-specific results and information are regularly disseminated in the country/system.	Country/system-specific results and information are regularly and widely disseminated in the country/system.
Products to provide feedback to schools and educators about the ILSA results are not made available.	This option does not apply to this dimension.	Products to provide feedback to schools and educators about the ILSA results are sometimes made available. ¹¹ *	Products to provide feedback to schools and educators about ILSA results are systematically made available.
There is no media coverage of the ILSA results.	There is limited media coverage of the ILSA results. ¹² *	There is some media coverage of the ILSA results.	There is wide media coverage of the ILSA results.
If any, country/system-specific results and information from the ILSA are not used to inform decision making in the country/system. ¹³ *	Results from the ILSA are used in a limited way to inform decision making in the country/system.	Results from the ILSA are used in some ways to inform decision making in the country/system.	Results from the ILSA are used in a variety of ways to inform decision making in the country/system.
It is not clear that decisions based on ILSA results have had a positive impact on students' achievement levels. ¹⁴ *	This option does not apply to this dimension.	This option does not apply to this dimension.	Decisions based on the ILSA results have had a positive impact on students' achievement levels.

International Large Scale Assessment (ILSA): Development level rating justifications

1. The Kyrgyz Republic has participated twice in PISA (2006 & 2009).
2. The Kyrgyz Republic has not taken concrete steps to participate in an ILSA in the next five years.
3. Participation in the two rounds of PISA was done through financing under a World Bank project. The Kyrgyz Republic itself, however, has not put in place a policy that addresses participation in ILSAs.
4. Both rounds of PISA were funded by a World Bank project.
5. Funding for the Kyrgyz Republic's participation in PISA was provided by the World Bank and covered all core activities.
6. Funding for the Kyrgyz Republic's participation in PISA was provided by the World Bank, and it did not cover research and development.
7. There was a team to carry out the ILSA activities.
8. Staff working on the ILSA lack capacity to carry out the ILSA effectively.
9. Opportunities to learn about ILSA, including training in the basic testing techniques, are available to the country's/system's ILSA team members only.
10. Results and information were presented in a consolidated report, and distributed across the country in order to reach decision makers and schools. Findings of the survey and sample tasks are publicly available and widely discussed across the country.
11. The consolidated report with results was prepared as a complete paper with attachments. It was distributed across the country in order to reach decision makers and all schools. Findings of the survey and sample tasks were publicly available and widely discussed across the country.
12. There is limited media coverage of the ILSA results in the Kyrgyz Republic.
13. ILSA results have not been used to inform decision making in the Kyrgyz Republic.
14. It is not clear whether and to what extent decisions based on ILSA results have had a positive impact on students' achievement levels.

Acknowledgements

This report was prepared by The World Bank SABER-Student Assessment team in collaboration with Dingyong Hou, World Bank Senior Education Specialist and Task Team Leader for education projects in the Kyrgyz Republic, and Aisuluu Bedelbayeva, World Bank Education Specialist.

References

Clarke, M. 2012. "What Matters Most for Student Assessment Systems: A Framework Paper." READ/SABER Working Paper Series. Washington, DC: World Bank.

Kyrgyz Republic. 2009. "Self-diagnosis in the Area of Student Assessment in Kyrgyzstan." Report prepared under the Russia Aid for Development Trust Fund program.

Organisation for Economic Co-operation and Development (OECD). 2010. "PISA 2009 at a Glance." Paris: OECD.

———. 2007. "PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis." Paris: OECD.

United Nations Children's Fund (UNICEF). 2010. "Country profile 2010: Education in Kyrgyzstan." Data retrieved from http://www.unicef.org/ceecis/Tajikistan_2010.pdf on March 12, 2013.

United Nations Educational, Scientific, and Cultural Organization (UNESCO) - Institute for Statistics. Kyrgyz Republic Country Indicator Data. Montreal, QC: UNESCO. Data retrieved from <http://stats.uis.unesco.org/unesco> on March 12, 2013

World Bank. World Bank Development Indicators. Kyrgyz Republic Country Indicator Data. Washington, DC: World Bank. Data retrieved from <http://databank.worldbank.org/data> on March 12, 2013.

The Systems Approach for Better Education Results (SABER) initiative produces comparative data and knowledge on education policies and institutions, with the aim of helping countries systematically strengthen their education systems. SABER evaluates the quality of education policies against evidence-based global standards, using new diagnostic tools and detailed policy data. The SABER country reports give all parties with a stake in educational results—from administrators, teachers, and parents to policymakers and business people—an accessible, objective snapshot showing how well the policies of their country's education system are oriented toward ensuring that all children and youth learn.

This report focuses specifically on policies in the area of student assessment.

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.



THE WORLD BANK