



Key Policy Areas for Student Assessment

1. Classroom Assessment

Although school-based assessments are part of the education system, there is no official system-level document that provides guidelines for classroom assessment activities in general. Several resources are available for teachers to use in carrying out classroom assessment activities, including descriptions of the performance levels that students are expected to reach in different subject areas at different grade and age levels. Some classroom assessment information is used as a required input to the external examination program, although it is unclear whether the results from these school-based assessments are moderated prior to combining them with scores from the external examination papers.

Status

Emerging



2. Examinations

The General Certificate of Education (GCE) Advanced Level Examination has been administered since 1964. It measures the performance of Grade 13 students on the national school curriculum, and is used for determining school graduation and student selection to university. GCE results are also used to monitor education quality levels and for planning education policy reforms. Students have access to various resources to prepare for the examination, including examples of the types of questions that are on the examination, information on how to prepare for the examination, and the framework document which explains what is measured on the examination. While teachers are involved in some examination-related tasks, such as examination administration and scoring, there are no training or professional development courses to prepare teachers for these tasks.

Established



3. National Large-Scale Assessment (NLSA)

The Grade 4 National Assessment of Achievement was administered in 2003, 2007, and 2009 to a representative sample of Grade 4 students. Students were assessed in first language (Sinhala or Tamil), Mathematics, and English. Funding for national assessment activities is made possible through donor support. While the National Education Research and Evaluation Centre (NEREC), which is in charge of the national assessment, is a permanent unit, it mainly employs temporary and part-time staff from institutions such as the Faculty of Education at the University of Colombo, the Department of Examinations, and the Department of Census and Statistics.

Emerging



4. International Large-Scale Assessment (ILSA)

Sri Lanka has not participated in an ILSA, and it does not have plans to do so in the near future.

Latent



Introduction

Sri Lanka has focused on increasing student learning outcomes by improving the quality of education in the country. An effective student assessment system is an important component to improving education quality and learning outcomes as it provides the necessary information to meet stakeholders' decision-making needs. In order to gain a better understanding of the strengths and weaknesses of its existing assessment system, Sri Lanka decided to benchmark this system using standardized tools developed 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 Sri Lanka

Sri Lanka is a lower middle income country in South Asia. GDP per capita (current US\$, 2012) is \$2,923, with annual growth of approximately 6.4 percent. After ending a 26-year internal military conflict in May 2009, Sri Lanka has demonstrated strong economic performance. The country's public debt and deficit have

gradually decreased as Sri Lanka transitions to middle income country status. Since the attainment of peace, the Sri Lankan Government can now focus on long-term strategic and structural development challenges.

The education system consists of primary school (grades 1 to 5), junior secondary (grades 6 to 9), and senior secondary (grades 9 to 13). Relative to countries with similar income levels, Sri Lanka performs well on education indicators, especially related to access and completion. In 2005, universal primary education was achieved with the net enrollment rate reaching 96 percent, and current primary completion rates are above 97 percent; in 2011, the net secondary enrollment rate was 85 percent. Sri Lanka has also demonstrated significant improvement in quality and learning outcomes. The National Assessment of Learning shows that achievement scores for grade 4 in language improved from 69 percent (2005) to 83 percent (2011.)

Sri Lanka currently faces challenges in transitioning to a knowledge-based economy. In general, there is a lack of workers with skills in information technology and the English language, as well as soft skills, such as problem-solving, strong communication, and the ability to work in teams. To address this skill gap, the overall objective of the Sri Lanka Education Sector Development Framework and Programme (ESDFP) 2012-2016 is to improve the education system by diversifying the secondary education curriculum to enable students to acquire the cognitive, affective, and psychomotor skills and values demanded by society.

Detailed information was collected on Sri Lanka's student assessment system using the SABER-Student Assessment questionnaires and rubrics in 2012. It is important to remember that these tools 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 Sri Lanka, 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 Sri Lanka's immediate interests and needs. Detailed, completed rubrics for each assessment type in Sri Lanka are provided in Appendix 5.

Classroom Assessment

Level of Development



In Sri Lanka, classroom assessment is used to diagnose student learning issues, provide feedback to students on their learning, and inform parents about their child's learning. Classroom assessment information is required to be disseminated to students and parents. Classroom assessment information is also used as an input to the external examination program, although it is unclear whether the results from the school-based assessments are moderated prior to combining them with the score from the external examination papers.

Although there is no official system-level document in place that provides guidelines for classroom assessment, several types of resources are available to teachers to carry out classroom assessment activities. For example, teachers are provided with Teacher Instruction Manuals (TIM) and Assessment and Evaluation guidelines that outline the performance levels that students are expected to reach in different subject areas at different grade and age levels. Teachers are also provided with books that include sample questions, and guidance on using appropriate scoring criteria when grading students' work.

In order to ensure that teachers develop expertise in classroom assessment, they are provided with pre- and in-service training through the National Colleges of Education and the National Institute of Education. There are currently no formal mechanisms for monitoring the quality of classroom assessment activities.

Suggested policy options:

1. Develop a system-level document that provides guidelines for carrying out classroom assessment activities in all grade levels and subject areas; make the document available to teachers, other key stakeholders, and the general public.
2. Introduce system-level mechanisms to monitor the quality of classroom assessment practices, such as making classroom assessment a required component of a teacher's performance evaluation and/or school inspection exercises.
3. Ensure suitable mechanisms are in place to moderate results from school-based assessments prior to integrating with external examination scores; provide training on moderation practices.

Examinations

Level of Development



The *General Certificate of Education (GCE) Advanced Level Examination* has been administered in Sri Lanka since 1964. It measures the performance of Grade 13 students on the national school curriculum, and is used for determining school graduation and student selection to university. Examination results are also used for monitoring education quality levels and planning education policy reforms. In order to prepare for the GCE, students have access to various resources, including sample questions that are on the examination, information on how to prepare for the examination, and a framework document that explains what is measured on the examination. Inappropriate behavior surrounding the examination process is low.

The *Public Examination Act No. 25 of 1968* document formally authorizes the GCE. This document addresses key aspects of the examination such as the distribution of power and responsibility among key entities, procedures to address security breaches, procedures to include all students in the examination, rules about examination preparation, the methodology for grading and marking the examination, and the use of examination results.

The Department of Examinations has been in charge of the GCE since 1968. While there is permanent staff within the Department of Examinations, it is insufficient to meet the needs of the examination. Staff must be brought in from universities to conduct and evaluate the examination, and staff from the National Institute of Education (NIE) participates in setting of the examination papers.

While teachers are involved in some examination-related tasks, such as administering and scoring the GCE, there are no training or professional development courses that prepare teachers for these activities.

The government of Sri Lanka allocates regular funding for the examination (funding is also sometimes provided by non-government sources). Funding covers all core examination activities such as design,

administration, data processing, reporting, as well as research and development. The World Bank provides technical assistance for item bank and examination guideline development.

There have been independent attempts to improve the examination by different stakeholders. Policymakers and universities have attempted to reform the examination through research, and NGOs have attempted to reform the examination through funding. In general, these efforts are welcomed by the Department of Examinations and the Ministry of Education.

Limited systematic mechanisms, such as internal review and translation verification, are in place to ensure the quality of the examination. While there are focus groups and research studies on the GCE, these activities are conducted on an ad hoc basis. There are no systematic mechanisms in place to monitor the consequences of the examination.

Suggested policy options:

1. Better ensure the quality of examination-related tasks for which teachers are responsible by introducing regular and mandatory courses for teachers involved in these activities.
2. Introduce additional systematic mechanisms to ensure the quality of the exam, including piloting and field testing of items and use of external review or observers.
3. Introduce systematic mechanisms to monitor the consequences of the exam by, for example, providing funding for research on its impact, and convening expert review groups on a regular basis.

National Large-Scale Assessment (NLSA)

Level of Development



The *Grade 4 National Assessment of Achievement* was administered in 2003, 2007, and 2009 to a representative sample of Grade 4 students. The plan for future assessment rounds is in the process of being prepared and will be available in April-May 2012.

Funding for the *Grade 4 National Assessment of Achievement* is made possible through donor support (as opposed to through government funding). Funding covers all core assessment activities, as well as research and development. Funding also covers staff training and participation in international programs on education measurement and evaluation.

The National Education Research and Evaluation Centre (NEREC) is the permanent unit in charge of the *Grade 4 National Assessment of Achievement*. However, NEREC mainly employs temporary and part-time staff from institutions such as the Faculty of Education at the University of Colombo, the Department of Examinations, and the Department of Census and Statistics. Various issues have been identified with the performance of staff responsible for carrying out assessment activities, including poor training of test administrators, unclear guidelines for administering the assessment, and errors in scoring, which have led to delays in reporting results.

In general, the *Grade 4 National Assessment of Achievement* measures performance against the national curriculum. Internal reviews of the alignment between the National Assessment of Achievement and what it is supposed to measure are conducted on a regular basis.

Sri Lanka employs a variety of mechanisms to ensure the quality of the *Grade 4 National Assessment of Achievement*. For example, all administrators are trained according to a protocol and are provided with a standardized manual for administration. A pilot is conducted before the main data collection takes place, all scorers are trained to ensure high inter-rater reliability, and there is double processing of data. After

the administration of the *Grade 4 National Assessment of Achievement*, a comprehensive technical report is prepared; however, its circulation is restricted.

Grade 4 National Assessment of Achievement results are disseminated to all stakeholder groups within 12 months of the administration of the assessment. Results are also featured in newspapers, radio, and other forms of media. Workshops on the results are held for key stakeholders and a dissemination seminar is held by the NEREC and the Ministry of Education in the provinces. The main reports on the results contain information on achievement levels and trends over time overall and by subgroups.

Suggested policy options:

1. Introduce regular government funding for national assessment activities, including for core activities such as test design, administration, analysis, and reporting, as well as for research and development.
2. Regularize the administration schedule of the national assessment program (for example, determine whether it should be held every two years or every three years).
3. Hire key permanent staff in the NEREC to manage national assessment activities; provide targeted training to the permanent staff on key aspects of the national assessment, including test administration and scoring.
4. Make high-quality courses or workshops on the national assessment available to teachers on a regular basis.

International Large-Scale Assessment (ILSA)

Level of Development



Sri Lanka has not participated in an ILSA, and it does not have plans to do so in the near future.

Suggested policy options:

1. Create an opportunity for high-level discussion among key stakeholders on key education policy questions or problems for which ILSA data could be useful.
2. Determine the need for, and possible next steps in relation to, participation in an ILSA exercise.

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 Sri Lanka

This appendix provides the completed SABER-Student Assessment rubrics for each type of assessment activity in Sri Lanka. 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.

Sri Lanka *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. ¹ *	There is an informal system-level document that provides guidelines for classroom assessment.	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.	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.	There are scarce system-wide resources for teachers for classroom assessment.	There are some system-wide resources for teachers for classroom assessment. ² *	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.	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. ³ *	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.

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ASSESSMENT QUALITY 1: <i>Ensuring the quality of classroom assessment</i></p>			
<p>Classroom assessment practices suffer from widespread weaknesses or there is no information available on classroom assessment practices. *</p>	<p>Classroom assessment practices are known to be weak.</p>	<p>Classroom assessment practices are known to be of moderate quality.</p>	<p>Classroom assessment practices are known to be generally of high quality.</p>
<p>There are no mechanisms to monitor the quality of classroom assessment practices. *</p>	<p>There are ad hoc mechanisms to monitor the quality of classroom assessment practices.</p>	<p>There are limited systematic mechanisms to monitor the quality of classroom assessment practices.</p>	<p>There are varied and systematic mechanisms in place to monitor the quality of classroom assessment practices.</p>
<p>ASSESSMENT QUALITY 2: <i>Ensuring effective uses of classroom assessment</i></p>			
<p>Classroom assessment information is not required to be disseminated to key stakeholders.</p>	<p>This option does not apply to this dimension.</p>	<p>Classroom assessment information is required to be disseminated to some key stakeholders.⁴ *</p>	<p>Classroom assessment information is required to be disseminated to all key stakeholders.⁸</p>
<p>There are no required uses of classroom assessment to support student learning.</p>	<p>There are limited required uses of classroom assessment to support student learning.</p>	<p>There are adequate required uses of classroom assessment to support student learning, excluding its use as an input for external examination results.</p>	<p>There are adequate required uses of classroom assessment to support student learning, including its use as an input for external examination results.⁵ *</p>

Classroom Assessment: Development-level rating justifications

1. While Sri Lanka does have School-Based Assessment, there are no system-level guidelines for classroom assessment in general.
2. Resources include the curriculum document that outlines what students are expected to learn in different subject areas at different grade/age levels, Teacher Instruction Manuals (TIM) and Assessment and Evaluation guidelines that outline the level(s) of performance that students are expected to reach in different subject areas at different grade/age levels, textbooks or workbooks that provide support for classroom assessment, mark schemes for students' work, and books with prototype questions.
3. Mechanisms include pre- and in-service teacher training.
4. Classroom assessment information is required to be disseminated to students and parents.
5. Classroom assessment activities are required to be used for diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child's learning, and providing input to an external examination program (e.g., school-based assessment with moderation and quality audit).

SRI LANKA *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. *	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. ³ *	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. ⁴ *	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.	There is regular funding allocated for the examination. ⁶ *	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.	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. ⁷ *
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.	The examination office is a stable organization. ⁸ *	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. ⁹ *	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.	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)

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

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. ¹² *	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. ¹³ *	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. ¹⁴ *	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.	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. ¹⁵ *	Teachers are involved in most examination-related tasks.

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.	There is a comprehensive technical report but with restricted circulation. ¹⁶ *	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. ¹⁷ *	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.	Inappropriate behavior surrounding the examination process is low. ¹⁸ *	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.	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.	All students can take the examination; there are no language, gender or other equivalent barriers. *

(CONTINUED)

<p>LATENT</p> <p>●○○○</p>	<p>EMERGING</p> <p>●●○○</p>	<p>ESTABLISHED</p> <p>●●●○</p>	<p>ADVANCED</p> <p>●●●●</p>
<p>ASSESSMENT QUALITY 3: <i>Using examination information in a fair way</i></p>			
<p>Examination results are not used in a proper way by all stakeholder groups.</p>	<p>Examination results are used by some stakeholder groups in a proper way.</p>	<p>Examination results are used by most stakeholder groups in a proper way.</p>	<p>Examination results are used by all stakeholder groups in a proper way. *</p>
<p>Student names and results are public.</p>	<p>This option does not apply to this dimension.</p>	<p>Students' results are confidential.¹⁹ *</p>	<p>This option does not apply to this dimension.</p>
<p>ASSESSMENT QUALITY 4: <i>Ensuring positive consequences of the examination</i></p>			
<p>There are no options for students who do not perform well on the examination, or students must leave the education system.</p>	<p>There are very limited options for students who do not perform well on the examination.</p>	<p>There are some options for students who do not perform well on the examination.²⁰ *</p>	<p>There is a variety of options for students who do not perform well on the examination.</p>
<p>There are no mechanisms in place to monitor the consequences of the examination.²¹ *</p>	<p>This option does not apply to this dimension.</p>	<p>There are some mechanisms in place to monitor the consequences of the examination.</p>	<p>There is a variety of mechanisms in place to monitor the consequences of the examination.</p>

Examinations: Development-level rating justifications

1. The GCE Advanced Level Examination has been administered since 1964. It is administered to grade 13 students and is used as certification for school completion and selection to higher education. The examination is also used to monitor education quality levels as well as for planning education policy reforms.
2. The policy document is the "Public Examination Act No. 25 of 1968."
3. The policy document outlines governance, distribution of power, and responsibilities among key entities; describes the purpose of the examination and the authorized uses of results; outlines procedures to investigate and address security breaches, cheating, or other forms of inappropriate behavior; outlines procedures for special/disadvantaged students; identifies rules about preparation; and addresses issues of confidentiality, security, the setting and marking of papers, as well as the powers of the Commissioner General.
4. While almost all stakeholder groups strongly support the examination, teacher unions sometimes oppose the examination of the results.
5. Policymakers and universities have attempted to reform the examination through research, and NGOs have attempted to reform the examination through funding.
6. There is regular funding allocated by the government and irregular funding allocated by non-government sources. There is also technical assistance provided by the World Bank for item bank and examination guideline development.
7. Activities also include long- or medium-term planning of program milestones.
8. The Department of Examinations has been operating independently since 1968.
9. The examination office is accountable to the Ministry of Education.
10. There is permanent or full-time staff, but it is insufficient to meet needs of the examination. Additional staff is brought from universities to conduct the examination and for evaluation, and staff from the National Institute of Education (NIE) is brought in for the setting of the examination papers. Issues that have sometimes been observed include poor training of test administrators or unclear instructions and guidelines in administering the examination, errors in scoring that have led to delays in results being reported, weaknesses in test design, errors in the application and efficiency of supervisors in entering admission numbers, etc. Additionally, schools do not always pay attention to all of the details in the information that they are required to provide, and sometimes submit incorrect information. Issues, however, are not pervasive.
11. One opportunity is available. Specifically, funding to attend international programs, courses, or workshops on educational measurement and evaluation. For example, five officers have been sent to Melbourne for courses in measurement and evaluation.

12. The examination measures the national school curriculum guidelines or standards.
13. However, some stakeholder groups question the examination in the sense that there are requests for further scrutiny of papers, etc.
14. Material includes examples of the types of questions that are on the examination, information on how to prepare for the examination, the framework document explaining what is measured on the examination, and a report on the strengths and weaknesses in student performance to the zones.
15. Tasks include the administering of the examination, scoring of the examination, acting as a judge (i.e., in orals), and supervising the examination procedures.
16. This is a confidential document with restricted circulation. There is no link, only a hard copy.
17. Mechanisms include internal review or observers and translation verification.
18. Inappropriate behaviors include impersonation, when an individual other than the registered candidate takes the examination, and copying from other candidates. Using unauthorized materials, such as prepared answers and notes, and issuing forged certificates or altering results has also been observed.
19. Only the student and persons with a legitimate, professional interest in the test-taker (e.g., his or her educators, parents, and authorized potential employers) can know the results.
20. Options include that students may retake the examination, attend remedial or preparatory courses in order to prepare to retake the examination, or opt for less selective schools/universities/tracks.
21. There are some focus groups, but they are not regular, as well as ad hoc studies on the examination.

SRI LANKA
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.	The NLSA is a stable program that has been operating regularly. ¹ *	This option does not apply to this dimension.
There is no policy document pertaining to NLSA.	There is an informal or draft policy document that authorizes the NLSA.	There is a formal policy document that authorizes the NLSA. 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. *	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. ³ *	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.	Most stakeholders groups support the NLSA. ⁴ *	All stakeholder groups support the NLSA.

(CONTINUED)

<p>LATENT</p> <p>●○○○</p>	<p>EMERGING</p> <p>●●○○</p>	<p>ESTABLISHED</p> <p>●●●○</p>	<p>ADVANCED</p> <p>●●●●</p>
<p>ENABLING CONTEXT 3: <i>Having regular funding for NLSA</i></p>			
<p>There is no funding allocated to the NLSA.</p>	<p>There is irregular funding allocated to the NLSA.⁵</p> <p style="text-align: right;">*</p>	<p>There is regular funding allocated to the NLSA.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Funding covers some core NLSA activities: design, administration, analysis and reporting.</p>	<p>Funding covers all core NLSA activities: design, administration, analysis and reporting.</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Funding does not cover research and development activities.</p>	<p>This option does not apply to this dimension.</p>	<p>Funding covers research and development activities.⁶</p> <p style="text-align: right;">*</p>
<p>ENABLING CONTEXT 4: <i>Having strong organizational structures for NLSA</i></p>			
<p>There is no NLSA office, ad hoc unit or team.</p>	<p>The NLSA office is a temporary agency or group of people.</p>	<p>The NLSA office is a permanent agency, institution or unit.⁷</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Political considerations regularly hamper technical considerations.</p>	<p>Political considerations sometimes hamper technical considerations.</p>	<p>Political considerations never hamper technical considerations.⁸</p> <p style="text-align: right;">*</p>
<p>This option does not apply to this dimension.</p>	<p>The NLSA office is not accountable to a clearly recognized body.</p>	<p>The NLSA office is accountable to a clearly recognized body.⁹</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>

(CONTINUED)

<p>LATENT</p> <p>● ○ ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○ ○</p>	<p>ADVANCED</p> <p>● ● ● ● ●</p>
<p>ENABLING CONTEXT 5: <i>Having effective human resources for NLSA</i></p>			
<p>There is no staff allocated for running an NLSA.</p>	<p>The NLSA office is inadequately staffed to effectively carry out the assessment.¹⁰</p> <p style="text-align: right;">*</p>	<p>The NLSA office is adequately staffed to carry out the NLSA effectively, with minimal issues.</p>	<p>The NLSA office is adequately staffed to carry out the NLSA effectively, with no issues.</p>
<p>The country does not offer opportunities that prepare individuals for work on NLSA.</p>	<p>This option does not apply to this dimension.</p>	<p>The country offers some opportunities to prepare individuals for work on the NLSA.¹¹</p> <p style="text-align: right;">*</p>	<p>The country offers a wide range of opportunities to prepare individuals for work on the NLSA.</p>

SYSTEM ALIGNMENT

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

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>SYSTEM ALIGNMENT 1: <i>Aligning the NLSA with learning goals</i></p>			
<p>It is not clear if the NLSA is based on curriculum or learning standards.</p>	<p>This option does not apply to this dimension.</p>	<p>The NLSA measures performance against curriculum or learning standards.¹²</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>What the NLSA measures is generally questioned by stakeholder groups.</p>	<p>This option does not apply to this dimension.</p>	<p>What the NLSA measures is questioned by some stakeholder groups.</p>	<p>What the NLSA measures is largely accepted by stakeholder groups.</p> <p style="text-align: right;">*</p>
<p>There are no mechanisms in place to ensure that the NLSA accurately measures what it is supposed to measure.</p>	<p>There are ad hoc reviews of the NLSA to ensure that it measures what it is intended to measure.</p>	<p>There are regular internal reviews of the NLSA to ensure that it measures what it is intended to measure.¹³</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the NLSA</i></p>			
<p>There are no courses or workshops on the NLSA.</p>	<p>There are occasional courses or workshops on the NLSA.¹⁴</p> <p style="text-align: right;">*</p>	<p>There are some courses or workshops on the NLSA offered on a regular basis.</p>	<p>There are widely available high quality courses or workshops on the NLSA offered on a regular basis.</p>

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. ¹⁵ *
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.	There are a variety of mechanisms in place to ensure the quality of the NLSA. ¹⁶ *
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.	There is a comprehensive technical report but with restricted circulation. ¹⁷ *	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.	NLSA results are disseminated in an effective way. ¹⁸ *	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. *	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. ¹⁹ *	There are a variety of mechanisms in place to monitor the consequences of the NLSA.

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

1. The Grade 4 National Assessment of Achievement was carried out in 2003, 2007, and 2009.
2. The document is the "Education Sector Development Framework and Programme" document, 2006.
3. The Plan for the coming years is in the process of being prepared and will become available in April/May 2012.
4. Department of Examinations, the National Institute of Education, universities, think tanks and NGOs strongly support the NLSA; policymakers, educators, students, and parents support the NLSA; the media is neutral to the NLSA.
5. The funding is allocated only by non-government sources.
6. Funding also covers long- or medium-term planning of program milestones, and staff training.
7. The agency is called the National Education Research and Evaluation Centre (NEREC).
8. Results from large-scale assessments have never been withheld from publication for political reasons.
9. The Ministry of Education (which instructs NEREC), the World Bank (which sponsors the assessment), and the University of Colombo (in which NEREC is part of the Faculty of Education).
10. There is mainly temporary or part-time staff involved. Specifically, NEREC employs members of the Faculty of Education, members of other universities, personnel from the Department of Examinations and the Department of Census and Statistics for running the assessments. Issues that have been identified with the performance of the staff include poor training of test administrators or unclear instructions and guidelines for administering the assessment, and there have been errors in scoring that have led to delays in results being reported.
11. Funding for attending international programs or courses or workshops on educational measurement and evaluation is provided to university staff from the Faculty of Education who are working on the NEREC projects. Research academics have been sent on three short-term visits and two long-term visits to the University of West of England for training in capacity building. Two academics have also been sent to the same university for a one-year master's degree. NEREC researchers were also sent on short-term visits in capacity building in areas such as preparation of questionnaires, data analysis, and sampling.
12. The NLSA measures performance against national/system or state-level curriculum guidelines or learning standards.
13. There are regular internal reviews of the alignment between the assessment instrument and what it is supposed to measure.

14. Most teachers have access to the live courses or workshops, most of which provide relevant resources that teachers can use in their classrooms. The courses are generally of high quality.
15. Special plans are made to ensure that the large-scale assessment is administered to students in hard-to-reach areas and the assessment is offered in the language of instruction for almost all student groups.
16. The mechanisms that are in place to ensure the quality of the assessment include: All proctors or administrators are trained according to a protocol; there is a standardized manual for large-scale assessment administrators; discrepancies must be recorded in a journal by administrators; a pilot is conducted before the main data collection takes place; all booklets are numbered; scorers are trained to ensure high inter-rater reliability; there is double processing of data; there are internal and external reviewers or observers. Specifically, observers are appointed for test administrators, and monitors are appointed in addition to supervisors. Panel marking is also carried out to ensure fairness.
17. There are reports made of each assessment incorporating the technical details.
18. Specifically, results are disseminated within 12 months after the large-scale assessment is administered, and reports with results are made available for all stakeholder groups. Results are also featured in newspapers, magazines, radio, or television. Additionally, the main reports on the results contain information on overall achievement levels and subgroups, as well as trends over time overall and for subgroups and standard errors. Workshops or presentations on the results are held for key stakeholders. One dissemination seminar is held by NEREC and the Ministry of Education in the provinces.
19. Where there are plans to introduce a permanent oversight committee, currently expert review groups and regular focus groups/surveys of key stakeholders are held.

SRI LANKA
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. ¹ *	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.
The country/system has not taken concrete steps to participate in an ILSA in the next 5 years. *	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.	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.	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.	This option does not apply to this dimension.
Funding does not cover research and development activities.	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.	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.	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.

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

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.	This option does not apply to this dimension.
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.	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.	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.	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.	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.	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. Sri Lanka has not participated in an ILSA, and it does not have plans to do so in the near future.

Acknowledgements

This report was prepared by The World Bank SABER-Student Assessment team in collaboration with Harsha Aturupane, World Bank Lead Education Specialist and Task Team Leader for education projects in Sri Lanka.

References

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

Ministry of Education. 2013. "2012 Annual Performance Report." Battaramulla, Sri Lanka. Data retrieved from http://www.moe.gov.lk/web/images/stories/publication/Annual_performance_2012/English.pdf on February 20, 2014.

World Bank. 2012. Sri Lanka Country Indicator Data. Washington, DC: World Bank. Data retrieved from <http://data.worldbank.org/> on February 20, 2014.

World Bank. 2012. *Sri Lanka - Country Partnership Strategy for the Period FY2013 - FY2016*. Report No. 66286-LA. Washington, DC: World Bank.

World Bank. 2012. *Project Performance Assessment Report Democratic Socialist Republic of Sri Lanka: Improving Relevance and Quality of Undergraduate Education and Education Sector Development Project*. Report No. 71893. Washington, DC: World Bank.

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.

