



Key Policy Areas for Student Assessment

1. Classroom Assessment

There is no system-level document that provides guidelines for classroom assessment activities. There is, however, an official curriculum that specifies what students are expected to learn, although the level of performance required is unclear. Regional Education Bureaus receive federal support that may be applied to classroom assessment activities within the context of their curriculum implementation and teacher development activities. The use of classroom assessment to support student learning, while required, is very limited.

Status



2. Examinations

The national examinations program is stable and has been operating on a regular basis. There is a formal, publicly-available policy document that authorizes the examinations. Regular funding, mainly from the registration fee that students must pay in order to take the examination, supports the examinations program. There is a clear understanding of what the examinations measure because they are based on the curriculum materials used to teach students. The examination results are credible for all stakeholder groups, and are used by most of these groups in an appropriate way.



3. National Large-Scale Assessment (NLSA)

The NLSA program has been operating on a regular basis with support from various donors. Although there is no formal document authorizing the NLSA, there is a general understanding that it will take place. While Ethiopia offers some courses to prepare individuals for work on the NLSA, such as those offered at Addis Ababa University, there are no courses on the NLSA geared to teachers. There are no mechanisms in place to monitor the consequences or impact of NLSA activities on education quality or achievement levels.



4. International Large-Scale Assessment (ILSA)

Ethiopia has not yet participated in an ILSA, although there is an informal plan to participate in one in the next few years.



Introduction

In 2009, Ethiopia 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, Ethiopia participated in a formal exercise to benchmark this 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 Ethiopia

Ethiopia is a low income country in East Africa. GDP per capita (current US\$, 2009) is \$350.8, with annual growth of approximately 6.5 percent. At the primary level, gross school enrollment is around 102 percent; at the secondary level, it is around 34 percent. While enrollment is high for both genders in primary school (98 percent for girls and 106 percent for boys), only 30 percent of girls and 38 percent of boys are enrolled in secondary school.

The *Ethiopia Education and Training Policy and Strategy*, introduced in 1994, created a new structure for primary and secondary schooling consisting of an 8-year primary education cycle, followed by two years of general secondary education and two years of preparatory secondary education. As a result of this strategy, the education system has a renewed focus on student learning and problem-solving skills, an emphasis on technical and vocational education and training, and support for the teaching of primary students in their mother tongue. The strategy also has supported the creation of self-contained classes in grades 1–4 (i.e., one teacher to cover all core subjects).

In recent years, Ethiopia's education sector has undergone significant expansion. At the same time, it continues to face challenges, especially in the area of quality. In response, the Government of Ethiopia introduced a comprehensive reform program to improve the quality of general education (Grades 1 through 12), referred to as the *General Education Quality Improvement Program* (GEQIP). The key objectives of GEQIP are to: (i) improve student learning outcomes; (ii) increase primary completion rates; and (iii) increase secondary enrolment rates.

Detailed information on Ethiopia'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 Ethiopia, 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 Ethiopia's immediate interests and needs. Detailed, completed rubrics for each assessment type are provided in Appendix 5.

Classroom Assessment

Level of development



In Ethiopia, there is no system-level document that provides guidelines for classroom assessment activities. There is, however, an official curriculum document that specifies what students are expected to learn, although the level of performance required is not clear.

Regional Education Bureaus (REB) receive federal support that may be applied to classroom assessment activities within the context of their curriculum implementation and teacher development activities (under the framework of the school improvement program).

There are no system-wide resources for teachers for classroom assessment, and there are no system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.

In general, classroom assessment practices suffer from widespread weaknesses, and monitoring of their quality is carried out only on an ad hoc basis.

Classroom assessment information is not required to be disseminated to key stakeholders. The use of classroom assessment to support student learning, while required, is in fact very limited.

Suggested policy options:

1. Build teacher capacity by making available system-wide resources/materials to make it easier for them to engage in classroom assessment, and by introducing a variety of system-level mechanisms to ensure that teachers develop the necessary skills and expertise in classroom assessment.
2. Improve the quality of classroom assessment practices by introducing systematic mechanisms to monitor their quality.

Examinations

Level of development

Established



In Ethiopia, students take the *Primary School Certificate Examination* at the end of grade 8, the *Ethiopian General Secondary Education Certificate Examination* at the end of grade 10, and the *Ethiopian Higher Education Entrance Examination* at the end of grade 12. These examinations have been operating on a regular basis. The *1994 Education and Training Policy* is a publically available policy document that authorizes these examinations and addresses some of their key aspects, including the grades in which they should be given, and the criteria for their grading.

Most stakeholders groups support the examinations. Funding for the examinations comes from the registration fees that examinees are required to pay in order to take them. Funding is usually sufficient to cover all core examination activities (design, administration, data processing, and reporting), but not research and development.

The examinations office has been operating since 1950, and is accountable to the Ministry of Education. It is currently undergoing reorganization in accordance with the Business Process Re-engineering (BPR) study undertaken by the Ministry of Education in 2009. The examinations office is adequately staffed for carrying out core examination activities. It has some, but not all, of the required facilities and equipment.

Some opportunities that prepare individuals for work on the examinations are available in the country, including academic courses on relevant topics offered by education and psychology departments in universities and colleges throughout the country.

While teachers are involved in some examination-related tasks, such as question writing and examination administration, there are no up-to-date courses or workshops on the examinations available to teachers.

Comprehensive material to prepare for the examinations is accessible by all students as the examinations are based on the regular curriculum taught to students in school.

Limited systematic mechanisms are in place to ensure the quality of the examinations. There has not been a tradition of producing a formal technical report on the examinations, but some informal documentation is available.

In appropriate behavior surrounding the examinations process is moderate and the results are viewed as credible by all stakeholder groups. All students can take the examinations; there are no language, gender, or other equivalent barriers. Student results are confidential. Examination results are used by most stakeholder groups in an appropriate way.

Currently, there are no mechanisms in place to monitor the consequences of the examinations.

Suggested policy options:

1. Make available comprehensive, high-quality technical reports on the examinations to the general public, and introduce varied and systematic mechanisms to better ensure the quality of the examinations.
2. Provide teachers with opportunities to learn about the examinations by introducing up-to-date compulsory courses or workshops on examinations and by involving teachers in more examination-related tasks.

National Large-Scale Assessment (NLSA)

Level of development

Emerging



The NLSA program in Ethiopia has been operating on a regular basis, with an NLSA exercise carried out every three to four years.

The first, second, and third *Ethiopian Baseline National Learning Assessments* (EBNLA) were conducted in the 1999/2000, 2003/2004, and 2006/2007 school years, respectively. The EBNLA assessed grade 4 students in English, Basic Reading, Mathematics, and Environmental Science, and grade 8 students in English, Mathematics, Physics, Chemistry, and Biology. The *Ethiopian First National Learning Assessment* (EFNLA) of students in grades 10 and 12 was conducted at the end of the 2008/2009 school year in English, Mathematics, Physics, Chemistry, and Biology.

While there is no policy document pertaining to NLSA activity in the country, there is a general understanding that the exercises will take place.

Funding for the NLSA exercises has been provided on an ad-hoc basis by the Ministry of Education and donors. Funding has tended to cover only basic aspects of the core assessment activities. At times, there has been insufficient funding to cover all core activities. Funding tends not to cover research and development.

In 2009, there was no permanent unit for the NLSA and an institutional home was still being worked out. In general, the NLSA office is a temporary group of people. The current team is comprised of four staff, some of whom have no background or training in NLSAs. However, Ethiopia does offer some opportunities to prepare individuals for work on the NLSA, such as assessment-related courses offered by Addis Ababa University. There are currently no courses or workshops on the NLSA available to teachers. Courses or workshops associated with previous NLSA exercises were only for policymakers and high-level educators.

The NLSA is based on the existing national curriculum and standards. Reviews to ensure that the assessment measures what it is intended to measure in relation to this curriculum are conducted on an informal and ad hoc basis.

Efforts are made to include all groups of students in the NLSA in that the assessment is translated into the language of instruction for each region.

There is some documentation about the technical aspects of the NLSA, but it is not in a formal report format.

NLSA results have not been widely disseminated to key stakeholders. Few copies of the reports are available. The 2009 NLSA results were, to a certain extent, used for curriculum development and teacher training.

Currently, there are no mechanisms in place to monitor the consequences or impact of the NLSA on education quality or achievement levels.

Suggested policy options:

1. Institutionalize the NLSA by creating a publicly-available policy document authorizing the program, and introducing regular government funding for carrying out all core NLSA activities.
2. Ensure that the NLSA supports student learning by conducting regular internal reviews to ensure that it measures what it is intended to measure, and introducing a variety of mechanisms to monitor the consequences and impact of the NLSA.

International Large-Scale Assessment (ILSA)

Level of development



Ethiopia has not participated in an ILSA in the last 10 years, and the country has not taken concrete steps to participate in one in the next 5 years (although apparently, there are informal plans to do so).

Suggested policy options:

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

	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 Ethiopia





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

ETHIOPIA

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.	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.	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.	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.	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. *	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. ² *	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. There are no classroom assessment guidelines at the federal level except indirectly through support to the Regional Education Bureaus (REBs) for curriculum implementation and teacher development activities under the framework of the school improvement program.
2. Use of classroom assessment information is required, but limited.

ETHIOPIA

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.	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. ⁵ *	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. ⁵ *	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. ⁷ *	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development. ⁸ *	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.	The examination office is a stable organization. ⁹ *	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. ¹⁰ *	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. ¹¹ *	Examination results are recognized by two or more certification or selection system in another country. ¹¹ *
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. ¹² *	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.¹³ *</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.¹⁴ *</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. *	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. ¹⁵ *	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. *	There is comprehensive material to prepare for the examination that is accessible to all students. ¹⁶ *
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. ¹⁷ *	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. ¹⁸ *	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. ¹⁹ *	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. ²⁰ *	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.	The examination results are credible for all stakeholder groups. ²¹ *	This option does not apply to this dimension.
The majority of the students (over 50 percent) 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. ²² *

(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.	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.	There are some options for students who do not perform well on the examination. ²³ *	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. ²⁴ *	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, 2, 3, 4. The grades 8, 10 and 12 examinations have been operating for many years. The school leaving certificate examination has been in place since 1950. The 1994 Education and Training Policy provides for examinations at three levels (Grades 8, 10, and 12). The above-referenced policy document is available to the public. (But there is no specific examinations policy document). The above document determines the grades in which the examinations should be given, and the criteria for grading.
5. The leadership employs additional personnel, invests in modern technology and works to make the examination entity semi- autonomous.
6. Every examinee is required to pay a registration fee in order to take the examination. The fund for examinations comes from fees collected from the examinees.
7. In 2009, funding for Grades 10 and 12 examinations was sufficient to cover core activities.
8. Based on Dr. Gregory's 2011 draft report on the status of national examinations activity in Ethiopia, it seems that the agency in charge of the Grades 10 and 12 examinations has not had sufficient funds over the past few years to carry out R&D or upgrade equipment or examination techniques.
9. The examination office has been in existence since 1950 in one form or another. It is currently undergoing reorganization in accordance with the BPR study undertaken by the MoE in 2009.
10. The examination agency is accountable to the MoE.
11. There is recognition by many systems in other countries.
12. Based on Dr. Gregory's 2011 draft report on the status of national examinations activity in Ethiopia, it seems that the agency in charge of the Grades 10 and 12 examinations does not have all of the required facilities for carrying out secure and high-quality examinations, and that this has been the case for several years.
13. Staffing seems to be adequate for core activities. But there are issues in terms of the ability to attract the right staff and retain them.
14. In 2009, teachers, university lecturers and subject matter experts were involved at various stages of the examinations process. Education and Psychology departments in universities and colleges throughout the country offer courses that prepare students for work in examinations.
15. Complaints about exam results are very minimal.
16. The examinations are based on the curriculum materials with which the students are taught, which are provided by the schools to the students.
17. With regard to national examinations, teachers are involved in exam writing and exam administration tasks.

18. Based on Dr. Gregory's 2011 draft report on the status of national examinations activity in Ethiopia, it seems that there has not been a tradition of producing a formal technical report on the Grades 10 and 12 examinations, but that some informal documentation is available.
19. Based on Dr. Gregory's 2011 draft report on the status of national examinations activity in Ethiopia, it seems that there has not been a strong tradition of quality control or assurance mechanisms for the Grades 10 and 12 examinations.
20. There have been allegations of cheating or other inappropriate behavior in relation to the national examinations. But these allegations were not high when seen in relation to the total number of exam takers.
21. Exam results are credible for all groups. There is no evidence showing otherwise.
22. There are no barriers for students.
23. Students who do not join the preparatory programme may join TVET programmes or take the examinations again and improve their grades. Students who fail (score below the pass mark) may take short skills training to join the labor market or retake the exam.
24. Based on Dr. Gregory's 2011 draft report on the status of national examinations activity in Ethiopia, it seems that there are no mechanisms in place to monitor the consequences of these examinations.

ETHIOPIA
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. ² *	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. ³ *	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. ⁴ *	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.	All stakeholder groups support the NLSA. ⁵ *

(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. ⁶ *	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. ⁷ *	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. ⁸ *	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. ⁵	The NLSA office is a temporary agency or group of people. ⁹ *	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. ¹⁰ *	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. ¹¹ *	The NLSA office is accountable to a clearly recognized body.	This option does not apply to this dimension.

(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.¹² *</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.¹³ *</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.

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. ¹⁴ *	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. ¹⁵ *
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. ¹⁶ *	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. ¹⁷ *	There are occasional courses or workshops on the NLSA.	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. ¹⁸ *	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. ¹⁹ *	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. ²⁰ *	There is a comprehensive technical report but with restricted circulation. ¹⁶	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. ²¹ *	NLSA results are disseminated in an effective way.	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. ²² *	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. ²³ *	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. In 2009, the NLSA program in Ethiopia was operating on a regular basis. But funding for the various NLSA exercises was being sourced from different donors, and the assessments were taking place roughly every 3 to 4 years.
2. At the start of READ (2009), Ethiopia did not have any kind (formal, informal, draft) of policy document on NLSA activity.
3. There was no policy document available in 2009.
4. Although there was no formal policy document underpinning the NLSA in 2009, there was a general understanding that the NLSA would take place every 3 to 4 years.
5. Based on our information, there is no opposition to NLSA.
6. NLSA activity is partially funded by the MOE and partially by donors. The funding is still ad hoc.
7. Funding has tended to cover only basic aspects of NLSA activities. Sometimes, there has been insufficient funding to cover all core activities.
8. Funding has primarily focused on supporting the actual carrying out of NLSA activities and not on R&D or secondary analysis.
9. In 2009, the NLSA team was comprised of a small number of staff (4), some of whom had no background or training in NLSA. There was no permanent unit, and an institutional home was still being worked out.
10. No precedence of political considerations hampering technical considerations.
11. In 2009, the NLSA office was not accountable to a clearly recognized body as it was in transition.
12. In 2009, the NLSA office did not have enough staff to effectively carry out NLSA activities.
13. There were some assessment-related courses offered by Addis Ababa University
14. NLSA was aligned with existing curriculum and standards.
15. MOE and other stakeholders have accepted the NLSA
16. In 2009, there were some procedures in place for reviewing the alignment of the NLSA test with the constructs/content it was intended to measure, but these procedures were not formalized or standardized.

17. The only courses or workshops associated with previous NLSA exercises have been for policymakers and high-level educators, and not for classroom teachers.
18. The NLSA is translated into the language of instruction for each region.
19. In 2009, there were some procedures in place for reviewing the alignment of the NLSA test with the constructs/content it was intended to measure. This would allow us to say that there were 'some mechanisms in place to ensure the quality of the NLSA.'
20. In 2009, no formal technical reports were available for the NLSA.
21. In 2009, NLSA results were not being widely disseminated to key stakeholders. Few copies of the report were available.
22. In 2009, the NLSA results were, to a certain extent, used for curriculum development and teacher training.
23. In 2009, there were no mechanisms in place to monitor the consequences of the NLSA.

ETHIOPIA
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: Setting clear policies for ILSA			
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. 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.	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: Having regular funding for ILSA			
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.	Funding covers all core activities of the ILSA.
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)

<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 effective human resources for ILSA</i></p>			
<p>There is no team or national/system coordinator to carry out the ILSA activities.</p>	<p>There is a team or national/system coordinator to carry out the ILSA activities.</p>	<p>There is a team and national/system coordinator to carry out the ILSA activities.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>The national/system coordinator or other designated team member may not be fluent in the language of the assessment.</p>	<p>The national/system coordinator is fluent in the language of the assessment.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>The ILSA office is inadequately staffed or trained to carry out the assessment effectively.</p>	<p>The ILSA office is adequately staffed or trained to carry out the ILSA effectively, with minimal issues.</p>	<p>The ILSA office is adequately staffed and trained to carry out the ILSA effectively, with no issues.</p>

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.	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.	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. As of 2009 (the reference year for this baseline rating), Ethiopia had yet to participate in an ILSA exercise. Nor did it have concrete plans to participate in any upcoming exercises. The overall rating for this assessment area, therefore, is Latent.
2. There is an informal plan by the MOE to participate in an ILSA in the coming two years.

Ethiopia has not participated in an international large-scale assessment and has not taken concrete steps to participate in an international large-scale assessment in the next five years. Therefore, there is a sufficient amount of information to determine the development level of the ILSA system in Ethiopia, and the remainder of the ILSA rubric is intentionally left blank.

Acknowledgements

This report was prepared by The World Bank SABER-Student Assessment team in collaboration with Rajendra Dhoj Joshi, World Bank Senior Education Specialist and Task Team Leader for education projects in Ethiopia.

References

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

Matters, G., and J. Pitman. 2010. "Review of National Learning Assessment Activity in Ethiopia: Status and Next Steps." Melbourne: ACER. Report prepared under the Russia Education Aid for Development Trust Fund program.

Netherlands Organisation for International Cooperation in Higher Education (NUFFIC). 2012. Country Module – Ethiopia. The Hague: NUFFIC.

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

World Bank. 2005. "Education in Ethiopia: Strengthening the Foundation for Sustainable Progress." Washington, DC: World Bank.

———. World Bank Development Indicators. Ethiopia 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.

