80066

Nepal

STUDENT ASSESSMENT

Key Policy Areas for Student Assessment

Status

1. Classroom Assessment

The National Curriculum Framework for School Education provides guidelines for classroom assessment. This document is widely available to the public online and through libraries, and is directly available to teachers through in-service teacher training courses. There are a variety of system-level mechanisms to help ensure that teachers develop the necessary expertise in classroom assessment, including in- and pre-service teacher training, and opportunities to participate in conferences and workshops. However, classroom assessment practices are generally not aligned with the national curriculum framework, and are considered weak. It is common to observe errors in the grading of students' work, teachers tend to provide little useful feedback to students, and parents are poorly informed about students' grades.

Emerging
●●○○

2. Examinations

The Higher Secondary Education Board Examination has been administered since 1992. The examination is formally used for high school graduation and university entrance, as well as informally for monitoring education quality levels, planning education policy reforms, school and educator accountability, and promoting competition among schools. The examination is funded entirely by student fees, which cover all core examination activities, but do not cover research and development. The government also allocates funds that could be used for the examination program, but it does not provide funding specifically for the examination. While the examination office has state-of-the-art facilities, there is an insufficient number of qualified staff to adequately carry out examination activities.

Established
●●●○

3. National Large-Scale Assessment (NLSA)

The Ministry of Education is currently conducting the National Assessment of Student Achievement (NASA) at the eighth grade, which is the country's first nationally representative, large-scale assessment exercise. The students in the sample are being tested in Nepali, Mathematics, and Social Studies, and the results will be published by August 2012. The NASA unit, which reports to the Education Review Office in the Ministry of Education, has been established as a permanent body responsible for carrying out NASA activities. Funding for NASA has been provided by the Government of Nepal, and covers all core, as well as research and development, activities. The government of Finland is also providing technical support for research and development.

Emerging
●●○○

4. International Large-Scale Assessment (ILSA)

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

Latent
●○○○

Systems Approach for Better Education Results

THE WORLD BANK

Introduction

Nepal 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, Nepal 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

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 Nepal

Nepal is a low income country in South Asia. GDP per capita (current US$, 2012) is $690, with annual growth of approximately 5 percent. Nepal is recovering from a 10-year armed conflict that ended in 2006 and is going

through a period of political instability and transition. Despite modest economic performance over the past 10 years, Nepal has shown impressive progress on poverty reduction and human development outcomes.

Nepal has made significant improvements in access to education. Primary school enrollment is close to universal, and primary completion is at 78 percent. Net enrollment rate disparities in access to primary education across regions, income levels, and ethnic groups have narrowed. In particular, girls have increasingly gained access to schooling: gender parity has been achieved in primary and secondary education. Furthermore, completion rates for girls at the primary level are slightly higher than for boys.

In order to further its reform, the Government of Nepal's School Sector Reform Plan (2009-2015) for Grades 1-12 focuses on the three pillars of enhancing access, promoting inclusion, and improving quality. The program supports the following areas: Basic Education (grades 1-8), including Early Childhood Education and Literacy and lifelong learning; Secondary Education (grades 9-12), which includes technical education and vocational training; and Institutional Capacity Strengthening for delivery and monitoring of education services and products. The key policy reforms being supported include increased community management of schools, public-private partnership for improved delivery of textbooks, and improved learning outcomes through the establishment of a standardized student assessment system.

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

Classroom Assessment

Level of Development	Emerging

The 2005 *National Curriculum Framework for School Education* provides guidelines for classroom assessment. This document is widely available to the public online and through libraries, and is available to teachers through in-service teacher training courses.

There are a variety of system-wide resources available for teachers to engage in classroom assessment activities, including a document that outlines what students are expected to learn and to what level of performance in different subject areas at different grade and age levels. Resources also include workbooks that provide support for classroom assessment, scoring criteria for students' work, as well as item banks with examples of questions.

System-level mechanisms are in place to ensure that teachers develop the necessary expertise in classroom assessment, including in- and pre-service teacher training, and opportunities to participate in conferences and workshops. All teacher training programs include a required course on classroom assessment. Teacher supervision also includes a component that is focused on classroom assessment.

In general, classroom assessment practices are not aligned with the national curriculum framework, and are considered weak. It is common to observe errors in the grading of students' work, teachers tend to provide little useful feedback to students, and parents are poorly informed about students' grades.

There are various systematic mechanisms in place to monitor the quality of classroom assessment activities. Classroom assessment is a required component of a teacher's performance evaluation and of school inspection. In addition, there are system-wide reviews of the quality of education, which include a focus on classroom assessment. Government funding is available for research on the quality of classroom assessment activities and on how to improve the quality of classroom assessment practices.

In Nepal, there are adequate required uses of classroom assessment to support student learning. Such uses include diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child's learning, planning the next steps in instruction, grading students for internal classroom uses, and providing input to an external examination program (e.g., school-based assessment with moderation and quality audit).

Suggested policy options:

1. Review existing system-level mechanisms for helping teachers develop expertise in classroom assessment to ensure that they are aligned with the national curriculum framework and the required uses of assessment data.

2. Clearly specify the required uses of classroom assessment information and develop detailed guidelines for carrying out classroom assessment activities that support student learning.

3. Communicate to key stakeholders about the mechanisms in place to monitor the quality of classroom assessment practices, and ensure that monitoring of classroom assessment activities takes place.

Examinations

Level of Development	Established
	●●●○

The *Higher Secondary Education Board Examination* program was authorized in 1989 and has been administered since 1992. The examination is formally used for high school graduation and university entrance, as well as informally for monitoring education quality levels, planning education policy reforms, ensuring that schools and educators accountable, and promoting competition among schools.

The *Higher Secondary Education Board Examination Manual 2010* document (original year of authorization, 1989), is certified by the Higher Secondary Education Board and authorizes the examination. This document is available to the public and outlines governance, distribution of power, and responsibilities among key entities. It also outlines procedures to investigate and address security breaches, cheating, or other forms of inappropriate behavior, as well as procedures for special and disadvantaged students.

The examination is funded entirely by student fees, which cover all core examination activities, but do not cover research and development. The government also allocates funds that could be used for the examination program, but it does not provide funding specifically for the examination.

While the examination office has state-of-the-art facilities, there is an insufficient number of qualified staff to adequately carry out examination activities. Nepal offers some opportunities to prepare for work on the examination, including university graduate programs and courses specifically focused on educational measurement and evaluation.

In general, the *Higher Secondary Education Board Examination* measures the national school curriculum guidelines and standards (which are based on internationally recognized guidelines and standards). Prior to the examination, students have access to examples of the types of questions that will be on the examination and general information on how to best prepare for the examination.

Teachers are involved in most examination-related tasks. Teachers select and create the examination questions and scoring guides, administer and score the examination, and supervise the examination procedures. There are up-to-date compulsory courses and workshops on the examinations offered to teachers.

Significant types of inappropriate behavior diminish the credibility of the examination. Such behavior includes the leakage of the content of an examination paper prior to the examination; impersonation (when an individual other than the registered candidate takes the examination); copying from other candidates; using unauthorized materials such as prepared answers and notes; collusion among candidates (via mobile phones, passing of paper, or equivalent); intimidation of examination supervisors, markers, or officials; issuing forged certificates or altering results information; and the provision of external assistance (via the supervisor, mobile phone, etc.). While inappropriate behavior surrounding the examination process is high, provisions to combat inappropriate behavior have been put in place.

Suggested policy options:

1. Introduce government funding to be used specifically for carrying out activities related to the examination, including research and development.

2. Build the capacity of individuals responsible for developing the examination questions by making available specific training opportunities in the area of test item design.

3. Develop and make available to all test administrators a manual that clearly outlines the process of test administration and the roles, responsibilities, and expectations of the test administrator.

National Large-Scale Assessment (NLSA)

Level of Development	Emerging ●●○○

The Ministry of Education is currently conducting the National Assessment of Student Achievement (NASA) at the eighth grade, which is the country's first nationally representative, large-scale assessment exercise. The students in the sample are being tested in Nepali, Mathematics, and Social Studies, and the results will be published by August 2012.

The NASA unit, which reports to the Education Review Office in the Ministry of Education, has been established as a permanent body responsible for carrying out NASA activities. In order to prepare for work on the NASA, Nepal offers university graduate programs, graduate and non-graduate university courses, and non-university training courses specifically focused on educational measurement and evaluation, as well as funding for attending international programs and workshops on educational measurement and evaluation.

Funding for NASA has been provided by the Government of Nepal, and covers all core, as well as research and development, activities. The government of Finland is also providing technical support for research and development.

NASA measures performance against national curriculum and is largely accepted by stakeholders. Mechanisms have been put in place to ensure that NASA accurately measures what it is supposed to measure. Such measures include regular independent review by qualified experts, regular internal review, and ad-hoc review of the alignment between the NASA and the national curriculum.

Nepal offers different options to include all groups of students in the NASA. Accommodations or alternative assessments are provided for students with disabilities, and special plans are made to ensure that NASA is administered to students in hard-to-reach areas. Additionally, NASA is offered in the language of instruction for almost all student groups.

In order to ensure the quality of the NASA instrument, all proctors and administrators are trained according to a set protocol, all discrepancies must be recorded on a standard sheet, data is double processed, and there are external reviewers and observers.

Nepal has also put in place various mechanisms to monitor the consequences of NASA, including providing funding for independent research on the impact of NASA, introducing a permanent oversight committee, regularly convening focus groups or surveys of key stakeholders, holding themed conferences that provide a forum to discuss research and other data on the consequences of NASA, and organizing expert review groups.

Suggested policy options:

1. Publicly assure the technical quality of NASA by making available a comprehensive, high-quality technical report on the assessment.

2. Facilitate effective uses of NASA results by making available a clearly written report that includes detailed information on overall student achievement levels and related factors, as well as common errors or weaknesses in student responses and how to address them.

3. Inform key stakeholders about NASA results in a timely manner by ensuring that results are widely disseminated within 12 months after the assessment is administered through, for example, briefings, workshops, and various forms of media (such as radio, television, and newspapers).

International Large-Scale Assessment (ILSA)

| Level of Development | Latent ●○○○ |

Nepal 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
	Absence of, or deviation from, the attribute	*On way to meeting minimum standard*	*Acceptable minimum standard*	*Best practice*
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 Nepal

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

NEPAL
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: *Setting clear guidelines for classroom assessment*			
There is no system-level document that provides guidelines for classroom assessment.	There is an informal system-level document that provides guidelines for classroom assessment.	There is a formal system-level document that provides guidelines for classroom assessment. [1] ＊	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. [2] ＊
ENABLING CONTEXT AND SYSTEM ALIGNMENT 2: *Aligning classroom assessment with system learning goals*			
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. [3] ＊
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. [4] ＊
ENABLING CONTEXT AND SYSTEM ALIGNMENT 3: *Having effective human resources to carry out classroom assessment activities*			
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. [5] ＊

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

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
ASSESSMENT QUALITY 1: *Ensuring the quality of classroom assessment*			
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.[6] *	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.[7] *
ASSESSMENT QUALITY 2: *Ensuring effective uses of classroom assessment*			
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.[8] *	Classroom assessment information is required to be disseminated to all key stakeholders.[8]
There are no required uses of classroom assessment to support student learning.	There are limited required uses of classroom assessment to support student learning.	There are adequate required uses of classroom assessment to support student learning, excluding its use as an input for external examination results.	There are adequate required uses of classroom assessment to support student learning, including its use as an input for external examination results.[9] *

Classroom Assessment: Development-level rating justifications

1. There is a formal, system level document that provides guidelines for classroom assessment. This document is called the "National Curriculum Framework for School Education" (2005).

2. The document is widely available. It is available online, in the public library, and through in-service teacher training.

3. There are a variety of system-wide resources that are typically available to teachers. These resources include a document that outlines what students are expected to learn in different subject areas at different grade/age levels, and a document that outlines the level(s) of performance that students are expected to reach in different subject areas at different grade/age levels. Resources also include textbooks or workbooks that provide support for classroom assessment, as well as scoring criteria or rubrics for students' work. Item banks or pools with examples of selection/multiple-choice or supply/open-ended questions are also made available to teachers.

4. There is an official curriculum or standards document that specifies what students are expected to learn and to what level of performance.

5. The system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment include pre-service and in-service teacher training and opportunities for teachers to participate in conferences and workshops. All teacher training programs also include a required course on classroom assessment. School inspection or teacher supervision includes a component that is focused on classroom assessment.

6. Classroom assessment activities are mainly about recalling information. It is also common to observe errors in the scoring or grading of students' work. Additionally, parents are poorly informed about students' grades, and teachers provide little useful feedback to students. Classroom assessment is mainly used as an administrative or a control tool, rather than a pedagogical resource, and it is not aligned with the pedagogical or curricular framework.

7. Various systematic mechanisms are in place to monitor the quality of classroom assessment activities. These mechanisms include that classroom assessment is a required component of a teacher's performance evaluation and of school inspection or teacher supervision. There are also system-wide reviews of the quality of education which include a focus on classroom assessment, and government funding is available for research on the quality of classroom assessment activities and on how to improve the quality of classroom assessment.

8. Teachers are required to report on individual student's performance to parents and students, but not to the school district or Ministry of Education officials.

9. Required uses of classroom assessment include diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child's learning, planning the next steps in instruction, grading students for internal classroom uses, and providing input to an external examination program (e.g., school-based assessment with moderation and quality audit).

NEPAL
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: *Setting clear policies*			
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: *Having strong leadership*			
All stakeholder groups strongly oppose the examination or are indifferent to it.	Most stakeholder groups oppose the examination. [5] ＊	Most stakeholders groups support the examination.	All stakeholder groups support the examination.
There are no attempts to improve the examination by stakeholder groups. [6] ＊	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. [7] ＊	This option does not apply to this dimension.

(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 3: *Having regular funding*			
There is no funding allocated for the examination.	There is irregular funding allocated for the examination.	There is regular funding allocated for the examination.[8] ✱	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. [9] ✱	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development. [10] ✱	This option does not apply to this dimension.	Funding covers research and development.
ENABLING CONTEXT 4: *Having strong organizational structures*			
The examination office does not exist or is newly established.	The examination office is newly established.	The examination office is a stable organization. [11] ✱	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. [12] ✱	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. [13] ✱

(CONTINUED)

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
ENABLING CONTEXT 5: *Having effective human resources*			
There is no staff to carry out the examination.	The examination office is inadequately staffed to effectively carry out the examination, issues are pervasive. [14] ✱	The examination office is adequately staffed to carry out the examination effectively, with minimal issues.	The examination office is adequately staffed to carry out the assessment effectively, with no issues.
The country does not offer opportunities that prepare for work on the examination.	This option does not apply to this dimension.	The country offers some opportunities that prepare for work on the examination. [15] ✱	The country offers a wide range of opportunities that prepare for work on the examination.

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

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●

SYSTEM ALIGNMENT 1: *Aligning examinations with learning goals and opportunities to learn*			
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.[16] ✱	This option does not apply to this dimension.
What the examination measures is questioned by some stakeholder groups.	This option does not apply to this dimension.	What is measured by the examination is largely accepted by stakeholder groups. ✱	This option does not apply to this dimension.
Material to prepare for the examination is minimal and it is only accessible to very few students.	There is some material to prepare for the examination that is accessible to some students. [17] ✱	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.
SYSTEM ALIGNMENT 2: *Providing teachers with opportunities to learn about the examination*			
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.	Teachers are involved in most examination-related tasks. [18] ✱

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: *Ensuring quality*			
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. [19] ✱	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.[20] ✱	There are varied and systematic mechanisms in place to ensure the quality of the examination.
ASSESSMENT QUALITY 2: *Ensuring fairness*			
Inappropriate behavior surrounding the examination process is high. [21] ✱	Inappropriate behavior surrounding the examination process is moderate.	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. ✱	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)

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
ASSESSMENT QUALITY 3: _Using examination information in a fair way_			
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: _Ensuring positive consequences of the examination_			
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. [22] ✻	There are some options for students who do not perform well on the examination.	There is a variety of options for students who do not perform well on the examination.
There are no mechanisms in place to monitor the consequences of the examination.	This option does not apply to this dimension.	There are some mechanisms in place to monitor the consequences of the examination. [23] ✻	There is a variety of mechanisms in place to monitor the consequences of the examination.

Examinations: Development-level rating justifications

1. The *Higher Secondary Education Board Examination* program was authorized in 1989 and has been operating since 1992. It is used formally for high school graduation and university entrance, as well as informally for monitoring education quality levels, planning education policy reforms, ensuring that schools and educators accountable, and promoting competition among schools.

2. The document is the Higher Secondary Education Board Examination Manual 2010, authorized by the Higher Secondary Education Board. The original year of authorization is 1989.

3. The policy document is available to the public.

4. The policy document outlines governance, distribution of power, and responsibilities among key entities. It also outlines procedures to investigate and address security breaches, cheating, or other forms of inappropriate behavior, as well as procedures for special/disadvantaged students. The policy document does not, however, describe the purpose of the examination or the authorized uses of results. It also does not specify who can sit for the examination. The policy document does not explain how the examination is aligned with curricula and standards, or the format of the examination questions.

5. Based on publicly available evidence, most of the stakeholder groups oppose the examination.

6. There have been no attempts to reform the examination. While stakeholder groups criticize the examination when the results are released, they do not take steps to reform the examination program.

7. Efforts to improve the examination are generally welcomed by the leadership in charge of the examination. For example, the leadership conducts workshops to gather suggestions from the different stakeholder groups.

8. Funding for the examination is provided through student fees. The government allocates funds that could be used for the examination, but it does not provide funding directly to the examination.

9. Funding also covers long- or medium-term planning of program milestones, as well as staff training.

10. Funding does not cover research and development.

11. The examination office has been operating since 1989.

12. The examination office is accountable to the Ministry of Education.

13. The examination office has computers for all technical staff, secure buildings and storage facilities, and adequate communication tools. The examination office generally has sufficient access to computer servers and the ability to back up data.

14. While the examination office is staffed with permanent or full-time staff, it is insufficient to meet the needs of the examination.

15. Opportunities include university graduate programs and courses that specifically focus on educational measurement and evaluation.

16. The examination measures the national school curriculum guidelines and standards (which are based on internationally recognized guidelines and standards).

17. Publicly available material on the examination includes examples of the types of questions that are on the examination and information on how to prepare for the examination.

18. Teachers mainly select or create examination questions and examination scoring guides, administer and score the examination, and supervise the examination procedures.

19. There is a formal technical report but it does not provide sufficient detail about the examination.

20. The systematic mechanisms that are in place include internal and external review or observers, as well as translation verification. However, there is no external certification or audit, pilot or field testing.

21. Significant types of inappropriate behavior diminish the credibility of the examination. Such behavior includes leakage of the content of an examination paper or part of a paper prior to the examination; impersonation when an individual other than the registered candidate takes the examination; copying from other candidates; using unauthorized materials such as prepared answers and notes; collusion among candidates via mobile phones, passing of paper, or equivalent; intimidation of examination supervisors, markers, or officials; issuing forged certificates or altering results information; and the provision of external assistance via the supervisor, mobile phone, etc. While inappropriate behavior surrounding the examination process is high, provisions to combat inappropriate behavior have been put in place.

22. Students can retake the examination only if they fail it.

23. Mechanisms to monitor the consequences of the examination include funding for independent research on the impact of the examination, existence of a permanent oversight committee, and regular focus groups or surveys of key stakeholders.

NEPAL
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: *Setting clear policies for NLSA*			
No NLSA exercise has taken place.	The NLSA has been operating on an irregular basis. [1]	The NLSA is a stable program that has been operating regularly.	This option does not apply to this dimension.
	✳		
There is no policy document pertaining to NLSA.	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.	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.	There is a written NLSA plan for the coming years.
			✳
ENABLING CONTEXT 2: *Having strong public engagement for NLSA*			
All stakeholder groups strongly oppose the NLSA or are indifferent to it. [2]	Some stakeholder groups oppose the NLSA.	Most stakeholders groups support the NLSA.	All stakeholder groups support the NLSA.

(CONTINUED)

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

(CONTINUED)

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
ENABLING CONTEXT 5: *Having effective human resources for NLSA*			
There is no staff allocated for running an NLSA.[6]	The NLSA office is inadequately staffed to effectively carry out the assessment.	The NLSA office is adequately staffed to carry out the NLSA effectively, with minimal issues.	The NLSA office is adequately staffed to carry out the NLSA effectively, with no issues.
The country does not offer opportunities that prepare individuals for work on NLSA.	This option does not apply to this dimension.	The country offers some opportunities to prepare individuals for work on the NLSA.	The country offers a wide range of opportunities to prepare individuals for work on the NLSA.[7] ✳

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

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
SYSTEM ALIGNMENT 1: *Aligning the NLSA with learning goals*			
It is not clear if the NLSA is based on curriculum or learning standards.	This option does not apply to this dimension.	The NLSA measures performance against curriculum or learning standards. ＊	This option does not apply to this dimension.
What the NLSA measures is generally questioned by stakeholder groups.	This option does not apply to this dimension.	What the NLSA measures is questioned by some stakeholder groups.	What the NLSA measures is largely accepted by stakeholder groups. [8] ＊
There are no mechanisms in place to ensure that the NLSA accurately measures what it is supposed to measure.	There are ad hoc reviews of the NLSA to ensure that it measures what it is intended to measure.	There are regular internal reviews of the NLSA to ensure that it measures what it is intended to measure. [9] ＊	This option does not apply to this dimension.
SYSTEM ALIGNMENT 2: *Providing teachers with opportunities to learn about the NLSA*			
There are no courses or workshops on the NLSA.	There are occasional courses or workshops on the NLSA. [10] ＊	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: *Ensuring the quality of the NLSA*			
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. [11] ✻
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. [12] ✻	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. [13]	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.	There is a comprehensive, high quality technical report available to the general public.
ASSESSMENT QUALITY 2: *Ensuring effective uses of the NLSA*			
NLSA results are not disseminated. [14]	NLSA results are poorly disseminated.	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. [15]	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.	There are a variety of mechanisms in place to monitor the consequences of the NLSA. [16] ✻

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

1. The National Assessment of Student Achievement (NASA) Program has been in place for less than five years.

2. Because NASA is being conducted for the first time, it is not yet possible to tell whether stakeholder groups support or oppose it.

3. There is regular funding for NASA that is allocated by the government of Nepal.

4. The NLSA office is called the National Assessment of Student Achievement (NASA) unit, and it reports to the Ministry of Education.

5. NASA is accountable to the Ministry of Education.

6. It is not possible to tell as the entire cycle of NASA has not yet been completed.

7. A variety of opportunities are available in Nepal on an annual basis to prepare individuals for work on the NLSA. These opportunities include university graduate programs, graduate and non-graduate university courses, and non-university training courses or workshops specifically focused on educational measurement and evaluation, as well as funding for attending international programs or courses, or workshops on educational measurement and evaluation.

8. Although NASA is in an early stage, what the NASA measures is largely accepted by stakeholders.

9. Mechanisms have been put in place to ensure that NASA accurately measures what it is supposed to measure, including having regular independent review by qualified experts, regular internal review, and ad-hoc review of the alignment between the assessment instrument and what it is supposed to measure.

10. There are occasional courses, workshops, and presentations that are offered to teachers on NASA.

11. Various options are offered to include all student groups in NASA. Accommodations or alternative assessments are provided for students with disabilities and special plans are made to ensure that NASA is administered to students in hard-to-reach areas. Additionally, NASA is offered in the language of instruction for almost all student groups.

12. Certain mechanisms are in place to ensure the quality of the NASA instrument. These mechanisms include that all proctors or administrators are trained according to a set protocol, the discrepancies must be recorded on a standard sheet, data is double processed, and there are external reviewers or observers.

13. It is not possible to tell as the entire cycle of NASA has not yet been completed.

14. It is not possible to tell as the entire cycle of NASA has not yet been completed.

15. It is not possible to tell as the entire cycle of NASA has not yet been completed.

16. Mechanisms in place to monitor the consequences of NASA include funding for independent research on the impact of NASA, a permanent oversight committee, regular focus groups or surveys of key stakeholders, themed conferences that provide a forum to discuss research and other data on the consequences of NASA, as well as expert review groups.

NEPAL
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. ✱	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.	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: *Having effective human resources for ILSA*			
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.

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

ASSESSMENT QUALITY

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

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
ASSESSMENT QUALITY 1: *Ensuring the quality of ILSA*			
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: *Ensuring effective uses of ILSA*			
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. Nepal has not participated in an ILSA, and it does not have plans to do so in the near future.

Acknowledgements

References

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

World Bank. 2012. Nepal Country Indicator Data. Washington, DC: World Bank. Data retrieved from http://data.worldbank.org/ on February 27, 2014.

World Bank. 2013. *Nepal- Country Program Snapshot July 2013*. Kathmandu, Nepal: World Bank.

World Bank. 2013. *Nepal: School Sector Reform Program: Additional Financing Project Information Document*. Report No. AB7049. 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.

READ
Russia Education Aid for Development
When children learn, nations prosper

THE WORLD BANK