



Key Policy Areas for Student Assessment

1. Classroom Assessment

In Lebanon, a formal, publicly available system-level document provides guidelines for classroom assessment. Some system-wide resources and materials (such as scoring criteria for evaluating students' work) are also available to teachers for carrying out classroom assessment activities. In Lebanon, classroom assessment activities tend to be aligned with the pedagogical and curricular framework and are considered to be of moderate quality. While there are adequate required uses of classroom assessment information to support student learning, the information is not used as an input to external examination programs.

Status

Established



2. Examinations

The General Secondary Diploma examination has been administered to grade 12 students since 1925. The results are used to certify student completion of the school cycle and to determine selection to higher-education institutions. The results are officially recognized by certification and selection systems in Lebanon and abroad. Regular funding for the examination is provided by the government and covers all core examination activities. There are limited systematic mechanisms in place to ensure the quality of the examination. Additionally, there are no mechanisms, such as a permanent oversight committee, in place to monitor the consequences of the examination for students and other stakeholder groups.

Established



3. National Large-Scale Assessment (NLSA)

The Measuring Learning Achievement study was administered for the first time in 1994. Since then, it has been operating on an irregular basis, each time assessing a different grade and set of subjects. Lebanon does not have a policy document on the NLSA, or a plan for future NLSA activities. Funding for NLSA activities, provided by the Center for Educational Research and Development, covers core NLSA activities as well as research and development. While there are some mechanisms in place to ensure the quality of the NLSA, there is no publicly available technical report or similar document. There also are no mechanisms in place to monitor the consequences of the NLSA in terms of its impact on education quality or student learning.

Emerging



4. International Large-Scale Assessment (ILSA)

In the last 10 years, Lebanon has participated in TIMSS (2003, 2007, 2011) and PASEC (2009), and has taken concrete steps to participate in PISA 2015 and TIMSS 2015. Funding for ILSAs is provided by the Centre for Educational Research and Development and donors. The ILSA team typically attends international meetings on ILSAs; however, no opportunities to learn about ILSAs are offered in Lebanon. Lebanon-specific ILSA results have not been disseminated in the country, and it is not clear that decisions based on ILSA results have had a positive impact on achievement levels.

Emerging



Introduction

Lebanon 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, Lebanon decided to benchmark this system using standardized tools developed under The World Bank's Systems Approach for Better Education Results (SABER) program. SABER is an evidence-based program to help countries systematically examine and strengthen the performance of different aspects of their education systems.

What is SABER-Student Assessment?

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

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

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

SABER-Student Assessment methodology

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

systems: the types/purposes of assessment activities and the quality of those activities.

Assessment types and purposes

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

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

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

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

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

Quality drivers of an assessment system

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

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

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

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

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

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

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

Source: World Bank.

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

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

Levels of development

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

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

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

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

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

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

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

Education in Lebanon

Lebanon is an upper-middle-income country in the Middle East and North Africa region. GDP per capita

(current US\$) is \$9,705, with annual growth of 1.4 percent in 2012.

The Lebanese education system is comprised of kindergarten (3- to 6-year-olds); elementary education from grades 1 through 6; intermediate education for grades 7 through 9; and secondary education for grades 10 through 12. Lebanon invests heavily in education, committing 13 percent of its GDP for this purpose. This strong commitment is reflected in Lebanon's primary net enrollment rate (95 percent), which is above the regional average. Although Lebanon has low pupil/teacher ratios of 14:1 at the primary level and 8:1 at the secondary level, there is a lack of qualified teachers (only 4 percent of public school teachers hold a specialized degree), and little or no in-service training is offered to new teachers.

The Ministry of Education has made important efforts to improve its education system. The government's priorities for education include expanding access to early childhood education; reducing dropout and repetition rates; enhancing the qualifications and performance of the teaching workforce; establishing a National Qualifications Framework that will link the various education levels to the labor market; and introducing quality assurance at the higher-education level.

Detailed information was collected on Lebanon's student assessment system using the SABER-Student Assessment questionnaires and rubrics. 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 Lebanon, 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 Lebanon's immediate interests and needs. Detailed, completed rubrics for each assessment type are provided in Appendix 5.

Classroom Assessment

Level of development: ESTABLISHED

In Lebanon, a formal system-level document provides guidelines for classroom assessment for primary and secondary education. The document is available to the public and during in-service courses for teachers. The document is also distributed to all school principals.

There are some system-wide resources available to teachers for carrying out classroom assessment activities. These include an official curriculum/learning standards document that outlines what students are expected to learn and their expected levels of performance in different subject areas at different grade levels; textbooks or workbooks that provide support for classroom assessment; scoring criteria or rubrics for students' work; and explanations about the school exams.

There are limited opportunities available to primary and secondary school teachers for learning about or developing more effective classroom assessment practices. The existing opportunities are provided during pre-service teacher-training programs in institutes and universities, and during in-service teacher training programs offered by the Centre for Educational Research and Development (CERD). There are also opportunities for teachers to participate in conferences and workshops, and school inspection or teacher supervision includes a component focused on classroom assessment. In addition, teachers have opportunities to participate in item development for, or scoring of, large-scale assessments or exams.

Classroom assessment practices are perceived as being of moderate quality. While these practices tend to be aligned with the pedagogical and curricular framework, it is also common for them to be focused primarily on information recall. Apart from classroom assessments being a component of school inspection or teacher supervision, there are ad hoc mechanisms in place to monitor the quality of classroom-related practices. In addition, government funding is available for research on the quality of classroom assessment activities and how to improve classroom assessment with the CERD and the Faculty of Education at the Lebanese University.

Information from classroom assessment is required to be made available to students, parents, and school directors. Teachers provide feedback in an assessment grid, which allows space for quantitative as well as qualitative input.

There is no requirement for classroom assessment data to be used for grading students or as an input to an external examination program (e.g., school-based assessment with moderation and quality audit).

Suggested policy options:

1. The Program "Assessment of Student Achievement and Curriculum Development," which is part of the Education Sector Development Plan (2010-2015), is to be implemented with the cooperation of CERD and the Faculty of Education of the Lebanese University. This program aims to develop the curriculum in a manner consistent with national needs and global trends, through setting clear foundations for the reform of the curriculum and the identification of learning outcomes, developing a Question Bank System (QBS) to be used for examinations conducted at the school level, and developing standards for the production of textbooks and teacher guides. In implementing this program, ensure participation and full involvement of the staff of the regional resource centers for teacher professional development with the selected school principals, as well as teacher trainers from universities outside of the Lebanese University. In addition, ensure that resources developed in this program to support teachers in classroom assessment practices are widely disseminated.

2. Introduce a variety of mechanisms to systematically monitor the quality of classroom assessment practices. For example, include classroom assessment as a component of a teacher's performance evaluation.

3. Ensure that classroom assessment information is effectively used to inform policymaking by requiring that it be disseminated to all key stakeholders, including district or Ministry of Education officials.

Examinations

Level of development: ESTABLISHED

The *General Secondary Diploma (Baccalaureate)* examination has been administered to grade 12 students since 1925. The results are used for certifying student completion of grade 12 and for determining selection to university or other higher-education institutions.

There is a policy document authorizing and addressing some key aspects of the *Baccalaureate*. However, the document does not describe authorized uses of the examination results.

While policymakers and educators support the examination, it is unclear to what extent other stakeholder groups (teacher unions, students, parents, media, think tanks and NGOs, employers, and universities) support it. The Ministry of Education's Directorate General has coordinated efforts with educators from all sectors to improve the *Baccalaureate*. These efforts are generally welcomed by the leadership in charge of the examination.

The Ministry of Education and Higher Education's Department of Examinations has been running the *Baccalaureate* since 1949. Funding for the examination is allocated by the government and covers all core examination activities (design, administration, data processing, and reporting), but does not cover research and development. The *Baccalaureate* office has some of the required facilities to carry out the examination and is adequately staffed in terms of the quality (qualifications and knowledge) of those working on the examination, but not necessarily in terms of their quantity (number of staff).

The Central Inspectorate and the General Directorate of General Education is the external body to which the Examination Department is accountable. The examination results are officially recognized by certification and selection systems in Lebanon and abroad.

Lebanon offers some opportunities that prepare individuals for work on the *Baccalaureate*. These

opportunities include university graduate programs (masters or doctoral level) and non-university training courses or workshops focused on educational measurement and evaluation.

There is a clear understanding and acceptance that the *Baccalaureate* measures national curricula guidelines. Teachers have the opportunity to learn about the *Baccalaureate* through regularly updated, voluntary courses or workshops that are offered as part of the in-service training provided by the Center for Educational Research and Development. Teachers are involved in most examination-related tasks, including selecting or creating examination questions and scoring guides. They also administer, supervise, and score the examination.

Baccalaureate results are credible for all stakeholder groups, and there is no systematic evidence of their improper use. However, students' names and results are not confidential and are made public on the website of the Ministry of Education.

Suggested policy options:

1. In order to create linkages between curriculum reform led by the CERD and the formal examination system, introduce mechanisms, such as expert review groups, which monitor the consequences of the examination and inform curriculum reform activities. Ensure that the administrative procedures within the Department of Examinations at the Ministry of Education and Higher Education are streamlined to allow for effective implementation of such linkages.
2. Ensure that the Department of Examinations has all of the required facilities to carry out the examination effectively. For example, ensure that all technical staff have access to computers.
3. Support system-level mechanisms to ensure the quality of the examination by, for example, conducting a pilot as well as by continuing the ongoing work to ensure expansion and improvement of the Question Bank System (QBS).

National Large-Scale Assessment (NLSA)

Level of development: EMERGING

The *Measuring Learning Achievement* study is a national large-scale assessment that was administered for the first time in 1994. Since then, it has been implemented on an irregular basis (1995, 1997, 2003, and 2012), each time targeting a different grade level and different set of subjects.

There is no policy document, technical report, or other available documentation on the *Measuring Learning Achievement* study. In order to conduct a round of the assessment, the Centre for Educational Research and Development (CERD) must submit a proposal to, and receive approval from, the Minister of Education.

The office responsible for the *Measuring Learning Achievement* study is adequately staffed with mostly part-time and temporary staff. Funding for the assessment is irregular and allocated through the budget of the Center for Educational Research and Development. Funding covers all core assessment activities (design, administration, analysis, and reporting), as well as staff training and research and development activities. The *Measuring Learning Achievement* office is accountable to the Central Inspectorate, which is a monitoring body, and to a special studies committee mandated by the Center for Educational Research and Development.

As a way to prepare individuals for work on the *Measuring Learning Achievement* study, Lebanon offers funding opportunities for attending international programs, courses, or workshops on educational measurement and evaluation.

There are some mechanisms in place to ensure the quality of the *Measuring Learning Achievement* study. However, discrepancies are not recorded on a standard sheet, and there is no double data scoring, no external reviewers or observers, and no external certification or audit. There are options to include all student groups in the study, such as plans to ensure that the assessment is administered to students in hard-to-reach areas and in the language of instruction for almost all student groups.

Reports on the *Measuring Learning Achievement* results, which contain information on overall achievement levels and subgroups, are available for all stakeholder groups. There also are workshops or presentations for key stakeholders on the results. However, the reports do not contain information on trends over time, nor do they provide information on standard errors. Additionally, there is no media briefing or coverage of the *Measuring Learning Achievement* results.

Measuring Learning Achievement data tend not to be used or else are used in ways inconsistent with the technical characteristics of the assessment.

Suggested policy options:

1. In order to institutionalize the *Measuring Learning Achievement* assessment, introduce a policy document pertaining to key aspects of the NLSA (sources of funding, uses of results); ensure that the policy document is approved by the relevant authorizing body and is publicly available.
2. Develop a written plan for future NLSA activities, which includes a strategy for collaboration among the DNE (Dispositif National d'Evaluation - National Evaluation Body), the CERD Evaluation and Assessment Unit, and the Faculty of Education at the Lebanese University.
3. Ensure that NLSA results are effectively disseminated in Lebanon. For example, ensure that reports with results include information on trends over time overall and for subgroups, and that the media provides coverage of NLSA results.

International Large-Scale Assessment (ILSA)

Level of development:

In the last 10 years, Lebanon has participated in a number of ILSA exercises, including PASEC (2009) and TIMSS (2003, 2007, and 2011). The formal policy document that addresses participation in ILSAs is the *Letter from the Head of the Center for Educational Research and Development*, authorized by the Ministry of Education and Higher Education on February 2, 2011. Lebanon has taken concrete steps to participate in PISA 2015 and TIMSS 2015.

There is funding for participation in ILSAs allocated at discretion by the Centre for Educational Research and Development (CERD) in Lebanon and from loans or external donors. Funding covers international participation fees, in-country implementation of the assessment exercise (e.g., printing booklets, travel to schools), data processing and analysis, and attendance at international meetings for the assessment exercise. Funding does not cover reporting and dissemination of the assessment results in Lebanon or research and development activities.

The national coordinator oversees a sufficiently staffed team that has previous experience working on international assessments as well as the necessary training to carry out the required ILSA activities effectively. The national coordinator is fluent in the language (English, in the case of TIMSS) in which the international-level meetings are conducted and related documentation is available.

No issues were identified with the carrying out of the most recent ILSA activities in Lebanon. In addition, the ILSA team was able to attend all international workshops and meetings related to the exercise. However, Lebanon only met the standards necessary to have its data presented beneath the main display of the international report or in an annex.

Country-specific results and information from ILSAs have not been disseminated in Lebanon and have only been published in the international report. Products to provide feedback to schools and educators about the ILSA results are not made available. While it is planned for ILSA results to be used to inform curriculum improvement, teacher-training programs, and other assessment activities in the system, at the present time, results from ILSAs have not been used to inform decision making at the national level.

It is not clear that decisions based on ILSA results have had a positive impact on student achievement levels.

Suggested policy options:

1. Continue efforts to strengthen capacity within the CERD, which include Lebanon's involvement in the Regional Program on Student Assessment and Quality of Education, part of the Arab Regional Agenda for Improving Education Quality (ARAIEQ). In addition, introduce opportunities to learn about ILSA to relevant stakeholders. For example, host workshops on using international assessment databases that are available to staff working on ILSA and other individuals interested in assessment.
2. Ensure that Lebanon-specific ILSA results are widely disseminated in the country. For example, develop a national report and make it available to all key stakeholders, including school principals, teachers (through teacher-training programs), and the media.
3. Organize regular conferences and workshops involving a variety of key stakeholder groups—including academia, practitioners, media, and civil society organizations—to discuss how ILSA results can be used to improve learning in Lebanon.
4. Continue ongoing analytical work on TIMSS data and ensure that it is used for informing policymaking, such as curriculum improvement and teacher-training programs.

Appendix 1: Assessment Types and Their Key Differences

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

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

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

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

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

Appendix 4: Methodology for Assigning Development Levels

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

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

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

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

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

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

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

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

6. Scores for certain key dimensions under 'Enabling Context' (in the case of EXAM, NLSA, and ILSA) and under 'System Alignment' (in the case of CLASS) were set as ceiling scores, i.e., the overall mean score for the

particular assessment type cannot be greater than the score for these key dimensions. These key variables include formal policy, regular funding, having a permanent assessment unit, and the quality of assessment practices.

Appendix 5: SABER-Student Assessment Rubrics for Lebanon

This appendix provides the completed SABER-Student Assessment rubrics for each type of assessment activity in Lebanon. 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 of the selection (as indicated by a thick border and an asterisk), which is provided 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, such superscript indicates that insufficient information was available to determine the relevant selection in the row.

LEBANON

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.

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ENABLING CONTEXT AND SYSTEM ALIGNMENT 1: <i>Setting clear guidelines for classroom assessment</i></p>			
<p>There is no system-level document that provides guidelines for classroom assessment.</p>	<p>There is an informal system-level document that provides guidelines for classroom assessment.</p>	<p>There is a formal system-level document that provides guidelines for classroom assessment.¹ *</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>This option does not apply to this dimension.</p>	<p>The availability of the document is restricted.</p>	<p>The document is widely available.² *</p>
<p>ENABLING CONTEXT AND SYSTEM ALIGNMENT 2: <i>Aligning classroom assessment with system learning goals</i></p>			
<p>There are no system-wide resources for teachers for classroom assessment.</p>	<p>There are scarce system-wide resources for teachers for classroom assessment.</p>	<p>There are some system-wide resources for teachers for classroom assessment.³ *</p>	<p>There are a variety of system-wide resources available for teachers for classroom assessment.</p>
<p>There is no official curriculum or standards document.</p>	<p>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.</p>	<p>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.</p>	<p>There is an official curriculum or standards document that specifies what students are expected to learn and to what level of performance.⁴ *</p>
<p>ENABLING CONTEXT AND SYSTEM ALIGNMENT 3: <i>Having effective human resources to carry out classroom assessment activities</i></p>			
<p>There are no system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.</p>	<p>This option does not apply to this dimension.</p>	<p>There are some system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.</p>	<p>There are a variety of system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.⁵ *</p>

ASSESSMENT QUALITY

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

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ASSESSMENT QUALITY 1: <i>Ensuring the quality of classroom assessment</i></p>			
<p>Classroom assessment practices suffer from widespread weaknesses or there is no information available on classroom assessment practices.</p>	<p>Classroom assessment practices are known to be weak.</p>	<p>Classroom assessment practices are known to be of moderate quality.⁶</p>	<p>Classroom assessment practices are known to be generally of high quality.</p>
<p>There are no mechanisms to monitor the quality of classroom assessment practices.</p>	<p>There are ad hoc mechanisms to monitor the quality of classroom assessment practices.⁷ *</p>	<p>There are limited systematic mechanisms to monitor the quality of classroom assessment practices.</p>	<p>There are varied and systematic mechanisms in place to monitor the quality of classroom assessment practices.</p>
<p>ASSESSMENT QUALITY 2: <i>Ensuring effective uses of classroom assessment</i></p>			
<p>Classroom assessment information is not required to be disseminated to key stakeholders.</p>	<p>This option does not apply to this dimension.</p>	<p>Classroom assessment information is required to be disseminated to some key stakeholders.⁸ *</p>	<p>Classroom assessment information is required to be disseminated to all key stakeholders.</p>
<p>There are no required uses of classroom assessment to support student learning.</p>	<p>There are limited required uses of classroom assessment to support student learning.</p>	<p>There are adequate required uses of classroom assessment to support student learning, excluding its use as an input for external examination results.⁹ *</p>	<p>There are adequate required uses of classroom assessment to support student learning, including its use as an input for external examination results.</p>

Classroom Assessment: Development-level rating justifications

1. The formal system-level document providing guidelines for classroom assessment is the Minister decision No. 666/m/2000, dated November 14, 2000, and its amendment by decision No. 940/m/2001, dated June 27, 2001, authorized by the Ministry of Education. The document applies for primary and secondary education.
2. The document is available to the public and has been distributed to all school principals in the public sector. The document is also available during in-service courses for teachers. Additionally, teachers have had the opportunity to review the document in the early phase of implementation of the new curricula in 2000 and 2001.
3. Resources teachers for classroom assessment include:
 - a. A document that outlines what students are expected to learn in different subject areas at different grade/age levels and the level(s) of performance that students are expected to reach in different subject areas at different grade/age levels (the document can be located in the Centre for Educational Research and Development's website at www.crdp.org);
 - b. Textbooks or workbooks that provide support for classroom assessment;
 - c. Scoring criteria or rubrics for students' work; and
 - d. Explanations about the school exams.
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 that ensure that teachers develop skills and expertise in classroom assessment include pre-service teacher training opportunities available in the institutes and in the universities, in-service teacher training provided continuously by the Centre for Educational Research and Development for all primary and secondary school teachers, opportunities for teachers to participate in conferences and workshops and in item development for, or scoring of, large-scale assessments or exams, and school inspection or teacher supervision includes a component focused on classroom assessment. Pre-service teacher training programs include courses on classroom assessment.
6. Classroom assessment practices are known to be of moderate quality. Specifically, it is not common for classroom assessment activities to rely mainly on multiple-choice, selection-type questions, to observe errors in the scoring or grading of students' work, or for classroom assessments to be mainly used as administrative or control tool rather than as pedagogical resource. It is common for classroom assessment practices to be aligned with the pedagogical and curricular framework. Uneven application of standards for grading students' work and grade inflation tend to not be a serious problem. Even though parents tend to be well informed of students' grades, classroom assessment information provides little feedback to students. Additionally, it is common for classroom assessment activities to be mainly about recalling information, and teachers tend to not use explicit or a priori criteria for scoring or grading students' work.

7. Although classroom assessment is a required component of school inspection or teacher supervision, classroom assessment is not a required component of a teacher's performance evaluation; there is no external moderation system in place that reviews the difficulty of classroom assessment activities or the appropriateness of scoring criteria in grading students' work; and national or other system-wide reviews of the quality of education that include a focus on classroom assessment do not take place. However, government funding is available for research on the quality of classroom assessment activities and how to improve classroom assessment with the CERD and through the Faculty of Education at the Lebanese University.
8. Teachers are required to report on individual student's performance to students, parents, and school directors. Teachers provide feedback in an assessment grid, which allows space for quantitative as well as qualitative input.
9. The required uses include diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child's learning, and planning next steps in instruction. However, classroom assessment is not required to be used for grading students for internal classroom uses or for providing input to an external examination program (e.g., school-based assessment with moderation and quality audit).

LEBANON

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.

<p>LATENT</p> 	<p>EMERGING</p> 	<p>ESTABLISHED</p> 	<p>ADVANCED</p> 
<p>ENABLING CONTEXT 1: <i>Setting clear policies</i></p>			
<p>No standardized examination has taken place.</p>	<p>The standardized examination has been operating on an irregular basis.</p>	<p>The examination is a stable program that has been operating regularly.¹ *</p>	<p>This option does not apply to this dimension</p>
<p>There is no policy document that authorizes the examination.</p>	<p>There is an informal or draft policy document that authorizes the examination.</p>	<p>There is a formal policy document that authorizes the examination.² *</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>The policy document is not available to the public</p>	<p>The policy document is available to the public.³</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>This option does not apply to this dimension.</p>	<p>The policy document addresses some key aspects of the examination.⁴ *</p>	<p>The policy document addresses all key aspects of the examination.</p>
<p>ENABLING CONTEXT 2: <i>Having strong leadership</i></p>			
<p>All stakeholder groups strongly oppose the examination or are indifferent to it.⁵</p>	<p>Most stakeholder groups oppose the examination.⁵</p>	<p>Most stakeholders groups support the examination.⁵</p>	<p>All stakeholder groups support the examination.⁵</p>
<p>There are no attempts to improve the examination by stakeholder groups.</p>	<p>This option does not apply to this dimension.</p>	<p>There are independent attempts to improve the examination by stakeholder groups.</p>	<p>There are coordinated attempts to improve the examination by stakeholder groups.⁶ *</p>
<p>Efforts to improve the examination are not welcomed by the leadership in charge of the examination</p>	<p>This option does not apply to this dimension.</p>	<p>Efforts to improve the examination are generally welcomed by the leadership in charge of the examination.⁷ *</p>	<p>This option does not apply to this dimension.</p>

(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. 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: <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. 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. ¹³ *
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. ¹⁴ *	The examination office has all of the required facilities to carry out the examination.	The examination office has state of the art facilities to carry out the examination.

(CONTINUED)

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT 5: <i>Having effective human resources</i>			
There is no staff to carry out the examination.	The examination office is inadequately staffed to effectively carry out the examination, issues are pervasive.	The examination office is adequately staffed to carry out the examination effectively, with minimal issues. ¹⁵ *	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. ¹⁶ *	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.

<p>LATENT</p> 	<p>EMERGING</p> 	<p>ESTABLISHED</p> 	<p>ADVANCED</p> 
--	--	---	--

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. ¹⁷ *	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. ¹⁸ *	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. ¹⁹ *	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. ²⁰	There are up-to-date voluntary courses or workshops on examinations available to teachers. ²⁰ *	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. ²¹ *

ASSESSMENT QUALITY

Degree to which the assessment meets 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>ASSESSMENT QUALITY 1: <i>Ensuring quality</i></p>			
<p>There is no technical report or other documentation.</p>	<p>There is some documentation on the examination, but it is not in a formal report format.²² *</p>	<p>There is a comprehensive technical report but with restricted circulation.</p>	<p>There is a comprehensive, high quality technical report available to the general public.</p>
<p>There are no mechanisms in place to ensure the quality of the examination.</p>	<p>This option does not apply to this dimension.</p>	<p>There are limited systematic mechanisms in place to ensure the quality of the examination.²³ *</p>	<p>There are varied and systematic mechanisms in place to ensure the quality of the examination.</p>
<p>ASSESSMENT QUALITY 2: <i>Ensuring fairness</i></p>			
<p>Inappropriate behavior surrounding the examination process is high.</p>	<p>Inappropriate behavior surrounding the examination process is moderate.</p>	<p>Inappropriate behavior surrounding the examination process is low.²⁴ *</p>	<p>Inappropriate behavior surrounding the examination process is marginal.</p>
<p>The examination results lack credibility for all stakeholder groups.</p>	<p>The examination results are credible for some stakeholder groups.</p>	<p>The examination results are credible for all stakeholder groups.²⁵ *</p>	<p>This option does not apply to this dimension.</p>
<p>The majority of the students (over 50%) may not take the examination because of language, gender, or other equivalent barriers.</p>	<p>A significant proportion of students (10%-50%) may not take the examination because of language, gender, or other equivalent barriers.</p>	<p>A small proportion of students (less than 10%) may not take the examination because of language, gender, or other equivalent barriers.</p>	<p>All students can take the examination; there are no language, gender or other equivalent barriers.²⁶ *</p>

(CONTINUED)

<p>LATENT</p> 	<p>EMERGING</p> 	<p>ESTABLISHED</p> 	<p>ADVANCED</p> 
<p>ASSESSMENT QUALITY 3: <i>Using examination information in a fair way</i></p>			
<p>Examination results are not used in a proper way by all stakeholder groups.</p>	<p>Examination results are used by some stakeholder groups in a proper way.</p>	<p>Examination results are used by most stakeholder groups in a proper way.</p>	<p>Examination results are used by all stakeholder groups in a proper way.²⁷</p>
<p>Student names and results are public.²⁸</p> <p>*</p>	<p>This option does not apply to this dimension.</p>	<p>Students' results are confidential.</p>	<p>This option does not apply to this dimension.</p> <p>*</p>
<p>ASSESSMENT QUALITY 4: <i>Ensuring positive consequences of the examination</i></p>			
<p>There are no options for students who do not perform well on the examination, or students must leave the education system.</p>	<p>There are very limited options for students who do not perform well on the examination.²⁹</p> <p>*</p>	<p>There are some options for students who do not perform well on the examination.</p>	<p>There is a variety of options for students who do not perform well on the examination.</p>
<p>There are no mechanisms in place to monitor the consequences of the examination.³⁰</p> <p>*</p>	<p>This option does not apply to this dimension.</p>	<p>There are some mechanisms in place to monitor the consequences of the examination.</p>	<p>There is a variety of mechanisms in place to monitor the consequences of the examination.</p>

Examinations: [Development-level rating justifications](#)

1. The examination, *General Secondary Diploma (Baccalaureate)*, has been administered since 1925 at grade 12 for student certification for grade completion and for selection to university or other higher-education institutions.
2. The formal policy document that authorizes the examination is Decree No. 5697 of June 15, 2001, authorized by the Presidency of the Republic in 2011.
3. The policy document is available to the public.
4. The policy document authorizing the examination outlines governance; distribution of power; responsibilities among key entities; procedures to investigate and address security breaches, cheating, or other forms of inappropriate behavior; and procedures for special/disadvantaged students. It also specifies who can sit for the examination and identifies rules about its preparation. The document also describes the purpose of the examination, states funding sources, and explains its alignment with curricula and standards and the format of the examination questions. However, the policy document does not describe authorized uses of results.
5. While policymakers and educators support the examination program, it is unclear to what extent other stakeholder groups (teacher unions, students, parents, media, think tanks and NGOs, employers, and universities) support it.
6. The Ministry of Education's Directorate General has coordinated efforts with educators from all sectors to improve the examination.
7. Efforts to improve the examination are generally welcomed by the leadership in charge of the examination.
8. There is regular funding allocated by the government for the examination.
9. Funding covers all core examination activities: examination design and administration, and data processing and reporting.
10. Funding does not cover research and development activities.
11. The Ministry of Education's Examination Department has been running the examination since 1949.
12. The Central Inspectorate and the General Directorate of General Education is the external body to which the Ministry's Examination Department is accountable to.
13. The examination results are officially recognized by certification and selection systems in the country and anywhere abroad. Lebanese students with Lebanese certification can access foreign universities in any country.

14. While the examination office usually has a secure building, adequate communication tools (phone, email, and internet), and secure storage facilities, it does not normally have computers for all technical staff, access to adequate computer servers, or the ability to back up data.
15. The examination office is adequately staffed in terms of quality, but not necessarily in terms of quantity. Mainly temporary and part time staff are responsible for running the examination.
16. The opportunities offered on an annual basis to prepare for work on the examination are university graduate programs (masters or doctoral level), as well as non-university training courses or workshops, specifically focused on educational measurement and evaluation.
17. The examination measures national school curriculum guidelines or standards.
18. What is measured by the examination is largely accepted by stakeholder groups.
19. The material available to prepare for the examination includes examples of the types of questions that are on the examination and information on how to prepare for the examination. This material is accessible by all students.
20. There are voluntary courses or workshops that are regularly updated and are offered through the continuous teacher in-service training provided by the Center for Educational Research and Development.
21. Teachers have the tasks of selecting or creating examination questions and scoring guides. They also administer, supervise, and score the examination.
22. There is some documentation on the examination, but it is not in a formal report format.
23. Despite there being internal and external reviewers or observers to ensure the quality of the examination, there is no external certification or audit, pilot or field testing, or translation verification.
24. The few inappropriate behaviors that typically occur are impersonation (when an individual other than the registered candidate takes the examination) and copying from other candidates. In the event of these conducts, students are forbidden to be admitted into the examination process for four consecutive terms.
25. The examination results are credible for all stakeholder groups.
26. All students can take the examination; there is no language, gender, or other equivalent barriers.
27. There is no systematic evidence of improper use of the examination results.

28. Student's names and results are not confidential and are made public on the website of the Ministry of Education.
29. The options for the students who do not perform well on the examination are retaking the examination or repeating the grade.
30. There are no mechanisms in place to monitor the consequences of the examination.

LEBANON
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.

<p>LATENT</p> 	<p>EMERGING</p> 	<p>ESTABLISHED</p> 	<p>ADVANCED</p> 
<p>ENABLING CONTEXT 1: <i>Setting clear policies for NLSA</i></p>			
<p>No NLSA exercise has taken place.</p>	<p>The NLSA has been operating on an irregular basis.¹</p>	<p>The NLSA is a stable program that has been operating regularly.</p>	<p>This option does not apply to this dimension.</p>
<p>There is no policy document pertaining to NLSA.²</p>	<p>There is an informal or draft policy document that authorizes the NLSA.</p>	<p>There is a formal policy document that authorizes the NLSA.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.³</p>	<p>The policy document is not available to the public.</p>	<p>The policy document is available to the public.</p>	<p>This option does not apply to this dimension.</p>
<p>There is no plan for NLSA activity.⁴</p>	<p>This option does not apply to this dimension.</p>	<p>There is a general understanding that the NLSA will take place.</p>	<p>There is a written NLSA plan for the coming years.</p>
<p>ENABLING CONTEXT 2: <i>Having strong public engagement for NLSA</i></p>			
<p>All stakeholder groups strongly oppose the NLSA or are indifferent to it.</p>	<p>Some stakeholder groups oppose the NLSA.</p>	<p>Most stakeholders groups support the NLSA.⁵</p>	<p>All stakeholder groups support the NLSA.</p>

(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. ⁷ *	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. ⁸ *
ENABLING CONTEXT 4: <i>Having strong organizational structures for NLSA</i>			
There is no NLSA office, ad hoc unit or team.	The NLSA office is a temporary agency or group of people. 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.	The NLSA office is accountable to a clearly recognized body. ¹¹ *	This option does not apply to this dimension.

(CONTINUED)

<p>LATENT</p> 	<p>EMERGING</p> 	<p>ESTABLISHED</p> 	<p>ADVANCED</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.</p>	<p>The NLSA office is adequately staffed to carry out the NLSA effectively, with minimal issues.¹² *</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 ● ● ● ●
<p align="center">SYSTEM ALIGNMENT 1: <i>Aligning the NLSA with learning goals</i></p>			
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.
<p align="center">SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the NLSA</i></p>			
There are no courses or workshops on the NLSA.	There are occasional courses or workshops on the NLSA. ¹⁷ *	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.

<p>LATENT</p> 	<p>EMERGING</p> 	<p>ESTABLISHED</p> 	<p>ADVANCED</p> 
<p>ASSESSMENT QUALITY 1: <i>Ensuring the quality of the NLSA</i></p>			
<p>No options are offered to include all groups of students in the NLSA.</p>	<p>This option does not apply to this dimension.</p>	<p>At least one option is offered to include all groups of students in the NLSA.¹⁸ *</p>	<p>Different options are offered to include all groups of students in the NLSA.</p>
<p>There are no mechanisms in place to ensure the quality of the NLSA.</p>	<p>This option does not apply to this dimension.</p>	<p>There are some mechanisms in place to ensure the quality of the NLSA.¹⁹ *</p>	<p>There are a variety of mechanisms in place to ensure the quality of the NLSA.</p>
<p>There is no technical report or other documentation about the NLSA.²⁰ *</p>	<p>There is some documentation about the technical aspects of the NLSA, but it is not in a formal report format.</p>	<p>There is a comprehensive technical report but with restricted circulation.</p>	<p>There is a comprehensive, high quality technical report available to the general public.</p>
<p>ASSESSMENT QUALITY 2: <i>Ensuring effective uses of the NLSA</i></p>			
<p>NLSA results are not disseminated.</p>	<p>NLSA results are poorly disseminated.²¹ *</p>	<p>NLSA results are disseminated in an effective way.</p>	<p>This option does not apply to this dimension.</p>
<p>NLSA information is not used or is used in ways inconsistent with the purposes or the technical characteristics of the assessment.²² *</p>	<p>This option does not apply to this dimension.</p>	<p>NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.</p>	<p>NLSA information is used by all stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.</p>
<p>There are no mechanisms in place to monitor the consequences of the NLSA.²³ *</p>	<p>This option does not apply to this dimension.</p>	<p>There are some mechanisms in place to monitor the consequences of the NLSA.</p>	<p>There are a variety of mechanisms in place to monitor the consequences of the NLSA.</p>

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

1. The Measuring Learning Achievement was administered for the first time in 1994, and has since been operating on an irregular basis. It continued to be implemented in 1995, 1997, 2003, and 2012, each year assessing a different grade level and different set of subjects. In the latest assessment (2012-2013), the NLSA measured the educational achievement of grades 4 and 7 in the subjects of Arabic, French, English, and Mathematics.
2. There is no policy document pertaining to NLSA. To conduct a national large-scale assessment, the Centre for Educational Research and Development (CERD) must submit a proposal to, and receive approval from the Minister of Education.
3. This option does not apply to this dimension as there is no policy document pertaining to NLSA.
4. There is no plan for NLSA activity.
5. While policymakers and educators support the national large-scale assessment program, it is unclear to what extent other stakeholder groups support it.
6. There is irregular funding allocated to the national large-scale assessment by the budget of the Center for Educational Research and Development.
7. Funding covers all core NLSA activities (design, administration, analysis, and reporting), as well as staff training.
8. Funding covers research and development activities.
9. The NLSA office is comprised of a group of people temporarily assigned to carry out the assessment exercise.
10. Political considerations sometimes hamper technical considerations.
11. The NLSA office is accountable to the Central Inspectorate, which is a monitoring body, and to a special issues committee mandated by the Center for Educational Research and Development.
12. The NLSA office is composed of mainly temporary or part-time staff. No issues have been identified with their performance.
13. Only funding opportunities for attending international programs or courses or workshops in education measurement and evaluation are available in Lebanon for preparing individuals for work on the NLSA.
14. The NLSA measures performance against national curriculum or learning standards.
15. What the NLSA measures is largely accepted by stakeholder groups.

16. There are ad hoc reviews of the NLSA to ensure that it measures what it is intended to measure.
17. There are occasional courses, workshops, or presentations on the NLSA.
18. Options to include all students groups in the national large-scale assessment exercise include special plans to ensure that the assessment is administered to students in hard-to-reach areas as well as offering the large-scale assessment in the language of instruction for most student groups.
19. The mechanisms that are in place to ensure the quality of the NLSA are training all proctors or administrators according to protocol, having a standardized manual for large-scale assessment administrators, conducting a pilot before the main data collection takes place, having all booklets numbered, double processing of data, and having internal reviewers or observers. However, discrepancies are not recorded on a standard sheet. There is also no double data scoring, external reviewers or observers, or external certification or audit.
20. There is no technical report or other documentation about the NLSA.
21. Reports with NLSA results are available for all stakeholder groups, as well as main reports containing information on overall achievement levels and subgroups. There are also workshops or presentations for key stakeholders on the results. However, reports do not contain information on trends over time overall and for subgroups or standard errors. Additionally, there is no media briefing or coverage to discuss or feature NLSA results, and the NLSA results are not disseminated within 12 months after the large-scale assessment is administered.
22. NLSA information is not used or is used in ways inconsistent with the purposes of the technical characteristics of the assessment.
23. There are no mechanisms in place to monitor the consequences of the NLSA.

LEBANON
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.

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ENABLING CONTEXT 1: <i>Setting clear policies for ILSA</i></p>			
<p>The country/system has not participated in an ILSA in the last 10 years.</p>	<p>This option does not apply to this dimension.</p>	<p>The country/system has participated in at least one ILSA in the last 10 years.</p>	<p>The country/system has participated in two or more ILSA in the last 10 years.¹ *</p>
<p>The country/system has not taken concrete steps to participate in an ILSA in the next 5 years.</p>	<p>This option does not apply to this dimension.</p>	<p>The country/system has taken concrete steps to participate in at least one ILSA in the next 5 years.² *</p>	<p>This option does not apply to this dimension.</p>
<p>There is no policy document that addresses participation in ILSA.</p>	<p>There is an informal or draft policy document that addresses participation in ILSA.</p>	<p>There is a formal policy document that addresses participation in ILSA.³ *</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>The policy document is not available to the public.⁴ *</p>	<p>The policy document is available to the public.</p>	<p>This option does not apply to this dimension.</p>
<p>ENABLING CONTEXT 2: <i>Having regular funding for ILSA</i></p>			
<p>There is no funding for participation in ILSA.</p>	<p>There is funding from loans or external donors.⁵ *</p>	<p>There is regular funding allocated at discretion.</p>	<p>There is regular funding approved by law, decree or norm.</p>
<p>This option does not apply to this dimension.</p>	<p>Funding covers some core activities of the ILSA.⁶ *</p>	<p>Funding covers all core activities of the ILSA.</p>	<p>This option does not apply to this dimension.</p>
<p>Funding does not cover research and development activities.⁷ *</p>	<p>This option does not apply to this dimension.</p>	<p>This option does not apply to this dimension.</p>	<p>Funding covers research and development activities.</p>

(CONTINUED)

<p>LATENT</p> 	<p>EMERGING</p> 	<p>ESTABLISHED</p> 	<p>ADVANCED</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.⁸ *</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.⁹ *</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.¹⁰ *</p>

SYSTEM ALIGNMENT

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

LATENT ● ○ ○ ○	EMERGING ● ● ○ ○	ESTABLISHED ● ● ● ○	ADVANCED ● ● ● ●
SYSTEM ALIGNMENT 1: <i>Providing opportunities to learn about ILSA</i>			
The ILSA team has not attended international workshops or meetings.	The ILSA team attended some international workshops or meetings.	The ILSA team attended all international workshops or meetings. ¹¹	This option does not apply to this dimension.
The country/system offers no opportunities to learn about ILSA. ¹² *	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. ¹³ *	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: <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. ¹⁴ *	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. ¹⁵ *	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. ¹⁶ *	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. ¹⁷ *	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. ¹⁸ *	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. ¹⁹ *	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. ²⁰ *	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. In the last 10 years, Lebanon has participated in TIMSS in 2003, 2007, and 2011, and in PASEC in 2009.
2. Lebanon has taken concrete steps to participate in PISA and TIMSS in 2015.
3. The formal policy document that addresses participation in ILSAs is the *Letter from the Head of the Center for Educational Research and Development*, authorized by the Ministry of Education and Higher Education on February 2, 2011.
4. The policy document is not available to the public.
5. Funding for the ILSA is provided by the regular budget of the government, allocated at the discretion of the Centre for Educational Research and Development, and from loans or external donors.
6. Funding for participating in TIMSS 2011 covers international participation fees, implementation of the assessment exercise in Lebanon (e.g., printing booklets, travel to schools), processing and analyzing data collected from implementation of the assessment exercise, and attendance at international expert meetings for the assessment exercise. However, funding does not cover reporting and disseminating the assessment results in Lebanon.
7. Funding does not cover research and development activities.
8. There is a team and national coordinator to carry out the ILSA activities.
9. The national coordinator is fluent in the language of assessment.
10. The national coordinator has a sufficiently staffed team that has previous experience working on international assessments and necessary training to carry out the required assessment activities effectively. The national coordinator is fluent in the language in which the international-level meetings are conducted and related documentation is available. No issues were identified with the carrying out of TIMSS 2011 in the country.
11. The ILSA team attended all international workshops and meetings related to the assessment.
12. There are no opportunities to learn about ILSA in Lebanon.
13. This option does not apply to this dimension because no opportunities to learn about ILSAs are offered in Lebanon.
14. Lebanon met sufficient standards to have its data presented beneath the main display of the international report or in an annex.

15. Lebanon has not contributed new knowledge on ILSA.
16. Country-specific results and information are not disseminated in Lebanon and have only been published in the international report. There are plans to publish a national report in Lebanon.
17. Products to provide feedback to schools and educators about the ILSA results are not made available.
18. There is no media coverage of the ILSA results.
19. While it is planned for the ILSA results to be used to inform curriculum improvement, teacher-training programs, and other assessment activities in the system, results from the ILSA have not been used to inform decision making at the national level.
20. It is not clear that decisions based on ILSA results have had a positive impact on students' achievement levels.

Acknowledgements

This report, part of a 16-country benchmarking exercise in the Middle East & North Africa and Africa regions, was prepared by the World Bank SABER Student Assessment team, in partnership with the Arab League Educational, Cultural, and Scientific Organization (ALECSO) and the United Nations Educational, Scientific, and Cultural Organization (UNESCO), which led data collection efforts. This effort is part of the Arab Regional Agenda for Improving Education Quality (ARAIEQ) led by ALECSO in partnership with the World Bank. It benefited from feedback and review from Juan Manuel Moreno, Lead Education Specialist and Task Team Leader for education projects in Lebanon at the World Bank, staff at the Education Sector Development Secretariat (ESDS) at the Ministry of Education and Higher Education in Lebanon, as well as comments received during a national validation workshop held in Lebanon.

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. 2010. Country Partnership Strategy for Lebanon for the Period FY11-FY14. Report No. 54690-LB. Washington, DC: World Bank.

— — —. Lebanon Country Indicator Data. Washington, DC: World Bank. Data retrieved from <http://data.worldbank.org/> on October 10, 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.

