Syria

STUDENT ASSESSMENT

Key Policy Areas for Student Assessment

Status

1. Classroom Assessment

In Syria, formal, system-level documents provide guidelines for classroom assessment. In addition, a variety of system-level mechanisms ensure that teachers develop skills and expertise in classroom assessment. Such mechanisms include pre- and in-service teacher training programs, all of which have a required component on classroom assessment. At the same time, classroom assessment practices are considered to be weak. Classroom assessment activities tend to provide little useful feedback to students and are used mainly as an administrative tool rather than as a pedagogical resource.

Emerging

2. Examinations

The General Secondary Certificate examination is administered to grade 12 students. The main purposes of this examination are student certification for school cycle completion and selection to higher-education institutions. The Examination Directorate is responsible for the examination. It is adequately staffed and has state-of-the art facilities for carrying out examination activities. Regular funding for all core activities is provided by the government. However, teachers have limited opportunities to learn about the examination. Additionally, few mechanisms are in place to monitor the consequences of the examination in terms of its effects on students, teachers, and the instructional process.

Established

3. National Large-Scale Assessment (NLSA)

Provincial assessments are administered to students in grades not assessed in the general examinations. Each year, provinces select a subject at a specific grade level and a provincial assessment is administered to all students at that grade level in the province. The main purposes of provincial assessments are to identify students' learning levels, diagnose learning difficulties, and evaluate teachers' performance and the curriculum in order to inform further curriculum development. The provincial assessments receive regular funding from the government that covers the cost of some activities. In addition, the assessment office responsible for the provincial assessments is adequately staffed. However, there are no mechanisms in place to ensure that these assessments accurately measure what they are intended to measure. There also are no mechanisms to ensure the quality of the testing process and the assessment results.

Emerging

4. International Large-Scale Assessment (ILSA)

Syria participated in the Trends in International Mathematics and Science Study (TIMSS) in 2003, 2007, and 2011, and has taken concrete steps to participate in TIMSS 2015. Funding for some core ILSA activities is allocated by the government and external donors. Although the ILSA team has previous experience working on international assessments, the ILSA office is inadequately staffed. There also are no opportunities to learn about ILSAs in the country. Despite the fact that results from ILSAs have informed curriculum improvement and teacher training programs, country-specific ILSA results are not disseminated in Syria.

Emerging

As of November 2015, in-country validation of this report had not taken place. Information on Syria's assessment system was collected in March 2013. Consequently, the findings in this report reflect the status of the country's assessment system prior to the worsening conflict situation. Potential policy options therefore may have diminished or varying relevance depending on current circumstances in Syria and in the sector.

Systems Approach for Better Education Results

WORLD BANK GROUP

Introduction

Syria 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, Syria 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- and 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 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 Syria

The Syrian Arab Republic is a lower-middle-income country in the Middle East region. Since March 2011, Syria has been engulfed in an escalating civil war with large humanitarian and economic impacts. More than 6.8 million people are in

4

need of humanitarian assistance, and the total number of registered Syrian refugees in neighboring countries is growing, with a large share of those being school-aged children. It is estimated that the economy contracted by about 30 percent in 2012.

General education in Syria consists of Primary Education (grades 1 to 6), Lower Secondary Education (grades 7 to 9), and Upper Secondary Education (grades 10 to 12). Education is compulsory and free from grades 1 to 9. At the end of grade 9, national examinations take place to determine if the student continues to a general or technical secondary school. In 2009, before the conflict, net primary school enrollment was 93 percent, and net secondary school enrollment was 68 percent.

Since 2011 the conflict has had serious impacts on the education system in Syria. One-fifth of the country's schools are no longer used as learning environments because they are either destroyed or used as shelters for displaced persons. Where schools are still open, average attendance is approximately two days per week, and in some of the most affected areas, the student attendance rate has dropped to 6 percent. Low attendance rates are attributed to insecurity, lack of teachers and resources, damaged buildings, and the pressure to drop out and earn an income or get married early. For those students still able to attend school, the quality of education has severely declined because of shortened hours, overcrowding, double shifting, and lack of materials. It is estimated that some Syrian children have missed out on as much as two years of education during the ongoing civil struggle.

Detailed information on Syria's student assessment system was collected 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 macrolevel. Additional data would need to be collected to determine actual, on-the-ground practices in Syria, particularly by teachers and students in schools. The following sections discuss the findings for each assessment type, accompanied by potential policy options. Detailed, completed rubrics for each assessment type are provided in Appendix 5.

Classroom Assessment

Level of Development: EMERGING

In Syria, the *By-laws for Basic Education Schools* document, authorized by the Ministry of Education in 2004, provides guidelines for classroom assessment at the basic education level. At the secondary level, the *Interior System for High Schools* document, authorized by the Ministry of Education in 1994, provides guidelines for classroom assessment.

Some system-wide resources are available to teachers to help them engage in classroom assessment activities. For example, standards documents, which specify general and subject-specific learning expectations for students, are available at the Ministry of Education in the Curriculum and Supervision Directorate. Textbooks also provide support for classroom assessment activities, and sample questions and scoring instructions are circulated annually to schools.

Syria also has a variety of system-level mechanisms in place to ensure that teachers develop skills and expertise in classroom assessment. For example, opportunities to learn about classroom assessment are made available through pre- and in-service teacher training programs, all of which include a required course on classroom assessment. Preservice teacher training in Syria is available either in the form of an education diploma after graduation from university (for high school teachers) or following four years of schooling in a faculty of education (for basic education teachers). In-service teacher training opportunities are available to all teachers in primary and secondary schools. In the 2009–10 school year, the Ministry of Education embarked on a three-year curriculum development project for the General Education System. During each academic year, the Ministry of Education trained teachers and educators on the developed curricula. Among the most prominent and important training issues was classroom assessment. The Center for Education Measurement and Assessment was recently created for the purposes of evaluating all elements of the education process, putting in place appropriate tools to measure student learning, and training teachers to develop their classroom assessment skills.

At the same time, classroom assessment practices are considered to be weak, with a tendency to be overly focused on information recall and lacking alignment with the curricular framework. Errors in the scoring or grading of students' work are also frequently observed, and grade inflation and the uneven application of standards for grading students' work are serious problems.

Limited systematic mechanisms are in place to monitor the quality of classroom assessment practices. For example, classroom assessment is a required component of a teacher's performance evaluation, and national reviews of the quality of education include a focus on classroom assessment.

Potential policy options are the following:

1. Ensure a variety of system-wide resources are available for teachers to use in carrying out classroom assessment activities. For example, develop a document that outlines the level of performance that students are expected to reach in different subject areas at different grade levels. Ensure that this document receives the necessary authorization from the relevant body and disseminate it through teacher training programs. Strengthen and integrate different areas of assessment and ensure that resources provided for teachers are based on the national standards and are also aligned with the national examination frameworks.

2. Introduce a variety of mechanisms to systematically monitor the quality of classroom assessment practices. For example, introduce an external moderation system to review the difficulty of classroom assessment activities and the appropriateness of scoring criteria.

6

3. Ensure that classroom assessment information is disseminated to all key stakeholders by establishing a requirement for it to be disseminated to, for example, students, parents, and school district or Ministry of Education officials in a format that is most appropriate for each stakeholder group.

Examinations

Level of Development: ESTABLISHED

The General Secondary Certificate examination's main purposes are certifying student completion of the secondary school cycle and making decisions about student selection to higher-education institutions. The examination was first administered in 1945 and continues to be administered to students in grade 12 in Arabic language, English, French, religious education, and technical subjects. The examination is authorized by the Ministry of Education through a formal, system-level document known as the *Executive Instructions of General Exam for Secondary Education*.

Regular funding for the examination is allocated by the government. Funding covers all core examination activities including design, administration, data analysis, and reporting. Funding also covers planning of program milestones as well as staff training. However, funding does not cover research and development activities.

The Examination Directorate, a branch within the Ministry of Education, has had primary responsibility for running the examination since 1958. The Examination Directorate is responsible for administering the examination, identifying examination centers in the provinces, printing and distributing examination papers, issuing student results, and conducting data analysis. The Examination Directorate relies on specialized committees under the supervision of the Curriculum and Supervision Directorate to score the examination answer sheets. The Examination Directorate has state-of-the-art facilities and full-time staff, which enable it to carry out the examination effectively, with minimal issues.

In addition, Syria offers a wide range of opportunities that prepare individuals to work on the examination, including university graduate courses and programs on educational measurement and evaluation. Internships are also offered in the examination office.

Teachers are required to participate in courses on the examination. However, these courses are not regularly updated. Furthermore, teachers are involved in very few examination-related tasks, which include administering the examination, participating in setting the scoring rules, and resolving inconsistencies between examination scores and school grades.

Internal review is the only systematic mechanism in place to ensure the quality of the examination. Specifically, committees from the Central Administration (each is commissioned to a specific province), along with members of the educational directorates representing the Ministry of Education, conduct field visits during the examination.

Inappropriate behavior surrounding the examination process is moderate. Copying from candidates, using unauthorized materials such as prepared answers and notes, and collusion among candidates all occur during the examination process.

Once the examination results are finalized, students' names and results are made public. Although expert review groups monitor the consequences of the examination, no permanent oversight committee is in place. In addition, no funding is available for independent research on the impact of the examination on students, teachers, or the instructional process.

Potential policy options are the following:

1. Ensure that the available opportunities to prepare current and future staff for work on the General Secondary Certificate examination are adequate. Introduce additional opportunities to prepare for work on the examination, including funding for attending international programs, courses, or workshops on educational measurement and evaluation.

2. Introduce and make widely available up-to-date courses or workshops on the General Secondary Certificate examination for teachers. At the same time, involve teachers in a variety of examination-related tasks, such as scoring the examination and supervising examination procedures.

3. Introduce a variety of mechanisms to systematically ensure the quality of the examination. For example, commission external reviews or observers or conduct pilot or field testing.

4. Ensure that a variety of options are in place for students who do not perform well on the examination. For example, introduce funding for students to attend remedial or preparatory courses in preparation to retake the examination.

5. Establish a variety of mechanisms to monitor the consequences of the examination. For example, arrange a permanent oversight committee or commit regular funding for conducting independent research on the impact of the examination.

6. To create better links between classroom assessment and the formal examination system, review the design of the General Secondary Certificate and consider including a school-based assessment portion.

National Large-Scale Assessment (NLSA)

Level of Development: EMERGING

Provincial assessments, also known as standardized examinations, have been administered annually in all provinces since the 1998–99 academic year to students in transitional grades (grades that do not have general examinations). Each province chooses a subject and grade in which a standardized assessment is administered to all students. The main purposes of the provincial assessments are to identify students' learning levels, diagnose learning difficulties, evaluate teachers' performance, and review the curriculum for the purpose of further development.

The Ministry of Education authorized the large-scale assessment program in 1998 through the *Executive Instructions of Unified Exam for Basic Education and High School* document.

Regular funding for the provincial assessments is provided by the government and covers some core activities, including assessment design, administration, and data analysis. However, funding does not cover data reporting, long- or medium-term planning of program milestones, staff training, or research and development activities.

The Research Directorate, which is in charge of the provincial assessments, is a permanent unit within the Ministry of Education and is adequately staffed with permanent and full-time staff to carry out the assessment effectively. Some opportunities are available in Syria to prepare individuals for work on the provincial assessments, including university courses and graduate programs, nonuniversity courses, and workshops on educational measurement and evaluation. The Educational and Psychological Measurement and Assessment Center was created to further address capacity-building needs in the field of assessment and measurement.

Currently no mechanisms are in place to ensure the quality of the provincial assessments. For example, there is no external or internal review, double scoring of data, training for scorers to ensure high interrater reliability, or a pilot conducted before the main data collection takes place.

Furthermore, provincial assessment results are not reported or disseminated. Although a comprehensive technical report is produced, it has restricted circulation.

Only one mechanism, a permanent oversight committee, is in place to monitor the consequences of the provincial assessments. Neither funding for independent research on the impact of the large-scale assessment nor regular focus groups or surveys of key stakeholders are in place.

Potential policy options are the following:

1. Ensure regular funding to cover all core provincial assessment activities, including data reporting and research and development.

2. Introduce mechanisms, such as independent alignment reviews by experts, to ensure that the provincial assessments accurately measure what they are intended to measure.

3. Introduce a variety of mechanisms to systematically ensure the quality of the provincial assessments. For example, train all proctors or administrators according to a protocol, and develop and disseminate a standardized manual for the assessment administrators.

4. Establish a plan for effective dissemination of assessment results, which includes a strategy for developing and disseminating reports for all key stakeholder groups within 12 months after the assessment is administered.

International Large-Scale Assessment (ILSA)

Level of Development: EMERGING

In the last 10 years, Syria has participated in TIMSS 2003, 2007, and 2011 and taken concrete steps to participate in TIMSS 2015.

Funding for international assessment participation is provided by the government (allocated at discretion) and external donors. Funding covers some core ILSA activities, including international participation fees, processing and analyzing data collected from the implementation of the assessment exercise, reporting and disseminating the results in Syria, and attendance at international expert meetings. However, funding does not cover the implementation of the assessment exercise in Syria or research and development activities.

Although a team and a national coordinator are responsible for the international assessment, the ILSA office is inadequately staffed and trained to carry out the assessment effectively. The team has previous experience working on international assessments and has attended some international meetings, however issues have been identified with the carrying out of ILSAs in Syria. For example, there have been complaints about poor training of test administrators. In addition, the teams working on ILSA activities in the provinces are not dedicated full-time to completing these activities, but take on ILSA work in addition to their primary jobs. Team members have expressed dissatisfaction about the low compensation received for working on the international assessment. Opportunities to learn about ILSAs are also not offered in Syria.

Country-specific ILSA results, products, and information are not disseminated to key stakeholders (including schools and educators) in Syria. However, ILSA results are used in some ways, including to inform curriculum improvement, teacher training programs, and other assessment activities in Syria. However, it is not clear whether decisions based on ILSA results have had a positive impact on students' achievement levels.

Potential policy options are the following:

1. Develop a formal policy document that addresses Syria's participation in ILSAs, ensure that it is approved by the appropriate authorizing body, and make it publicly available.

2. Ensure that funding for ILSA covers all core activities, including implementation of the assessment exercise in Syria, and research and development activities.

3. Introduce opportunities for staff and others interested in ILSA to learn about the assessment and prepare for work on it. For example, offer workshops or meetings on using international assessment databases.

4. Develop a plan to ensure that Syria-specific ILSA results are regularly and widely disseminated in the country. For example, ensure that a national report with results is disseminated to key stakeholders, and products that provide feedback to schools and educators are also systematically made available.

5. Ensure that ILSA results are used in a variety of ways to inform decision making in Syria. For example, utilize ILSA results to inform decisions related to resource allocation.

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 or 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 Syria

SYRIA
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.[2]	The document is widely available.
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.[3]	There are a variety of system-wide resources available for teachers for classroom assessment.
There is no official curriculum or standards document.	There is an official curriculum or standards document, but it is not clear what students are expected to learn or to what level of performance.	There is an official curriculum or standards document that specifies what students are expected to learn, but the level of performance required is not clear.	There is an official curriculum or standards document that specifies what students are expected to learn and to what level of performance.[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.[7]	There are varied and systematic mechanisms in place to monitor the quality of classroom assessment practices.
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.
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. *The By-laws for Basic Education Schools*, amended by decision number 3053/443 in August 2004, and the *Interior System for High Schools*, amended by law number 3921/443 in 1994, are formal system-level documents that provide guidelines for classroom assessment.

2. The availability of the documents is restricted. These documents are available in in-service courses for teachers and at the Central Administration of the Ministry of Education, the education directorates of the provinces, and with school principals.

3. Some system-wide resources are available for teachers to engage in classroom assessment activities. Specifically, there are standards documents, which specify students' learning expectations in general, and in various subjects in particular, available at the Ministry of Education–Curriculum and Supervision Directorate. In addition, textbooks or workbooks provide support for classroom assessment, and sample questions and scoring instructions are circulated to schools annually. However, no online assessment resources or computer-based testing is done with instant reports on students' performance.

4. *The National Standards* document is an official document that specifies what students are expected to learn and to what level of performance. It was put forth in 2007 for all subjects in both primary and secondary schools. Based on this document, guides were constructed to outline educational aims for units and lessons in each subject and performance level.

5. A variety of system-level mechanisms ensure that teachers develop skills and expertise in classroom assessment. For instance, opportunities are available to learn about classroom assessment in preservice and in-service teacher training, and all teacher training programs include a required course on classroom assessment. Preservice teacher training in Syria is available either in the form of an education diploma after graduation from university (for high school teachers) or following four years of schooling in a faculty of education (for basic education teachers). In-service teacher training opportunities are offered to all teachers in primary and secondary schools. Additionally, opportunities are available to participate 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. As of the 2009–10 academic year, the Ministry of Education embarked on a three-year curricula development project for the General Education System. In the first year, the focus was on developing the curricula for grades 1–4, 7, and 10. The development of the curricula for grades 5, 8, and 11 occurred in the first and second year, while that of grades 6, 9, and 12 occurred in the third year. During each academic year, the Ministry of Education trained teachers and educators on the developed curricula. Among the most prominent and important training issues was classroom assessment. Recently a Center for Education Measurement and Assessment was created. The center will be in charge of evaluating all elements of the education process, including putting in place the appropriate measurement tools and training teachers to develop their classroom assessment skills.

6. Classroom assessment practices are known to be weak. It is very common for classroom assessment activities to be mainly about recalling information, and classroom assessment activities tend to lack alignment with a pedagogical or curricular framework. Additionally, it is common for classroom assessment activities to rely mainly on multiple-choice, selection-type questions, provide little useful feedback to students, and be mainly used as an administrative or control tool rather than as a pedagogical resource. It is common to observe errors in the scoring or grading of students' work, and grade inflation and the uneven application of standards for grading students' work are serious problems. However, it is not common for teachers to not use explicit or a priori criteria for

SYSTEMS APPROACH FOR BETTER EDUCATION RESULTS

scoring or grading students' work. Although teachers typically use a scale or criteria to correct written examinations, students' work throughout the year is assessed without clear tools and is measured based on written assignments and oral recitations.

7. Limited systematic mechanisms are in place to monitor the quality and ensure the appropriateness of classroom assessment practices. Such monitoring mechanisms include, for example, evaluating whether and to what extent teachers provide feedback to students on their work, and whether teachers are appropriately applying grading criteria for classroom-level tests. In Syria, classroom assessment is a required component of a teacher's performance evaluation and school inspection or teacher supervision, and national or other system-wide reviews of the quality of education include a focus on classroom assessment. However, no external moderation system is in use to review the difficulty of classroom assessment activities or appropriateness of scoring criteria, and no government funding is available for research on the quality of classroom assessment activities and how to improve classroom assessment.

8. Classroom assessment information is required to be disseminated to some key stakeholders, including parents and students. However, it is not required to be disseminated to school district or Ministry of Education officials.

9. There are adequate required uses of classroom assessment to support student learning, including its use as an input for external examination results. Classroom assessment activities are required to be used for diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child's learning, grading students for internal classroom uses, and providing input to an external examination program. However, classroom assessment activities are not required to be used in planning next steps of instruction.

SYRIA
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.	Most stakeholder groups support the examination.[5]	All stakeholder groups support the examination.
There are no attempts to improve the examination by stakeholder groups.	This option does not apply to this dimension.	There are independent attempts to improve the examination by stakeholder groups.	There are coordinated attempts to improve the examination by stakeholder groups.[6]
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.

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.[12]	This option does not apply to this dimension.	The examination office is accountable to an external board or agency.	This option does not apply to this dimension.
Examination results are not recognized by any certification or selection system.	Examination results are recognized by a 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 systems 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.	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.[14]

(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.	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 examination.	This option does not apply to this dimension.	The country offers some opportunities that prepare for work on the examination.	The country offers a wide range of opportunities that prepare for work on the examination.[16]

SYSTEM ALIGNMENT

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

SYSTEM ALIGNMENT 1: Aligning examinations with learning goals and opportunities to learn

LATENT ●OOO	EMERGING ●●OO	ESTABLISHED ●●●O	ADVANCED ●●●●
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.[18]	This option does not apply to this dimension.	What is measured by the examination is largely accepted by stakeholder groups.	This option does not apply to this dimension.
Material to prepare for the examination is minimal, and it is only accessible to very few students.	There is some material to prepare for the examination that is accessible to some students.	There is comprehensive material to prepare for the examination that is accessible to most students.[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

LATENT	EMERGING	ESTABLISHED	ADVANCED
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.	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.[21]	Teachers are involved in some examination-related tasks.	Teachers are involved in most examination-related tasks.

ASSESSMENT QUALITY

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

LATENT ○○○●	EMERGING ●●●○	ESTABLISHED ●●●○	ADVANCED ●●●●
ASSESSMENT QUALITY 1: *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.[22]	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.[23]	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.	Inappropriate behavior surrounding the examination process is moderate.[24]	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.[25]	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.[26]

(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.[27]
Student names and results are public.[28]	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.[29]	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.[30]	There is a variety of mechanisms in place to monitor the consequences of the examination.

Examinations: Development-Level Rating Justifications

1. The General Secondary Certificate was first administered in 1945, with the purposes of student certification for grade or school cycle completion and student selection to university or other higher-education institutions. It is administered to students in grade 12, and covers Arabic language, philosophy, geography, history, religious education, English, and French for the General Secondary–Literature section. For the General Secondary—Science section, subjects covered include Arabic language, biology, physics, chemistry, mathematics, religious education, English, and French.

2. The Ministry of Education authorized the examination through the *Executive Instructions of General Exam for Secondary Education* document in 2012.

3. *The Executive Instructions of General Exam for Secondary Education* document is available to the public.

4. *The Executive Instructions of General Exam for Secondary Education* addresses some key aspects of the examination, including outlining governance, distribution of power, and responsibilities among key entities. It also describes the purpose of the examination and authorized uses of results, as well as procedures to investigate and address security breaches, cheating, or other forms of inappropriate behavior. Additionally, it outlines procedures for special/disadvantaged students, specifies who can sit for the examination, and identifies rules about preparation. However, it does not state funding sources or explain alignment with curricula and standards or the format of the examination questions.

5. Most stakeholder groups support the examination. Although policy makers express strong support of the examination, teacher unions, educators, students, parents, media, and universities all support it. However, it is unclear whether think tanks, NGOs, and employers support the examination.

6. Coordinated attempts have been made to improve the examination by stakeholder groups. When the development of the pre-university education curricula was finalized, the Ministry of Education was keen to develop the assessment system, particularly with regard to the general examinations, which include the General Secondary Certificate examination. Thus, training was organized before the start of the academic year to address assessment issues, including holding discussion on the General Secondary Certificate examination. Subsequently, a workshop was organized at the Ministry of Education, with senior supervisors from the ministry and measurement and assessment specialists from the Faculty of Pedagogy at the University of Damascus participating. During this workshop, among other activities, a specifications table for the General Secondary Certificate examination for all educational subjects was elaborated. The findings of the workshop were disseminated via the ministry's website and to schools. Education Directorates of all provinces were asked to provide the ministry with samples of the achievement examinations that were conducted in the first semester. These samples were studied in line with the achievement examinations that were conducted in the first semester. These samples were studied in line with the specifications table and then debated in a subsequent central workshop that gathered specialized supervisors in all educational subjects. Debates focused on field samples in line with the specifications table with the view of improving them. During the last workshop, it was noted that most teachers still lean toward traditional examinations that measure lower levels (memorization and understanding, in the best of cases). Furthermore, most teachers face difficulties in formulating questions. These observations were provided to the specialized supervisors attending the workshop to be conveyed to teachers through weekly training sessions held on Saturdays.

7. Efforts to improve the examination are generally welcomed by the leadership in charge of the examination.

SYSTEMS APPROACH FOR BETTER EDUCATION RESULTS

8. Regular funding is allocated by the government for the examination.

9. Funding covers all core examination activities including examination design and administration as well as data analysis and reporting. Funding also covers long- or medium-term planning of program milestones and staff training.

10. Funding does not cover research and development activities.

11. The Examination Directorate, a branch within the Ministry of Education, has had primary responsibility for running the examination since 1958. The Directorate is responsible for administering the examination, specifying the examination centers in the provinces, printing and distributing questions, and producing and issuing the results and data analysis. However, it relies on specialized committees under the supervision of the Curriculum and Supervision Directorate to score the examination answer sheets. The Examination Directorate also includes subunits, called Examination Departments, in each province.

12. The examination office is not accountable to an external board or agency.

13. Examination results are recognized by certification or selection systems in Syria and in all Arab countries.

14. The examination office has state of the art facilities to carry out the examination including computers for all technical staff, a secure building and storage facilities, access to adequate computer servers, an ability to back up data, and adequate communication tools.

15. The examination office is adequately staffed with permanent or full-time staff to effectively carry out the examination with minimal issues. The minimal issues identified with the performance of human resources responsible for the examination include weakness in test design, omission of curricular topics, and errors in the examination.

16. Syria offers university graduate programs, university courses, and non-university training courses or workshops on educational measurement and evaluation. Internships in the examination office are offered as well. However, no funding is available for attending international programs, courses, or workshops on educational measurement and evaluation.

17. There is a clear understanding that the examination measures the national curriculum standards.

18. Some stakeholder groups question what the examination measures.

19. Comprehensive material is available to prepare for the examinations that is accessible to most students. Material includes examples of the types of questions that are on the examination, and information on how to prepare for the examination is available. However, resources such as the framework document explaining what is measured on the examination and the report on the strengths and weaknesses in student performance are not available.

20. Although compulsory courses or workshops on the examinations are available to teachers, these courses are not regularly updated.

21. Teachers are involved in very few examination-related tasks. Although they are involved in administering the examination, participating in setting the scoring rules, and resolving inconsistencies between examination scores and schools' grades, they are not involved in selecting or creating examination questions, scoring the examination, acting as a judge, or supervising examination procedures.

22. There is a comprehensive technical report but with restricted circulation.

23. Internal review or observation is the only systematic mechanism in place to ensure the quality of the examination. Internal review occurs as the quality of the examinations is verified by committees from the Central Administration (with each one commissioned to a specific province) along with some members of the educational directorates representing the Ministry, who conduct field visits during the examinations. There are no external reviews or observers, external certification or audits, pilot or field testing, or translation verification.

24. Inappropriate behavior surrounding the examination process is moderate. Behaviors such as copying from candidates, using unauthorized materials such as prepared answers and notes, and collusion among candidates occurs. However, other behaviors, such as leakage of the content of the examination paper or part of a paper before the examination and intimidation of examination supervisors, markers, or officials, have not been reported as diminishing the credibility of the examination.

25. The examination results are perceived as credible by all stakeholder groups.

26. All students can take the examination; there are no language, gender, or other equivalent barriers.

27. Examination results are used by all stakeholder groups in a proper way. There is no systematic evidence of improper use of examination results.

28. Student names and results are public.

29. Very limited options are open to students who do not perform well on the examination. Students who do not perform well may retake the examination or repeat the grade. For one time only, students with passing grades wanting to improve their overall score can choose three subjects to retake in the second round of examinations in the same year. Students can also retake the second round of examinations if they fail in three subjects. In this case, the student can retake the examination from within the school system. If he or she fails again, he or she can retake the examination as a guest auditor not enrolled in the school system. However, students do not have the option to attend remedial or preparatory courses to prepare for retaking the examination or opt for less selective schools, universities, or tracks.

30. Expert review groups are the only mechanism in place the monitor the consequences of the examination. There is no funding for independent research on the impact of the examination or a permanent oversight committee.

SYSTEMS APPROACH FOR BETTER EDUCATION RESULTS

SYRIA

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.	The NLSA is a stable program that has been operating regularly.[1]	This option does not apply to this dimension.
There is no policy document pertaining to NLSA.	There is an informal or draft policy document that authorizes the NLSA.	There is a formal policy document that authorizes the NLSA.[2]	This option does not apply to this dimension.
This option does not apply to this dimension.	The policy document is not available to the public.[3]	The policy document is available to the public.	This option does not apply to this dimension.
There is no plan for NLSA activity.	This option does not apply to this dimension.	There is a general understanding that the NLSA will take place.	There is a written NLSA plan for the coming years.[4]
ENABLING CONTEXT 2: *Having strong public engagement for NLSA*			
All stakeholder groups strongly oppose the NLSA or are indifferent to it.	Some stakeholder groups oppose the NLSA.	Most stakeholder groups support the NLSA.[5]	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.[6]	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding covers some core NLSA activities: design, administration, analysis, and reporting.[7]	Funding covers all core NLSA activities: design, administration, analysis, and reporting.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development activities.[8]	This option does not apply to this dimension.	Funding covers research and development activities.
ENABLING CONTEXT 4: *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.[9]	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.[10]
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)

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
ENABLING CONTEXT 5: *Having effective human resources for NLSA*			
There is no staff allocated for running an NLSA.	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.[12]	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.[13]	The country offers a wide range of opportunities to prepare individuals for work on the NLSA.

SYSTEM ALIGNMENT

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

LATENT ○○○●	EMERGING ●●●○	ESTABLISHED ●●●○	ADVANCED ●●●●
SYSTEM ALIGNMENT 1: *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.[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.[15]	What the NLSA measures is largely accepted by stakeholder groups.
There are no mechanisms in place to ensure that the NLSA accurately measures what it is supposed to measure.[16]	There are ad hoc reviews of the NLSA to ensure that it measures what it is intended to measure.	There are regular internal reviews of the NLSA to ensure that it measures what it is intended to measure.	This option does not apply to this dimension.
SYSTEM ALIGNMENT 2: *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.[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.

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.[18]	This option does not apply to this dimension.	At least one option is offered to include all groups of students in the NLSA.	Different options are offered to include all groups of students in the NLSA.
There are no mechanisms in place to ensure the quality of the NLSA.[19]	This option does not apply to this dimension.	There are some mechanisms in place to ensure the quality of the NLSA.	There are a variety of mechanisms in place to ensure the quality of the NLSA.
There is no technical report or other documentation about the NLSA.	There is some documentation about the technical aspects of the NLSA, but it is not in a formal report format.	There is a comprehensive technical report but with restricted circulation.[20]	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.[21]	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.	This option does not apply to this dimension.	NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.[22]	NLSA information is used by all stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.
There are no mechanisms in place to monitor the consequences of the NLSA.	This option does not apply to this dimension.	There are some mechanisms in place to monitor the consequences of the NLSA.[23]	There are a variety of mechanisms in place to monitor the consequences of the NLSA.

National (of System-Level) Large Scale Assessment (NLSA): Development-Level Rating Justifications

1. Standardized examinations, also known as "provincial assessments," have been administered in all provinces every year since the 1998–99 academic year to students in transitional grades (grades that do not have general examinations). Each province chooses a specific subject for a specific grade, in which a standardized assessment is administered to all students. The main purposes of the provincial assessments are to identify students' learning levels, diagnose learning difficulties experienced by students, and evaluate teachers' performance and the curriculum for the purpose of subject development.

2. Through the *Executive Instructions of Unified Exam for Basic Education and High School* document, the Ministry of Education authorized the large-scale assessment program in 1998.

3. The *Executive Instructions of Unified Exam for Basic Education and High School* document is available in schools along with implementation guidelines; however, it is not available to the public.

4. There is a written NLSA plan for the coming years, which is outlined in decision 547 of February 2, 2010, the *Standardized Examination Instructions* for the transitional grades. The plan is not publicly available.

5. Most stakeholder groups, including policy makers, teacher unions, educators, and employers, support the NLSA. Students, parents, media, think tanks, NGOs, and universities are neutral to the NLSA.

6. Regular funding for the NLSA is allocated by the government.

7. Funding covers some core NLSA activities, including assessment design and administration, and data analysis. However, funding does not cover data reporting, long- or medium-term planning of program milestones, or staff training.

8. Funding does not cover research and development activities.

9. The Research Directorate, which is in charge of the large-scale assessment, is a permanent unit within the Ministry of Education. The Research Directorate issues the necessary instructions on an annual basis and supervises the implementation process in cooperation with the Curriculum and Supervision Directorate. Within this framework, the specialized supervisors in the provinces prepare standardized questions and set the scoring principles. Scoring is carried out by teachers under the auspices of the specialized supervision body, while the Research Directorate analyzes and studies the results.

10. Political considerations never hamper technical considerations. Large-scale assessment results have never been withheld from publication because of political reasons.

11. The NLSA office, the Research Directorate, is accountable to a higher office in the Ministry of Education. The Research Directorate submits the results of the standardized examinations to the Deputy Minister for Research and Examination Affairs and sends a copy to the Curriculum and Supervision Directorate.

12. The number of permanent and full-time staff to carry out the NLSA effectively, with minimal issues, is adequate. Issues identified with the performance of human resources responsible for the large-scale assessment include poor training of test administrators as well as unclear instructions and guidelines for administering the assessment.

13. Syria offers some opportunities to prepare individuals for work on the NLSA, including through university graduate programs, university courses, and non-university courses and workshops on educational measurement and evaluation. However, academic courses offer more theoretical rather than practical knowledge. The Educational and Psychological Measurement and Assessment Center was created to further address the needs of capacity building in the field of assessment and measurement. No funding is available for attending international programs, courses, or workshops on educational measurement and evaluation. Additionally, no internships or short-term employment is offered in the large-scale assessment office.

14. The NLSA measures performance against national or state-level curriculum guidelines or learning standards.

15. What the NLSA measures is questioned by some stakeholder groups.

16. No mechanisms are in place to ensure that the NLSA accurately measures what it is supposed to measure.

17. Occasional courses or workshops on the NLSA are offered to new teachers by local supervisors. However, most teachers do not have access to live courses or workshops online.

18. No options are offered to include all groups of students in the NLSA. No accommodations are provided for students with disabilities or special plans made for students in hard-to-reach areas. In addition, the large-scale assessment is not offered in the language of instruction for all student groups.

19. No mechanisms are in place to ensure the quality of the NLSA. For example, there are no external or internal reviewers, double scoring of data, training for scorers to ensure high interrater reliability, or a pilot conducted before the main data collection takes place.

20. There is a comprehensive technical report, but with restricted circulation.

21. NLSA results are not reported or disseminated.

22. NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.

SYSTEMS APPROACH FOR BETTER EDUCATION RESULTS

23. Only one mechanism, a permanent oversight committee, is in place to monitor the consequences of the large-scale assessment. No funding is available for independent research on the impact of the large-scale assessment, regular focus groups or surveys of key stakeholders, themed conferences that provide a forum to discuss research and other data on the consequences of the large-scale assessment, or expert review groups.

SYRIA
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 in an ILSA in the last 10 years.	This option does not apply to this dimension.	The country/system has participated in at least one ILSA in the last 10 years.	The country/system has participated in two or more ILSA in the last 10 years.[1]
The country/system has not taken concrete steps to participate in an ILSA in the next five 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 five years.[2]	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.[3]	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.[4]	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.[5]	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.[6]	Funding covers all core activities of the ILSA.	This option does not apply to this dimension.
Funding does not cover research and development activities.[7]	This option does not apply to this dimension.	This option does not apply to this dimension.	Funding covers research and development activities.

ENABLING CONTEXT 3:
Having effective human resources for ILSA

LATENT	EMERGING	ESTABLISHED	ADVANCED
●○○○	●●○○	●●●○	●●●●
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.[8]	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[9]	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.[10]	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.[11]	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.[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: 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.[14]	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: Ensuring effective uses of ILSA				
	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.	There is limited media coverage of the ILSA results.[18]	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.[19]	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. Syria participated in TIMSS in 2003, 2007, and 2011.

2. Syria has taken concrete steps to participate in TIMSS 2015.

3. National strategy documents include information on policies related to ILSA.

4. Information is not available to provide a rating for this indicator.

5. Funding for the ILSA is provided by regular funding of the government that is allocated at discretion, as well as by external donors. The Ministry of Education provides for expenses such as printing brochures and disseminating them to participating schools in all provinces as well as the compensation of supervisors and the examination correctors. External donors, such as UNDP and the World Bank, cover expenses such as the airfare and accommodation for the assessment team coordinator and another team member to attend meetings.

6. Funding covers some core activities of the ILSA, including international participation fees, processing and analyzing data collected from the implementation of the assessment exercise, reporting and disseminating the results in Syria, and attendance at international expert meetings. However, funding does not cover the implementation of the assessment exercise in Syria.

7. Funding does not cover research and development activities.

8. A team and national coordinator are responsible for the international assessment. The team includes Senior Supervisors from the Ministry of Education, Specialized Science and Mathematics Supervisors from the provinces, and a technical committee from the IT Directorate, whose members include specialists in data entry and analysis.

9. The national coordinator is fluent in the language of the assessment.

10. The ILSA office is inadequately staffed and trained to carry out the assessment effectively. The team has previous experience working on international assessments, and team members have attended some international meetings. Some issues have been identified with the carrying out of ILSAs in Syria. For example, there have been complaints about poor training of test administrators. Additionally, teams working in the provinces are not dedicated full-time to completing ILSA activities because they take on ILSA work in addition to their principle jobs, and they express dissatisfaction for the low compensation received for working on the international assessment.

11. The ILSA team attended some international workshops and meetings. Some obstacles prevented the Syrian team from attending certain meetings because of visa issues.

SYSTEMS APPROACH FOR BETTER EDUCATION RESULTS

12. No opportunities to learn about ILSAs are offered in Syria.

13. This option does not apply to this dimension.

14. Syria met all technical standards required to have its data presented in the main displays of the international report.

15. Syria has not contributed new knowledge on ILSA.

16. Country-specific results and information are not disseminated in Syria.

17. Products to provide feedback to schools and educators about ILSA results are not made available.

18. Media coverage of the ILSA results is limited to a few small articles.

19. Results from the ILSA are used in some ways to inform decision making in Syria. Results have been used in informing curriculum improvement, teacher training programs, and other assessment activities in Syria. For example, the National Coordinator for the international assessment, who is the General Coordinator for Biology, has applied the knowledge and experience he gained in international assessments during the elaboration of the curricula, as well as communicated the knowledge and experience to the coordinators of other subjects. The National Coordinator also played a role in developing final exams and made presentations to the senior specialized supervisors regarding Syria's participation in TIMSS during the various stages of the curriculum development process.

20. It is not clear whether decisions based on ILSA results have had a positive impact on students' achievement levels.

Acknowledgments

References

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

United Nations Children's Fund. 2013. "Syria Crisis Bi-weekly Humanitarian Situation Report." UNICEF, Amman. Data retrieved from http://www.unicef.org/appeals/
files/UNICEF_Syria_Regional_Crisis_Humanitarian_SitRep_-_10_OCT_2013.pdf on October 15, 2013.

———. 2013. "Syria Crisis UNICEF Assessment Findings on the Situation in the Sectors of WASH, Education and Nutrition (Feb 2013)." UNICEF, Amman. Data retrieved from
http://www.unicef.org/appeals/files/UNICEF_Syria_Crisis_Assessment_Findings_Feb_2013.pdf on October 15, 2013.

World Bank. 2013. "Syrian Arab Republic Country Indicator Data." World Bank, Washington, DC. Data retrieved from http://data.worldbank.org/ on October 15, 2013.

———. "Syria Overview." World Bank, Washington, DC. Data retrieved from www.worldbank.org/en/
country/syria/overview on October 15, 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 policy makers 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

WORLD BANK GROUP