



Key Policy Areas for Student Assessment

Status

1. Classroom Assessment

In Tunisia, formal, system-level documents provide guidelines for classroom assessment. There are also a variety of mechanisms in place to systematically monitor the quality of classroom assessment practices, including quality reviews of the education system that include a focus on classroom assessment. In addition, there are system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment, such as pre-service teacher training at universities and special institutes. There are limited resources available to teachers for engaging in classroom assessment. These resources include scoring criteria and rubrics for evaluating students' work.

Established



2. Examinations

The Baccalaureate examination has been administered to grade 13 students since 1956. It is used primarily for certification of student completion of secondary school and for selection to higher-education institutions. Regular funding for the examination is provided by the government and from student fees. Funding covers core examination activities, but not research and development. The examination office is adequately staffed to carry out the examination effectively. However, there are limited opportunities that prepare individuals for future work on the examination. For example, Tunisia offers no university courses on educational measurement and evaluation.

Established



3. National Large-Scale Assessment (NLSA)

In Tunisia, the NLSA, "Diagnostic assessment of the student's acquisition at the beginning of the seventh grade," was first implemented in the 2012-13 academic year. Most stakeholder groups, including policymakers, teacher unions, educators, think-tanks, NGOs, and universities, support the NLSA, although what it measures is questioned by some stakeholder groups. Funding allocated by the government for the NLSA is not regular, and although there is a permanent, full-time staff, it is insufficient to meet the needs of the assessment.

Emerging



4. International Large-Scale Assessment (ILSA)

Tunisia has participated in several ILSAs, including TIMSS (1999, 2003, 2007, and 2011), and PISA (2003, 2006, and 2009). The country has taken concrete steps to participate in future ILSAs, including PISA. While there is a team and national coordinator responsible for carrying out ILSA activities, the ILSA office is inadequately staffed to carry out the assessment effectively. In addition, ILSA results specific to Tunisia are irregularly disseminated in the country and there is limited media coverage of the results.

Emerging



Introduction

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

What is SABER-Student Assessment?

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

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

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

SABER-Student Assessment methodology

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

Assessment types and purposes

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

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

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

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

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

Quality drivers of an assessment system

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

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

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

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

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

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

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

Source: World Bank.

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

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

Levels of development

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

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

The basic structure of the rubrics for evaluating data collected using the standardized questionnaires is summarized in Appendix 2. The goal of the rubrics is to provide a country with some sense of the development level of its assessment activities compared to best or

recommended practice in each area. For each indicator, the rubric displays four development levels—*Latent*, *Emerging*, *Established*, and *Advanced*. These levels are artificially constructed categories chosen to represent key stages on the underlying continuum for each indicator. Each level is accompanied by a description of what performance on the indicator looks like at that level.

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

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

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

Education in Tunisia

Tunisia is an upper-middle-income country in the Middle East & North Africa region. GDP per capita is \$4,237.

Growth was 3.6 percent in 2012, after a short contraction in 2011 which was precipitated by political and social instability in the country.

Education in Tunisia is organized into 4 levels: pre-primary, basic, (upper) secondary, and higher education. Education is compulsory for 13 years of primary and secondary schooling between the ages of 6 and 16. Between 2002 and 2007, the government of Tunisia implemented a major national education reform effort, the “School of Tomorrow.” The reform prioritized the following activities: (i) revamping of curricula and teaching and learning methods as well as improving the evaluation system; (ii) professionalizing teaching; (iii) decentralizing the education system and creating more space for civil society; (iv) making the education system more equitable and inclusive; and (v) modernizing schools by integrating new information technologies into curricula and management.

Due in part to this reform, in 2011, the net enrollment rate at the primary level was 99 percent and the transition rate from primary to secondary was 93 percent in 2009. As secondary school enrollment expands, Tunisia faces a number of challenges at this level of the system. One of these challenges is to provide a range of opportunities to a diverse student population while at the same time considering labor market needs. In order to address this challenge, which occurs alongside high unemployment (13 percent in 2010), Tunisia prioritized the strengthening of vocational training programs in its 2010-2014 plan for Economic and Social Development.

Detailed information was collected on Tunisia’s student assessment system using the SABER-Student Assessment questionnaires and rubrics. It is important to remember that these tools primarily focus on benchmarking a country’s policies and arrangements for assessment activities at the system or macro level. Additional data would be needed to determine actual, on-the-ground practices in Tunisia, particularly by teachers and students in schools. The following sections discuss the findings by each assessment type, accompanied by suggested policy options. The suggested policy options were determined in collaboration with key local stakeholders based on Tunisia’s immediate interests and needs. Detailed, completed rubrics for each assessment type are provided in Appendix 5.

Classroom Assessment

Level of development

Established



In Tunisia, there are formal, system-level documents that provide guidelines for classroom assessment. These documents are authorized by the Ministry of Education, and include the *Circular on the Evaluation and Promotion System in the Primary Level of Education* and the *Circular on the Timing of the Examination*. The availability of these documents, which pertain to the primary level of the education system, is restricted. While the documents are available in print-out form to teachers via educational institution directors and through teacher in-service courses, they are not available online, in public libraries, or in teacher training colleges.

Some system-wide resources are available to teachers to help them engage in classroom assessment. These include the *Programme of Programmes* document that outlines what students are expected to learn and to what level of performance in different subject areas at different grade levels. Teachers also have access to textbooks and workbooks that provide support for classroom assessment, scoring criteria and rubrics for grading students' work, and item banks with examples of selection/multiple-choice and supply/open-ended questions.

There are some mechanisms in place at the system-level to ensure that teachers develop skills and expertise in classroom assessment. These include pre-service teacher training at universities and special institutes, in-service teacher training, online-resources for classroom assessment, and school inspection and teacher evaluation, which have a component on classroom assessment. All teacher training programs include a required course on classroom assessment. Also, a teacher's promotion from one rank to another is based on their pedagogical score, which is assigned to teachers every two years by inspectors. This pedagogical score is partially based on the teacher's methods for conducting assessments.

Classroom assessment practices in Tunisia are viewed as being of moderate quality. There are varied and

systematic mechanisms in place to regularly monitor the quality of these assessment practices. For example, classroom assessment is a required component of a teacher's performance evaluation and of a school's inspection. Additionally, system-wide reviews of the quality of education include a focus on classroom assessment. Government funding is also available for research on the quality of classroom assessment activities and how to improve them.

Information from examinations administered by teachers at the classroom level is required to be disseminated to all key stakeholders, including school district/Ministry of Education officials, parents, and students. Parents and school districts/Ministry of Education officials are informed of individual students' performance through student score cards. However, other classroom assessment information is not required to be disseminated to key stakeholders.

At the same time, there are adequate required uses of classroom assessment to support student learning, including as an input for the Baccalaureate examination program, for diagnosing student learning issues, and for planning the next steps in instruction.

Suggested policy options

1. Introduce a variety of resources that are accessible to teachers throughout the system to help them engage in classroom assessment. For example, develop online assessment resources and make them widely available to teachers.
2. Ensure that formal, system-level documents, which provide guidelines for classroom assessment, are widely available to teachers. For example, disseminate the *Circular on the Evaluation and Promotion System in the Primary Level of Education* and the *Circular on the Timing of the Examination* in public libraries and teacher training colleges. In addition, make these documents available online for wider access.

Examinations

Level of development

Established



The Baccaalaureate is a national examination that students take upon completion of high school. It has been administered since 1956. The Baccaalaureate serves the purposes of selecting students to university or other higher-education institutions and certifying students for school cycle completion. The examination is also used to monitor education quality levels. A formal policy document, the *Decision regulating the Baccaalaureate examination*, from April 24, 2008, authorizes the examination.

Regular funding is allocated for the examination by the government and from student fees. Funding covers some core examination activities, including design, administration, data reporting, and staff training. However, it does not cover long- or medium-term planning of program milestones, data analysis, or research and development activities.

The Directorate General of Examinations in the Ministry of Education has had primary responsibility for running the Baccaalaureate examination since the 1950s. The Directorate General of Examinations is adequately staffed and has all of the required facilities to carry out the examination with no issues.

At the same time, there are very limited opportunities in the country that prepare individuals for work on the examination. For example, while there are some internships available in the examination office, there are no university graduate programs or courses on educational measurement and evaluation in Tunisia.

Teachers have access to up-to-date, compulsory courses and workshops on the examination, and they are involved in most examination-related tasks, such as creating or selecting examination questions and scoring guides, and administering and scoring the examination.

There are limited mechanisms in place to systematically ensure the quality of the examination. Existing

mechanisms include internal review, audit, and pilot testing. The Directorate General of Examinations is presently carrying out a comprehensive audit of all examination operations to evaluate their alignment with quality requirements and needed improvements. The goal of this audit is to maximize efficiency and improve related services in cooperation with a quality expert within the framework of French-Tunisian cooperation.

Inappropriate behavior surrounding the examination process is moderate. Nevertheless, certain mechanisms have been introduced to prevent or address any inappropriate behaviors, such as the use of identity cards to prevent impersonation, the reduction of classroom size to prevent copying from and collusion among candidates, and an increased level of security to prevent intimidation of examination supervisors.

In addition, there are very limited mechanisms in place to monitor the consequences of the examination in terms of its effects on student learning outcomes and opportunities. For example, while there are studies that are regularly updated, there is neither funding for independent research on the impact of the examination nor a permanent oversight committee to monitor the consequences of the examination.

Suggested policy options

1. Allocate resources (human capital and financial) in order to improve the research and analytical capacity of the Examination DG (Direction Generale des Examens) and introduce a variety of opportunities to prepare individuals for work on examinations. For example, collaborate with Tunisian universities to introduce courses and graduate programs on educational measurement and evaluation that would be available to staff of the Examination DG. Also, allocate funding for Examination DG staff to attend international programs, courses, or workshops on educational measurement and evaluation.

2. Ensure that funding is available for research and development activities related to the Baccaalaureate.

3. Disseminate to key stakeholders a comprehensive report that includes basic data on the Baccalaureate (and the Brevet, if available) that can be used to meet the purpose of monitoring education quality levels.

4. In order to strengthen the reliability and validity of national examination results for making inferences about student performance on curricular learning objectives, introduce varied and systematic mechanisms to ensure the quality of the examinations. For example, invite external reviewers or provide translation verification for quality assurance purposes.

National Large-Scale Assessment (NLSA)

Level of development



In Tunisia, the "Diagnostic assessment of the student's acquisition at the beginning of the seventh grade" was first implemented in the 2012-13 academic year. Its purpose is to monitor education quality at the system level as well as to support schools and teachers, policy design, evaluation, and decision making. The NLSA assesses mathematics, Arabic, and French, and is administered to a representative random sample of students.

A formal policy document, the *Orientation Law on Education and School Learning*, Law number 80-2002, authorized the assessment on July 23, 2002. In addition, there is a written NLSA plan for the coming years, specifying who will be tested and in which subject areas. However, the plan is not available to the public.

Although there is irregular funding allocated to the NLSA from the government, funding covers all core NLSA activities and research and development.

The National Center for Pedagogical Innovation and Educational Research is a permanent NLSA office. Although there is permanent full-time staff, it is insufficient to meet the needs of the assessment. Therefore, the center solicits the help of some inspectors and teachers to complete scoring, supervision, data analysis, and field work. There have been some challenges in carrying out the NLSA, including inspectors' lack of mobility due to security concerns.

The country offers some opportunities to prepare individuals for work on the NLSA. Non-university training courses and workshops, as well as funding for attending international training programs, courses, or workshops on educational measurement and evaluation are available. However, there are no university courses or graduate programs specifically focused on educational measurement and evaluation.

Some mechanisms are in place to ensure the quality of the NLSA. For example, a standardized manual for large-scale assessment administrators is in place, a pilot is conducted before the main data collection takes place, all booklets are numbered, and internal reviewers or observers are employed.

Specialized technical committees by subject area have been established to monitor the results of the NLSA. However, other potentially useful mechanisms for monitoring the impact of the NLSA on education quality and decision-making, such as expert review groups and permanent oversight committees, do not exist.

Suggested policy options

1. Institutionalize the implementation of the NLSA as a regular activity in key grades and subjects as a means to evaluate and monitor the performance of the system over time. Ensure that there is regular funding allocated by the government to cover all core assessment activities.
2. Ensure that there is adequate staff in the National Center for Pedagogical Innovation and Educational Research to effectively carry out the assessment, particularly in the areas of assessment design and data analysis. In addition, ensure that there are a wide range of opportunities available to prepare individuals to work on the assessment. For example, establish internships or short-term employment opportunities in the National Center for Pedagogical Innovation and Educational Research.
3. Introduce a variety of mechanisms to ensure the quality of the assessment. For example, train all assessment administrators according to a protocol and train all scorers to ensure high inter-rater reliability.
4. Introduce a variety of mechanisms for monitoring the consequences of the NLSA, such as expert review groups and a permanent oversight committee. Utilize the outputs of these reviews to inform policy design.

International Large-Scale Assessment (ILSA)

Level of development



Tunisia has participated in TIMSS 1999, 2003, 2007, and 2011, as well as in PISA 2003, 2006, and 2009, and has taken concrete steps to participate in PISA 2012 and 2015. A formal policy document, the *Orientation law on school education and learning* (Law number 80-2002), authorized by the Ministry of Education on July 23, 2002, addresses participation in ILSAs.

Regular funding for participation in ILSAs is sourced from loans and external donors and is allocated at discretion. Funding covers all core activities of the ILSA, including international participation fees, implementation of the assessment exercise in the country, processing and analysis of data collected from implementation of the assessment, attendance at international expert meetings for the assessment exercise, and research and development. However, funding is not provided for reporting and disseminating the assessment results in the country.

A national coordinator and team are responsible for carrying out the ILSA activities. Coordination is carried out by the Director of the Evaluation Department at the National Center for Pedagogical Innovation and Educational Research. ILSA team members have experience working on international assessments, have attended all international workshops and meetings, and have the necessary training and experience to carry out the required assessment activities effectively. However, some members of the team lack English proficiency, and the team is insufficiently staffed. There have been issues with translation of the assessment instruments, and a gap has been observed between the Ministry's annual data collection and the data required by the ILSA organizers.

Tunisia offers some opportunities to learn about ILSAs, including occasional workshops or meetings on using international assessment databases. In addition, Tunisia

provides funding for attending international workshops and training on international assessments, and occasionally makes courses on international assessments available online.

Tunisia-specific results from ILSAs are irregularly disseminated in the country and there is limited media coverage. The country's results for PISA (through 2009) and TIMSS (through 2007) were published during the International Forum on Educational Reform in Tunisia in March 2012. Copies of the national report were distributed to key stakeholders, and brochures and PowerPoint presentations with the country's results were made available online and distributed to key stakeholders. Products that provided information on the results were also made available to schools and educators. However, a national report was not made available online, and copies of the international report were not distributed to key stakeholders. Additionally, the country's results were not communicated to the media through a press release.

Suggested policy options

1. Establish regular government funding for carrying out ILSAs, which is approved by law, decree, or norm.
2. Ensure that the National Center for Pedagogical Innovation and Educational Research is adequately staffed to carry out the assessment effectively. In order to improve the office's research and analytical capacity, review existing opportunities for current and potential future staff, such as students, to learn about and develop skills needed for carrying out ILSAs. For example, collaborate with universities to develop and offer a course on international assessments. In addition, ensure regular funding is available for developing the capacity of current staff, including support for taking online courses on international assessments.
3. Strengthen national ownership of ILSA participation and results by: (i) facilitating Tunisian ILSA staff's greater involvement in the design of questions in relevant ILSA questionnaires (particularly the school, teacher, and student questionnaires) related to key issues in the country's education sector; (ii) establishing a plan for broad, on-time dissemination of ILSA results through, for

example, a variety of media sources; and (iii) developing incentives for national researchers to undertake analysis of ILSA data.

4. In order to use ILSA results for decision making, develop a plan and collaborate with relevant stakeholder groups.

Appendix 1: Assessment Types and Their Key Differences

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

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

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

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

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

Appendix 4: Methodology for Assigning Development Levels

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

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

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

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

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

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

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

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

6. Scores for certain key dimensions under 'Enabling Context' (in the case of EXAM, NLSA, and ILSA) and under 'System Alignment' (in the case of CLASS) were set as ceiling scores, i.e., the overall mean score for the particular assessment type cannot be greater than the score for these key dimensions. These key variables include formal policy, regular funding, having a permanent assessment unit, and the quality of assessment practices.

Appendix 5: SABER-Student Assessment Rubrics for Tunisia





This appendix provides the completed SABER-Student Assessment rubrics for each type of assessment activity in Tunisia. In each row of the rubric, the relevant selection is indicated by a thick border and an asterisk. The selection may include a superscript number that refers to the justification or explanation for the selection. The explanation or justification text can be located in the “Development-level rating justifications” section at the end of each rubric. If a row includes a superscript, but not a thick border and an asterisk, this means that insufficient information was available to determine the relevant selection in the row.

TUNISIA

Classroom Assessment





ENABLING CONTEXT AND SYSTEM ALIGNMENT

Overall policy and resource framework within which classroom assessment activity takes place in a country or system, and the degree to which classroom assessment activity is coherent with other components of the education system.

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT AND SYSTEM ALIGNMENT 1: <i>Setting clear guidelines for classroom assessment</i>			
There is no system-level document that provides guidelines for classroom assessment.	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: <i>Aligning classroom assessment with system learning goals</i>			
There are no system-wide resources for teachers for classroom assessment.	There are scarce system-wide resources for teachers for classroom assessment.	There are some system-wide resources for teachers for classroom assessment. 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: <i>Having effective human resources to carry out classroom assessment activities</i>			
There are no system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.	This option does not apply to this dimension.	There are some system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment. ⁵ *	There are a variety of system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment.

ASSESSMENT QUALITY

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

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

Classroom Assessment: Development-level rating justifications

1. There are formal system-level documents that provide guidelines for classroom assessment. These documents are authorized by the Ministry of Education, and include the *Circular on the Evaluation and Promotion System in the Primary Level of Education* and the *Circular on the Timing of the Examination*, which are both available for the primary level of education.
2. The availability of the documents is restricted. While the documents are available in print-out form to teachers via educational institution directors and in-service courses for teachers, they are not available online, in public libraries, or in teacher training colleges.
3. There are some system-wide resources for teachers to engage in classroom assessment. There is a document, *Programme of Programmes*, which outlines what students are expected to learn in different subject areas at different grade/age levels and what level(s) of performance students are expected to reach in different subject areas at different grade/age levels. Textbooks or workbooks that provide support for classroom assessment are available to teachers in addition to scoring criteria or rubrics for students' work, and item banks or pools with examples of selection/multiple-choice or supply/open-ended questions. However, online assessment resources, or computer-based testing with instant reports on students' performance are not available for teachers' use.
4. The *Programme of Programmes* document is an official curriculum or standards document that specifies what students are expected to learn and to what level of performance. This document covers all levels of education.
5. There are some system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment. These mechanisms include pre-service teacher training offered at universities and special institutes, in-service teacher training, online-resources for classroom assessment, and school inspection of teacher supervision, which includes a component focused on classroom assessment. Additionally, all teacher training programs include a required course on classroom assessment. In-service teacher trainings by pedagogical inspectors and/or pedagogical assistants occur through field visits or attendance at inspector-led training sessions during pedagogical days. Additionally, a teacher's promotion from one rank to another is based on their pedagogical score, assigned once every two years by inspectors. This pedagogical score contains elements of methods for conducting assessments.
6. Classroom assessment practices are known to be of moderate quality. While it is rare to observe errors in the scoring or grading of students' work, it is common for classroom assessment activities to be mainly about recalling information. Although the uneven application of standards for grading students' work is not a serious problem, classroom assessment activities are considered as administrative and monitoring tools, in addition to being pedagogical tools. Additionally, it is not common for classroom assessment activities to mainly rely on multiple-choice questions. Teachers generally use explicit or a priori criteria for scoring or grading students' work, and grade inflation is not a serious problem. Classroom assessment activities are generally aligned with the pedagogical or curricular framework, and parents tend to be informed about students' grades. Teachers also generally provide useful feedback to students on their classroom assessment activities.





7. There are varied and systematic mechanisms in place to monitor the quality of classroom assessment practices. Classroom assessment is a required component of a teacher's performance evaluation and school inspection or teacher supervision, national or other system-wide reviews of the quality of education include a focus on classroom assessment, and government funding is available for research on the quality of classroom assessment activities and how to improve classroom assessment. Many MOE structures, including the General Inspectorate of Pedagogy, the National Center for Pedagogical Innovation and Educational Research, the General Direction for Studies and Planning and Information Systems, and the General Directorate for Assessment and Quality, conduct research on classroom assessment.
8. Classroom assessment information from examinations administered by teachers at the classroom level is required to be disseminated to all key stakeholders, including school district/Ministry of Education officials, parents, and students. Parents and school districts/Ministry of Education officials are informed of individual students' performance through student score cards. However, other classroom assessment information is not required to be disseminated to key stakeholders.
9. There are adequate required uses of classroom assessment to support student learning, including its use as an input for an external examination program, the Baccalaureate examination, diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child's learning, and planning next steps in instruction.

TUNISIA

Examinations

ENABLING CONTEXT

Overall framework of policies, leadership, organizational structures, fiscal and human resources in which assessment activity takes place in a country or system and the extent to which that framework is conducive to, or supportive of, the assessment activity.

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

(CONTINUED)





LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 3: <i>Having regular funding</i>			
There is no funding allocated for the examination.	There is irregular funding allocated for the examination.	There is regular funding allocated for the examination. ⁸ *	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding covers some core examination activities: design, administration, data processing or reporting. ⁹ *	Funding covers all core examination activities: design, administration, data processing and reporting.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development. ¹⁰ *	This option does not apply to this dimension.	Funding covers research and development.
ENABLING CONTEXT 4: <i>Having strong organizational structures</i>			
The examination office does not exist or is newly established.	The examination office is newly established.	The examination office is a stable organization. ¹¹ *	This option does not apply to this dimension.
The examination office is not accountable to an external board or agency. 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 certification or selection system in the country.	Examination results are recognized by one certification or selection system in another country.	Examination results are recognized by two or more certification or selection system in another country. ¹³ *
The examination office does not have the required facilities to carry out the examination.	The examination office has some of the required facilities to carry out the examination.	The examination office has all of the required facilities to carry out the examination. ¹⁴ *	The examination office has state of the art facilities to carry out the examination.

(CONTINUED)

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ENABLING CONTEXT 5: <i>Having effective human resources</i></p>			
<p>There is no staff to carry out the examination.</p>	<p>The examination office is inadequately staffed to effectively carry out the examination, issues are pervasive.</p>	<p>The examination office is adequately staffed to carry out the examination effectively, with minimal issues.</p>	<p>The examination office is adequately staffed to carry out the assessment effectively, with no issues.¹⁵ *</p>
<p>The country does not offer opportunities that prepare for work on the examination.</p>	<p>This option does not apply to this dimension.</p>	<p>The country offers some opportunities that prepare for work on the examination.¹⁶ *</p>	<p>The country offers a wide range of opportunities that prepare for work on the examination.</p>

SYSTEM ALIGNMENT

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

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
<p>SYSTEM ALIGNMENT 1: <i>Aligning examinations with learning goals and opportunities to learn</i></p>			
It is not clear what the examination measures.	This option does not apply to this dimension.	There is a clear understanding of what the examination measures. ¹⁷ *	This option does not apply to this dimension.
What the examination measures is questioned by some stakeholder groups.	This option does not apply to this dimension.	What is measured by the examination is largely accepted by stakeholder groups. ¹⁸ *	This option does not apply to this dimension.
Material to prepare for the examination is minimal and it is only accessible to very few students.	There is some material to prepare for the examination that is accessible to some students.	There is comprehensive material to prepare for the examination that is accessible to most students. *	There is comprehensive material to prepare for the examination that is accessible to all students. ¹⁹ *
<p>SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the examination</i></p>			
There are no courses or workshops on examinations available to teachers.	There are no up-to-date courses or workshops on examinations available to teachers.	There are up-to-date voluntary courses or workshops on examinations available to teachers.	There are up-to-date compulsory courses or workshops on examinations for teachers. ²⁰ *
Teachers are excluded from all examination-related tasks.	Teachers are involved in very few examination-related tasks.	Teachers are involved in some examination-related tasks.	Teachers are involved in most examination-related tasks. ²¹ *

ASSESSMENT QUALITY

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

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

(CONTINUED)

LATENT ● ○ ○ ○	EMERGING ● ● ○ ○	ESTABLISHED ● ● ● ○	ADVANCED ● ● ● ●
ASSESSMENT QUALITY 3: <i>Using examination information in a fair way</i>			
Examination results are not used in a proper way by all stakeholder groups.	Examination results are used by some stakeholder groups in a proper way.	Examination results are used by most stakeholder groups in a proper way. ²⁷ *	Examination results are used by all stakeholder groups in a proper way.
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: <i>Ensuring positive consequences of the examination</i>			
There are no options for students who do not perform well on the examination, or students must leave the education system.	There are very limited options for students who do not perform well on the examination.	There are some options for students who do not perform well on the examination. ²⁹ *	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 Baccalaureate exam is a national examination that students undergo upon completion of high school, and has been administered since 1956. The Baccalaureate exam serves the purposes of selecting students to university or other higher-education institutions and certifying students for grade or school cycle completion, as well as for monitoring education quality levels.
2. The *Decision regulating the Baccalaureate examination* from April 24, 2008 is the formal policy document that authorizes the examination from the Ministry of Education.
3. The policy document is available to the public. It was published in the Official Gazette of the Republic of Tunisia (Number 34, 25 April 2008), and is available on the website of the Ministry of Education.
4. The policy document addresses some key aspects of the examination. It outlines governance, distribution of power, and responsibilities among key entities, describes the purpose of the examination, describes authorized use of results, outlines procedures to investigate and address security breaches, cheating, or other forms of inappropriate behavior, outlines procedures for special/disadvantaged students, specifies who can sit for the examination, and identifies rules about its preparation. However, the policy document does not state funding sources, explain alignment with curricula and standards, or explain the format of the examination questions.
5. All stakeholder groups, including policymakers, teacher unions, educators, students, parents, media, think-tanks and NGOs, universities, and employers, support the examination.
6. There are coordinated attempts to improve the examination by stakeholder groups.
7. Efforts to improve the examination are generally welcomed by the leadership in charge of the examination.
8. There is regular funding allocated for the examination, allocated by the government and from student fees.
9. Funding covers some core examination activities, including design, administration, data reporting, and staff training. However, it does not cover data analysis, long- or medium-term planning of program milestones, or research and development.
10. Funding does not cover research and development.
11. The Directorate General of Examinations in the Ministry of Education has primary responsibility for running the examination since the 1950s.





12. The examination office, which is part of the Ministry of Education, is not accountable to an external board or agency, as it is accountable to the Ministry of Education. The Directorate General of Examinations is still subject to the internal audit of the Inspection Generale at the administrative and financial levels.
13. Examination results are recognized by two or more certification or selection systems in another country. In addition to certification or selection systems in Tunisia, examination results are recognized by certification or selection systems in the Arab states.
14. The examination office has all of the required facilities to carry out the examination, including computers for all technical staff, a secure building, secure storage facilities, access to adequate computer servers, the ability to backup data, and adequate communication tools.
15. The examination office is adequately staffed to carry out the assessment effectively, with no issues.
16. There are very limited opportunities in the country that prepare for work on the examination. Available opportunities include internships in the examination office; however, there are no university graduate programs, university courses, or non-university training courses or workshops on educational measurement and evaluation in Tunisia. In addition, there is no funding for attending international programs, courses, or workshops on educational measurement and evaluation.
17. There is a clear understanding that the examination measures the national school curriculum guidelines and standards.
18. What is measured by the examination is largely accepted by stakeholder groups.
19. There is comprehensive material to prepare for the examination that is accessible by all students in a variety of learning contexts. Material includes examples of the types of questions that are on the examination, information on how to prepare for the examination, and the framework document explaining what is measured on the examination.
20. There are up-to-date compulsory courses or workshops on examinations for teachers.
21. Teachers are involved in most examination-related tasks, including selecting or creating examination questions and scoring guides, administering and scoring the examination, supervising examination procedures, and resolving inconsistencies between examination scores and school grades.
22. There is a comprehensive technical report but with restricted circulation.
23. There are limited systematic mechanisms in place to ensure the quality of the examination, including internal review or observers, external certification or audit, and pilot or field testing. Currently, the Directorate General of Examinations is carrying out a comprehensive audit of all examination operations in line with the quality requirements and improvements, and with the aim of maximizing efficiency and improving related services in cooperation with a quality expert within the framework of French-Tunisian cooperation. However, there are no systematic mechanisms for external review or observers, or translation verification.

24. Inappropriate behavior surrounding the examination process is moderate. Impersonation (when an individual other than the registered candidate takes the examination), copying from other candidates, using unauthorized materials, collusion among candidates, intimidation of examination supervisors or officials, and the provision of external assistance, all occur. However, the use of identity cards to prevent impersonation, the reduction of classroom size to prevent copying from and collusion among candidates, an increased level of security to prevent intimidation of examination supervisors, and the use of a cellphone jammer to hinder the provision of external assistance and collusion among candidates, are mechanisms that have been introduced to address inappropriate behaviors.
25. The examination results are credible for 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 most stakeholder groups in a proper way. For example, data can sometimes be used (not in line with the purposes or uses of the examination) by the media to compare province and national data.
28. Student names and results are public. They are published at the school level, on the internet, and made available via SMS.
29. There are some options for students who do not perform well on the examination. Students may retake the examination after the results are released or repeat a grade.
30. There are very limited mechanisms in place to monitor the consequences of the examination. For example, there are studies that are updated regularly; however, there is no funding for independent research on the impact of the examination, no permanent oversight committee, no regular focus groups or surveys of key stakeholders, and no expert review groups to monitor the consequences of the examination.





TUNISIA
National (or System-Level) Large-Scale Assessment (NLSA)

ENABLING CONTEXT

Overall framework of policies, leadership, organizational structures, fiscal and human resources in which NLSA activity takes place in a country or system and the extent to which that framework is conducive to, or supportive of, the NLSA activity.

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

(CONTINUED)

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT 3: <i>Having regular funding for NLSA</i>			
There is no funding allocated to the NLSA.	There is irregular funding allocated to the NLSA. ⁶ *	There is regular funding allocated to the NLSA.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding covers some core NLSA activities: design, administration, analysis and reporting. *	Funding covers all core NLSA activities: design, administration, analysis and reporting. ⁷ *	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development activities.	This option does not apply to this dimension.	Funding covers research and development activities. ⁸ *
ENABLING CONTEXT 4: <i>Having strong organizational structures for NLSA</i>			
There is no NLSA office, ad hoc unit or team.	The NLSA office is a temporary agency or group of people.	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)

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ENABLING CONTEXT 5: <i>Having effective human resources for NLSA</i></p>			
<p>There is no staff allocated for running an NLSA.</p>	<p>The NLSA office is inadequately staffed to effectively carry out the assessment.¹² *</p>	<p>The NLSA office is adequately staffed to carry out the NLSA effectively, with minimal issues.</p>	<p>The NLSA office is adequately staffed to carry out the NLSA effectively, with no issues.</p>
<p>The country does not offer opportunities that prepare individuals for work on NLSA.</p>	<p>This option does not apply to this dimension.</p>	<p>The country offers some opportunities to prepare individuals for work on the NLSA.¹³ *</p>	<p>The country offers a wide range of opportunities to prepare individuals for work on the NLSA.</p>





SYSTEM ALIGNMENT

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

<p>LATENT ● ○ ○ ○</p>	<p>EMERGING ● ● ○ ○</p>	<p>ESTABLISHED ● ● ● ○</p>	<p>ADVANCED ● ● ● ●</p>
<p>SYSTEM ALIGNMENT 1: <i>Aligning the NLSA with learning goals</i></p>			
<p>It is not clear if the NLSA is based on curriculum or learning standards.</p>	<p>This option does not apply to this dimension.</p>	<p>The NLSA measures performance against curriculum or learning standards.¹⁴ *</p>	<p>This option does not apply to this dimension.</p>
<p>What the NLSA measures is generally questioned by stakeholder groups.</p>	<p>This option does not apply to this dimension.</p>	<p>What the NLSA measures is questioned by some stakeholder groups.¹⁵ *</p>	<p>What the NLSA measures is largely accepted by stakeholder groups.</p>
<p>There are no mechanisms in place to ensure that the NLSA accurately measures what it is supposed to measure.</p>	<p>There are ad hoc reviews of the NLSA to ensure that it measures what it is intended to measure.</p>	<p>There are regular internal reviews of the NLSA to ensure that it measures what it is intended to measure.¹⁶ *</p>	<p>This option does not apply to this dimension.</p>
<p>SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the NLSA</i></p>			
<p>There are no courses or workshops on the NLSA.</p>	<p>There are occasional courses or workshops on the NLSA.¹⁷ *</p>	<p>There are some courses or workshops on the NLSA offered on a regular basis.</p>	<p>There are widely available high quality courses or workshops on the NLSA offered on a regular basis.</p>

ASSESSMENT QUALITY

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

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ASSESSMENT QUALITY 1: <i>Ensuring the quality of the NLSA</i>			
No options are offered to include all groups of students in the NLSA.	This option does not apply to this dimension.	At least one option is offered to include all groups of students in the NLSA. ¹⁸ *	Different options are offered to include all groups of students in the NLSA.
There are no mechanisms in place to ensure the quality of the NLSA.	This option does not apply to this dimension.	There are some mechanisms in place to ensure the quality of the NLSA. ¹⁹ *	There are a variety of mechanisms in place to ensure the quality of the NLSA.
There is no technical report or other documentation about the NLSA.	There is some documentation about the technical aspects of the NLSA, but it is not in a formal report format.	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: <i>Ensuring effective uses of the NLSA</i>			
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. ²²	This option does not apply to this dimension.	NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.	NLSA information is used by all stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.
There are no mechanisms in place to monitor the consequences of the NLSA.	This option does not apply to this dimension.	There are some mechanisms in place to monitor the consequences of the NLSA. 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. In Tunisia, the "Diagnostic assessment of the student's acquisition at the beginning of the seventh grade" was first implemented in the 2012-13 academic year. Its purpose is to monitor education quality at the system level, support schools and teachers, as well as policy design, evaluation, and decision making. The NLSA assesses mathematics, Arabic, and French, and is administered to a representative random sample of students.
2. The Ministry of Education authorized *Orientation Law on Education and School Learning*, Law number 80-2002, on July 23, 2002.
3. The Orientation Law on Education and School Learning is available on the Ministry's website. In addition, hard copies are available in the National Center for Pedagogical Innovation and Educational Research (NCPIERR) and other ministerial departments.
4. There is a written NLSA plan for the coming years specifying who will be tested and in which subject areas. However, the plan is not available to the public.
5. Most stakeholder groups, including policymakers, teacher unions, educators, think-tanks, NGOs, and universities, support the NLSA. Students and parents are neutral to it, and it is unclear what employers' positions are regarding the NLSA.
6. There is irregular funding allocated to the NLSA from the government.
7. Funding covers all core NLSA activities, including design, administration, analysis, reporting, long- or medium-term planning of program milestones, research and development, and staff training.
8. Funding covers research and development activities.
9. The NLSA office, the National Center for Pedagogical Innovation and Educational Research, is a permanent entity created for running the assessment.
10. Political considerations never hamper technical considerations.
11. The NLSA office is accountable to the Scientific Council of the National Center for Pedagogical Innovation and Educational Research.
12. Although there is permanent or full-time staff, it is insufficient to meet the needs of the assessment. Therefore, the center solicits the help of some inspectors and teachers to complete scoring, supervision, data analysis, and field work. There have been some obstacles in carrying out the NLSA, including inspectors' lack of mobility due to security concerns, arising from current difficulties in the country.
13. The country offers some opportunities to prepare individuals for work on the NLSA. Non-university training courses or workshops and funding for attending international training programs, courses, or workshops, on educational measurement and evaluation are available. However, there are no university courses or

graduate programs specifically focused on educational measurement and evaluation. Additionally, there are no internships or short-term employment opportunities in the large-scale assessment office.





14. The NLSA measures performance against curriculum guidelines or learning standards for Arabic, French, and mathematics.
15. What the NLSA measures is questioned by some stakeholder groups.
16. There are regular internal reviews, ad-hoc reviews, and independent reviews by qualified experts of the alignment between the assessment instrument and what it is supposed to measure.
17. There are occasional courses, workshops, or presentations on the NLSA that are made available to teachers.
18. At least one option is offered to include all groups of students in the NLSA. Specifically, the NLSA is offered in the language of instruction for almost all student groups.
19. There are some mechanisms in place to ensure the quality of the NLSA. For example, there is a standardized manual for large-scale assessment administrators, a pilot is conducted before the main data collection takes place, all booklets are numbered, and internal reviewers or observers are employed. However, there are no external reviewers or observers, not all proctors or administrators are trained according to protocol, and scorers are not trained to ensure high inter-rater reliability.
20. There is a comprehensive technical report but with restricted circulation.
21. Unable to tell given that the NLSA began in the 2012-2013 school year.
22. Unable to tell given that the NLSA began in the 2012-2013 school year.
23. Specialized technical committees by subject matter have been established to monitor results of the NLSA. However, other mechanisms, such as expert review groups and permanent oversight committees, do not exist to monitor the consequences of the NLSA.

TUNISIA

International Large-Scale Assessment (ILSA)

ENABLING CONTEXT

Overall framework of policies, leadership, organizational structures, fiscal and human resources in which ILSA takes place in a country or system and the extent to which that framework is conducive to, or supportive of, ILSA activity.

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

(CONTINUED)

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ENABLING CONTEXT 3: <i>Having effective human resources for ILSA</i></p>			
<p>There is no team or national/system coordinator to carry out the ILSA activities.</p>	<p>There is a team or national/system coordinator to carry out the ILSA activities.</p>	<p>There is a team and national/system coordinator to carry out the ILSA activities.⁸ *</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>The national/system coordinator or other designated team member may not be fluent in the language of the assessment.⁹ *</p>	<p>The national/system coordinator is fluent in the language of the assessment.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>The ILSA office is inadequately staffed or trained to carry out the assessment effectively.¹⁰ *</p>	<p>The ILSA office is adequately staffed or trained to carry out the ILSA effectively, with minimal issues.</p>	<p>The ILSA office is adequately staffed and trained to carry out the ILSA effectively, with no issues.</p>





SYSTEM ALIGNMENT

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

<p>LATENT ● ○ ○ ○</p>	<p>EMERGING ● ● ○ ○</p>	<p>ESTABLISHED ● ● ● ○</p>	<p>ADVANCED ● ● ● ●</p>
<p>SYSTEM ALIGNMENT 1: <i>Providing opportunities to learn about ILSA</i></p>			
<p>The ILSA team has not attended international workshops or meetings.</p>	<p>The ILSA team attended some international workshops or meetings.</p>	<p>The ILSA team attended all international workshops or meetings.¹¹</p>	<p>This option does not apply to this dimension.</p>
<p>The country/system offers no opportunities to learn about ILSA.</p>	<p>This option does not apply to this dimension.</p>	<p>The country/system offers some opportunities to learn about ILSA.¹²</p>	<p>The country/system offers a wide range of opportunities to learn about ILSA.</p>
<p>This option does not apply to this dimension.</p>	<p>This option does not apply to this dimension.</p>	<p>Opportunities to learn about ILSA are available to the country's/system's ILSA team members only.¹³</p>	<p>Opportunities to learn about ILSA are available to a wide audience, in addition to the country's/system's ILSA team members.¹³</p>

ASSESSMENT QUALITY

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

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ASSESSMENT QUALITY 1: <i>Ensuring the quality of ILSA</i>			
Data from the ILSA has not been published.	The country/system met sufficient standards to have its data presented beneath the main display of the international report or in an annex.	The country/system met all technical standards required to have its data presented in the main displays of the international report. 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: <i>Ensuring effective uses of ILSA</i>			
If any, country/system-specific results and information are not disseminated in the country/system.	Country/system-specific results and information are disseminated irregularly in the country/system. ¹⁶ *	Country/system-specific results and information are regularly disseminated in the country/system.	Country/system-specific results and information are regularly and widely disseminated in the country/system.
Products to provide feedback to schools and educators about the ILSA results are not made available.	This option does not apply to this dimension.	Products to provide feedback to schools and educators about the ILSA results are sometimes made available.	Products to provide feedback to schools and educators about ILSA results are systematically made available. ¹⁷ *
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. ¹⁹ *	Results from the ILSA are used in some ways to inform decision making in the country/system.	Results from the ILSA are used in a variety of ways to inform decision making in the country/system.
It is not clear that decisions based on ILSA results have had a positive impact on students' achievement levels. ²⁰ *	This option does not apply to this dimension.	This option does not apply to this dimension.	Decisions based on the ILSA results have had a positive impact on students' achievement levels.

International Large Scale Assessment (ILSA): Development-level rating justifications

1. Tunisia has participated in TIMSS in 1999, 2003, 2007, and 2011, as well as PISA in 2003, 2006, and 2009.
2. Tunisia has taken concrete steps to participate in PISA in 2012 and 2015.
3. There is a formal policy document that addresses participation in ILSA. The *Orientation law on school education and learning* (Law number 80-2002) authorized by the Ministry of Education on July 23, 2002, addresses participation in international assessments.
4. The policy document is available to the public on the official website of the Ministry of Education, as well as the website of the official printing house of the Tunisian Republic.
5. There is a regular funding program for international assessment participation, allocated at discretion, and funding sourced from loans and external donors. Funding for international large-scale assessment implementation activities in the country, including booklet translation and printing, training, and transportation, is made available through the budget of the National Pedagogical Innovation and Educational Research Centre within the Ministry of Education. However, funding for international participation is done through loans.
6. Funding covers all core activities of the ILSA, including international participation fees, implementation of the assessment exercise in the country, processing and analyzing data collected from implementation of the assessment exercise, attendance at international expert meetings for the assessment exercise, and research and development. However, funding is not provided for reporting and disseminating the assessment results in the country.
7. Funding covers research and development activities.
8. There is a team and national coordinator to carry out the ILSA activities. Coordination is carried out by the Director of the Evaluation Department at the National Center of Pedagogical Innovation and Educational Research. The assessment section within the Center of Pedagogical Innovation and Educational Research is supported by inspectors and researchers.
9. The national coordinator or other designated team members may not be fluent in the language of the assessment. The lack of English proficiency among all team members poses difficulties.
10. The ILSA office is inadequately staffed to carry out the assessment effectively. Team members have experience working on international assessments, have attended all international workshops or meetings, and have the necessary training or experience to carry out the required assessment activities. However, some team members lack English proficiency, and the team is insufficiently staffed. There have been issues with translation of the assessment instruments and a gap between the Ministry's annual data collection and the data required by organizers of the ILSA.

11. The ILSA team attended all international workshops or meetings.
12. Tunisia offers some opportunities to learn about ILSAs, including offering occasional workshops or meetings on using international assessment databases, providing funding for attending international workshops or training on international assessments, and occasionally making available online courses on international assessments. However, there are no university courses on the topic of international assessments and opportunities are dependent upon the availability of financial resources.
13. Opportunities to learn about ILSA are available to a wide audience, in addition to the country's ILSA team members. For instance, opportunities are available to individuals directly working on the ILSA and professionals or university staff interested in assessment.
14. Tunisia met all technical standards required to have its data presented in the main displays of the international report.
15. Tunisia has not contributed new knowledge on ILSA.
16. ILSA country specific results and information are disseminated irregularly in the country. Country results for PISA (through 2009) and TIMSS (through 2007) were published during the International Forum on Educational Reform in Tunisia in March 2012. Copies of the national report were distributed to key stakeholders, brochures and PowerPoint presentations with the country's results were made available online or distributed to key stakeholders, and products providing feedback to schools or educators about the results were made available as well. However, a national report was not made available online, and copies of the international report were not distributed to key stakeholders. Additionally, country results were not communicated in the media through, for example, a press release, on the television, radio, or in newspapers.
17. Recently, products to provide feedback to schools and educators about ILSA results have become systematically available.
18. There is limited media coverage of the ILSA results. TIMSS and PISA results were presented during the International Forum on Education Reform organized by the Ministry of Education in March 2012. However, the results of international assessments were not published before 2011.
19. Results from the ILSA are used in a limited way to inform decision making in Tunisia. For example, results have been used to inform curriculum improvement and other assessment activities in the country, as well as to allocate resources. However, results from the ILSA are not used for tracking the impact of reforms on student achievement levels or for informing teacher training programs.
20. It is not clear that decisions based on ILSA results have had a positive impact on students' achievement levels.

Acknowledgements

This report, part of a 16-country benchmarking exercise in the Middle East & North Africa and Africa regions, was prepared by the World Bank SABER-Student Assessment team, in partnership with the Arab League Educational, Cultural, and Scientific Organization (ALECSO) and the United Nations Educational, Scientific, and Cultural Organization (UNESCO), which led data-collection efforts. This effort is part of the Arab Regional Agenda for Improving Education Quality (ARAIEQ), led by ALECSO in partnership with the World Bank. It benefited from feedback and review by Kamel Braham, Human Development Sector Coordinator and Task Team Leader for education projects in Tunisia in the World Bank's Education Global Practice, as well as comments received from participants at a national validation workshop held in Tunisia.

References

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

Organization for Economic Cooperation and Development. 2013. *Integrity of Public Education in Tunisia: Restoring Trust*. Paris: OECD.

Republic of Tunisia. 2010. *Economic & Social Development in Tunisia: 2010-2014*. Tunis: Republic of Tunisia.

United Nations Educational, Scientific, and Cultural Organization – Institute for Statistics. 2011. UIS Statistics in Brief. "Education Profile: Tunisia." Montreal: UNESCO-UIS.

World Bank. 2011. *Tunisia - Second Phase of the Education Quality Improvement Program Project*. Washington, DC: World Bank.

-----, 2012. Tunisia Country Indicator Data. Washington, DC: World Bank. Data retrieved from <http://data.worldbank.org/> on October 15, 2013.

-----, 2013. "Tunisia Overview." Washington, DC: World Bank. Data retrieved from <http://www.worldbank.org/en/country/tunisia/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 policymakers and business people—an accessible, objective snapshot showing how well the policies of their country's education system are oriented toward ensuring that all children and youth learn.

This report focuses specifically on policies in the area of **student assessment**.

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

