

Sudan

80072



STUDENT ASSESSMENT

SABER Country Report
2013

Key Policy Areas for Student Assessment

Status

1. Classroom Assessment

In Sudan, the "Guidelines for the Two Levels (Basic and Secondary Education)" document provides guidelines for classroom assessment. Some system-wide resources, such as textbooks and workbooks, are available to teachers for carrying out classroom assessment activities. However, in-service teacher training opportunities to ensure that teachers develop skills and expertise in classroom assessment are limited. Classroom assessment practices are generally considered to be weak, as they provide little useful feedback to students. Limited systematic mechanisms are in place to monitor the quality of classroom assessment practices.

Emerging



2. Examinations

In Sudan, the Secondary Certificate Examination is administered to grade 11 students, with the main purposes of student certification for school cycle completion, student selection to higher-education institutions, and student preparation for the labor market. The examination covers seven academic and seven technical subjects. Regular funding for the examination is allocated by the government and through student fees. The General Directorate of Examinations and Assessment, which is responsible for the examination, is adequately staffed with permanent and full-time staff, and has all of the required facilities to carry out the examination. However, teachers have limited opportunities to learn about the examination, and they are involved in very few examination related tasks. There are also very limited options for students who do not perform well on the examination.

Established



3. National Large-Scale Assessment (NLSA)

The Survey of the State of Education in States (also known as provinces) in Sudan has been conducted twice, in 2009 to a non-representative sample of students in three states and in 2012 to a representative sample of students in four states, with the main purposes of monitoring education quality at the system level, and supporting schools, teachers, policy design, evaluation, and decision making. A formal policy document authorizes the assessment. Funding is provided by external sources such as grants. While no issues have been identified with the performance of the human resources that are responsible for the assessment, mainly temporary and part-time staff work on the NLSA. There are some mechanisms, such as a protocol for training proctors or administrators, in place to ensure the quality of the examination. However, there are no reviews of the alignment between the assessment instrument and what it is supposed to measure.

Emerging



4. International Large-Scale Assessment (ILSA)

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

Latent



Introduction

Sudan has focused on increasing student learning outcomes by improving the quality of education in the country. An effective student assessment system is an important component to improving education quality and learning outcomes as it provides the necessary information to meet stakeholders' decision-making needs. In order to gain a better understanding of the strengths and weaknesses of its existing assessment system, Sudan 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.

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.

Table 1: Framework for building an effective assessment system

	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.

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 Sudan

Sudan is a lower-middle-income country in Sub-Saharan Africa. GDP per capita (current US\$, 2012) is \$1,580 and average annual growth was -4% in 2011 due to loss of oil production, on which the economy had been reliant, and massive devaluation of the Sudanese pound. The 21-year conflict between north and south

Sudan, which ended in 2005, has also had significant economic effects on the country.

During the conflict between north and south Sudan, resources for the education sector were extremely limited. After signing the Comprehensive Peace Agreement (CPA) in 2005 (and prior to the formation of South Sudan in 2011), the Government of National Unity made great efforts to increase access to education, particularly in areas previously affected by conflict. The five year (2007-11) Education Sector Strategic Plan (ESSP) prioritized the country's commitment to the Education for All goals and outlined activities to meet the MDG education targets. There has been a considerable increase in public education spending since 2005. In real terms, total education spending almost doubled, from SDG 1.3 billion to SDG 2.4 billion between 2005 and 2009, equivalent to 16 percent average annual real growth. The positive impact of the CPA has been demonstrated through significant increases in basic school enrollment in Sudan, especially in communities previously affected by conflict. Basic school enrollments for Sudan increased by almost 1 million between 2004/05 and 2008/09, corresponding to average annual growth of 5.7 percent, compared to the pre-CPA period spanning from 2000/01 to 2004/05, when annual enrollments increased 4.2 percent on average. However, regional disparities in access to education persist, especially in rural areas, where girls are most disadvantaged.

After the secession of South Sudan in 2011, the Government of National Unity prepared a new interim Basic Education Strategy (iBES) which explicitly links urgent service delivery to improve access to and the quality of education alongside the development of the education system as a whole, in order to strengthen the foundation for the long-term sustainability of the sector. The expansion of the iBES into a five year Education Sector Strategic Plan (ESSP 2012-2016) will continue as the environment becomes more stable.

Detailed information was collected on Sudan's student assessment system using the SABER-Student Assessment questionnaires and rubrics in 2011. It is important to remember that these tools primarily focus on benchmarking a country's policies and arrangements for assessment activities at the system or macro level.

Additional data would need to be collected to determine actual, on-the-ground practices in Sudan, 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 Sudan's immediate interests and needs. Detailed, completed rubrics for each assessment type in Sudan are provided in Appendix 5.

Classroom Assessment

Level of Development



In Sudan, a formal, system-level document, the *Guidelines for the Two Levels*, authorized by the National Center for Curricula and Educational Research in 2007, provides guidelines for classroom assessment.

In addition, there are some system-wide resources available to teachers for engaging in classroom assessment activities. For example, there are textbooks or workbooks that provide support for classroom assessment, scoring criteria or rubrics for evaluating students' work, and a document that outlines the levels of performance that students are expected to reach in different subject areas at different grade or age levels.

Also, there are some system-level mechanisms in place to ensure that teachers develop skills and expertise in classroom assessment. In-service teacher training opportunities are available; however, few teachers are able to access them. In addition, school inspection or teacher supervision includes a component focused on classroom assessment.

However, classroom assessment practices are known to be weak, particularly because they provide little useful feedback to students, one of the main purposes of classroom assessment. In addition, classroom assessment activities are very commonly about recalling information and rely on multiple-choice/selection-type questions. It is also common for teachers to not use explicit or a priori criteria for scoring or grading students' work. At the same time, it is rare for grade inflation to be a problem or to observe errors in the scoring or grading of students' work.

In addition, limited systematic mechanisms are in place to monitor the quality of classroom assessment practices. Although classroom assessment is a required component of a teacher's performance evaluation, and school inspection or teacher supervision, there are no national reviews of the quality of education that focus on classroom assessment, government funding is not provided to conduct research on the quality of

classroom assessment activities or on how to improve classroom assessment, and an external moderation system that reviews the difficulty of classroom assessment activities and the appropriateness of scoring criteria is not in place.

Although classroom assessment information is required to be disseminated to all key stakeholders, including school district or Ministry of Education officials, parents, and students, there are limited required uses of classroom assessment to support student learning. While classroom assessment is used for diagnosing student learning issues, providing feedback to students on their learning, and providing input to an external examination program, it is not used for planning next steps in instruction or grading students for internal classroom uses, two of the main purposes of classroom assessment.

Suggested policy options:

1. Introduce varied and systematic mechanisms to monitor the quality of classroom assessment practices. For example, allocate government funding for conducting system-wide reviews on the quality of classroom assessment practices in Sudan and for identifying steps on how to improve them.
2. Establish policies which require teachers to use classroom assessment activities for a variety of activities which promote and inform student learning. For example, mandate that classroom assessment activities be used to plan next steps in instruction.
3. Expand opportunities for teachers to develop skills and expertise in conducting classroom assessment activities. For example, review the content and availability of current in-service training opportunities on classroom assessment. Based on the review, further develop the content of the in-service teacher training, and expand the reach of these opportunities to make them accessible to all teachers in Sudan. Additionally, develop pre-service teacher training opportunities on classroom assessment and introduce on-the-job training to ensure that new teachers have the necessary knowledge and skills to conduct classroom assessment activities.

Examinations

Level of Development



The Secondary Certificate Examinations program has been administered since 1956 to grade 11 students, with the main purposes of student certification for school cycle completion, student selection to higher-education institutions, and student preparation for the labor market. The examination covers seven academic and seven technical subjects and was authorized by The Examination Council and the Technical Committees for the Administration of Exams through the *Examination Regulations* and the *Student's Guide to Secondary Examination* documents in 2010.

Regular funding for the examinations is allocated by the government and from student fees. The government finances most of the examination costs although contributions are made by students with financial means. Funding covers all core examination activities, including examination design and administration, data analysis and reporting, and long- or medium-term planning of program milestones. However, funding does not cover staff training or research and development activities.

The General Directorate of Examinations and Assessment, which is an office or branch within the Ministry of Education, has had primarily responsibility for running the examinations since 1950. The General Directorate of Examinations and Assessment is accountable to an external board or agency, the Examinations Council, which includes representatives from universities, education districts, and the committees of education and services in the National Council (The Parliament). In addition, the General Directorate of Examinations and Assessment has all of the required facilities to carry out the examination and is adequately staffed with permanent and full-time staff to carry out the examination effectively, with minimal issues.

Sudan offers a wide range of opportunities that prepare for work on the Secondary Certificate Examinations. For example, there are university graduate programs and

university courses on educational measurement and evaluation at Khartoum University, Zaeem Al Azhari University, Oumdurman Islamic University, Al Jazeera University, and Kassala University (there are also university courses at federal and state level training institutes).

Although there are some workshops available on an irregular basis, there are no up-to-date courses or workshops on examinations available to teachers. At the same time, teachers are involved in very few examination-related tasks, which include selecting or creating examination questions, administering the examination, and supervising examination procedures.

In addition, there is only one systematic mechanism, internal review or observers, in place to ensure the quality of the examination.

There are also very limited options for students who do not perform well on the examinations. While they may retake the examinations or repeat the grade, there are no options for students to attend remedial or preparatory courses in order to prepare to retake the examinations, or to opt for less selective schools, universities, or tracks.

Suggested policy options:

1. Identify funding sources to support research and development activities related to the Secondary Certificate Examination.
2. Introduce opportunities for teachers to learn more about the Secondary Certificate Examination. For example, develop and implement regular, up-to-date workshops on the topic. In addition, provide teachers with opportunities to be involved in a variety of examination-related tasks, such as creating examination scoring guides.
3. Support a variety of options for students who do not perform well on the examination. For example, ensure funding and opportunities for students to attend remedial or preparatory courses in order to prepare to retake the examination.

4. Introduce varied and systematic mechanisms to monitor the consequences of the examination. For example, convene regular focus groups of key stakeholders to understand the consequences of the Secondary Certificate Examinations.

5. Review existing courses on educational measurement and evaluation that prepare for work on the Secondary Certificate Examinations and address any gaps in their content in order to ensure the capacity of existing and incoming staff of the General Directorate of Examinations and Assessment.

National Large-Scale Assessment (NLSA)

Level of Development



The Survey of the State of Education in States (provinces) in Sudan was conducted twice, with the main purposes of monitoring education quality at the system level, supporting schools and teachers, and informing policy design, evaluation, or decision making. These Service Delivery Surveys sought to assess resource availability and utilization in schools, as well as learning levels. It was administered to a non-representative sample of students in grades 4 and 5 in three states in the 2009-10 school year and to a representative sample of students in grades 4 and 5 in four states (including three conflict-affected states) in the 2010-11 school year. The Directorate of Educational Planning in the Ministry of Education authorized the NLSA with the document, *An Evaluation Study for Basic Education in the States of Northern Darfur, the Red Sea, South Kordofan and the Blue Nile*, in 2010.

Irregular funding is allocated to the assessment from grants, supported by the World Bank Sudan multi-donor trust fund and the Education Program Development Fund. However, funding covers all core assessment activities, including assessment design and administration, data analysis and reporting, long- or medium-term planning of program milestones, and staff training. Funding for the assessment does not cover research and development activities.

There was a group of people temporarily assigned to carry out the assessment exercise under the Director Generals for Planning and Teacher Training, which is accountable to a clearly recognized body. While no issues were identified with the performance of the human resources responsible for the large-scale assessment, mainly temporary or part-time staff worked on the NLSA.

Sudan does offer some opportunities to prepare individuals for work on the assessment. For example, university graduate programs and courses on educational measurement and evaluation are available

at all universities with faculties of education, which include Khartoum University, Zaeem Al Azhari University, Omdurman University, Al Jazeera University, and Kassala University. In addition, non-university training courses or workshops on educational measurement and evaluation, as well as funding for attending international programs, courses, or workshops on educational measurement and evaluation are available. However, many of the available opportunities are limited and not of high quality.

Although there are no mechanisms in place to ensure that the assessment accurately measures what it is supposed to measure, there are some mechanisms in place to ensure the quality of the assessment. For example, all proctors or administrators are trained according to a protocol, a pilot is conducted before the main data collection takes place, and scorers are trained to ensure high inter-rater reliability.

There are no mechanisms in place to monitor the consequences of the assessment.

Suggested policy options:

1. Ensure stable government funding for national large-scale assessment activities, which covers all core assessment activities, as well as research and development.
2. With the Ministry of Education providing guidance, develop a plan for effective dissemination of NLA results. For example, include plans for reporting results on subgroups and making results available through media to all stakeholder groups.
3. Formulate a plan within the Ministry of Education to ensure that appropriate mechanisms are in place for using NLA results and for monitoring its consequences in the country. The plan should allow for the engagement of various stakeholders. Include, for example, plans to convene regular focus groups of stakeholders to discuss possible actions on NLA results.
4. Provide funding for research on the impact of the NLA on education in Sudan and collaborate with higher education institutions to conduct such research studies.

5. Ensure the effectiveness of the human resources responsible for carrying out NLSA activities by allocating funding for carrying out training of NLSA staff on assessment design and data analysis on a regular basis.

Box 1: National Learning Assessment

In July 2013, the Global Partnership for Education financed Basic Education Recovery Project (BERP) became effective in Sudan, which will support the establishment of a National Learning Assessment (NLA). The NLA will assess students at the end of grade 3, which is the conclusion of the basic education cycle in Sudan. Students will be assessed in reading, speaking, and listening, through a modified version of the Early Grade Reading Assessment (EGRA). In addition, they will be assessed in concepts in early numeracy. At this time, it is proposed that at one to two rounds of the NLA will take place before the end of 2016, assessing a sample of students in all 15 states. In addition to carrying out the two rounds of the NLA, there will be a strong focus on capacity building within the Ministry of Education to ensure future sustainability of the NLA. A Task Team has been established within the General Directorate of Examinations to design and implement the NLA. In addition, the National Governing Committee (NGC) composed of members of the Ministry of General Education (MoGE) (drawn from Education Planning, National Center for Curriculum and Education Research (NCCER) representatives from State Ministries of Education and Sudan Teachers Union) was established to provide technical support to the development and implementation of the NLA.

International Large-Scale Assessment (ILSA)

Level of Development



Sudan has never participated in an international large-scale assessment and has not taken concrete steps to participate in one in the next five years. In addition, no policy document pertaining to ILSA is available in Sudan.

Suggested policy options:

1. Build momentum around the importance of an internationally comparable, system-level assessment of student achievement levels.
2. Ensure stable government funding for international large-scale assessment activities.
3. Ensure appropriate organizational structures and qualified staff to carry out international large-scale assessment activities.

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 or plans to participate in one.	Participation in an ILSA has been initiated, but there still is need to develop institutional capacity to carry out the ILSA.	There is more or less stable participation in an ILSA. There is institutional capacity to carry out the ILSA. The information from the ILSA is disseminated, but not always used in effective ways.	There is stable participation in an ILSA and institutional capacity to run the ILSA. The information from the ILSA is effectively used to improve education.

Appendix 4: Methodology for Assigning Development Levels

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

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

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

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

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

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

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

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

6. Scores for certain key dimensions under 'Enabling Context' (in the case of EXAM, NLSA, and ILSA) and under 'System Alignment' (in the case of CLASS) were

set as ceiling scores, i.e., the overall mean score for the particular assessment type cannot be greater than the score for these key dimensions. These key variables include formal policy, regular funding, having a permanent assessment unit, and the quality of assessment practices.

Appendix 5: SABER-Student Assessment Rubrics for Sudan

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

SUDAN

Classroom Assessment

ENABLING CONTEXT AND SYSTEM ALIGNMENT

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

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

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

Classroom Assessment: Development-level rating justifications

1. The *Guidelines for the Two Levels* document, authorized by the National Center for Curricula and Educational Research in 2007, provides guidelines for classroom assessment.
2. The *Guidelines for the Two Levels* document is widely available to the public through public libraries.
3. There are some system-wide resources available to teachers for engaging in classroom assessment activities. For example, there are scoring criteria or rubrics for students' work and a document that outlines the levels of performance that students are expected to reach in different subject areas at different grade or age levels. However, there are no item banks or pools with examples of selection /multiple-choice or supply/open-ended questions, online assessment resources, computer-based testing with instant reports on students' performance, or a document that outlines what students are expected to learn in different subject areas at different grade or age levels. Although there are textbooks or workbooks available to teachers to support classroom assessment activities, the textbook to student ratio is as low as 1:5 in some states in Sudan.
4. Curriculum documents and teachers' guides specify what students are expected to learn and to what level of performance.
5. There are some system-level mechanisms in place to ensure that teachers develop skills and expertise in classroom assessment. For example, in-service teacher training opportunities are available, although few teachers are able to access them. In addition, school inspection or teacher supervision includes a component focused on classroom assessment. Other system-level mechanisms — such as teacher training programs including a required course on classroom assessment, pre-service training, online resources on classroom assessment, and opportunities to participate in conferences and workshops — are not in place.
6. Classroom assessment practices are known to be weak, particularly because they provide little useful feedback to students, one of the main purposes of classroom assessment. In addition, classroom assessment activities are very commonly about recalling information and rely on multiple-choice/selection-type questions, and it is common for teachers to not use explicit or a priori criteria for scoring or grading students' work. However, it is rare for grade inflation to be a problem or to observe errors in the scoring or grading of students' work. It is also not common for parents to be poorly informed about students' grades or for the uneven application of standards for grading students' work to be a serious problem. It is rare for classroom assessment activities to be mainly used as an administrative or control tool rather than as a pedagogical resource, and for classroom assessment activities to not be aligned with a pedagogical or curricular framework.
7. There are limited systematic mechanisms in place to monitor the quality of classroom assessment practices. Although classroom assessment is a required component of a teacher's performance evaluation, and school inspection or teacher supervision, there are no national reviews of the quality of education which focus on classroom assessment, government funding is not provided to conduct research on the quality of classroom assessment activities or on how to improve classroom assessment, and an external moderation system which reviews the difficulty of classroom assessment activities and the appropriateness of scoring criteria is not in place.

8. Classroom assessment information is required to be disseminated to all key stakeholders, including school district or Ministry of Education officials, parents, and students.
9. There are limited required uses of classroom assessment to support student learning. While classroom assessment is used for diagnosing student learning issues, providing feedback to students on their learning, and providing input to an external examination program, it is not used for planning next steps in instruction or grading students for internal classroom uses, which are two of the main purposes of classroom assessment.

SUDAN

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. ³ *	The policy document is available to the public.	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. ⁵ *	All stakeholder groups support the examination.
There are no attempts to improve the examination by stakeholder groups.	This option does not apply to this dimension.	There are independent attempts to improve the examination by stakeholder groups. 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.	Funding covers all core examination activities: design, administration, data processing and reporting. ⁹ *	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding does not cover research and development. ¹⁰ *	This option does not apply to this dimension.	Funding covers research and development.
ENABLING CONTEXT 4: <i>Having strong organizational structures</i>			
The examination office does not exist or is newly established.	The examination office is newly established.	The examination office is a stable organization. ¹¹ *	This option does not apply to this dimension.
The examination office is not accountable to an external board or agency.	This option does not apply to this dimension.	The examination office is accountable to an external board or agency. ¹² *	This option does not apply to this dimension.
Examination results are not recognized by any certification or selection system.	Examination results are recognized by certification or selection system in the country.	Examination results are recognized by one certification or selection system in another country. ¹³ *	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)			

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

SYSTEM ALIGNMENT

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

SYSTEM ALIGNMENT			
LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
SYSTEM ALIGNMENT 1: <i>Aligning examinations with learning goals and opportunities to learn</i>			
It is not clear what the examination measures.	This option does not apply to this dimension.	There is a clear understanding of what the examination measures. 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
SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the examination</i>			
There are no courses or workshops on examinations available to teachers.	There are no up-to-date courses or workshops on examinations available to teachers. ²⁰	There are up-to-date voluntary courses or workshops on examinations available to teachers.	There are up-to-date compulsory courses or workshops on examinations for teachers.
Teachers are excluded from all examination-related tasks.	Teachers are involved in very few examination-related tasks. ²¹	Teachers are involved in some examination-related tasks.	Teachers are involved in most examination-related tasks.

ASSESSMENT QUALITY

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

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

(CONTINUED)

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

Examinations: Development-level rating justifications

1. The Secondary Certificate Examinations program has been administered since 1956 to grade 11 students, with the main purposes of student certification for school cycle completion, student selection to higher-education institutions, and student preparation for the labor market. The examination covers seven academic and seven technical subjects from the following:

Academic subjects: Islamic Education, Christian Education, Arabic Language, English Language, French Language, Geography, History, Physics, Chemistry, Biology, Family Sciences, Engineering Sciences, Computer Sciences, Agricultural and Livestock Production, Commercial Sciences, Islamic Studies, Military Sciences, Advanced Arabic language, Specialized Mathematics, Basic Mathematics and Arts.

Technical Subjects: Architectural Drawing, Costs and Taxes, Economy and Trade, Mechanical Drawing, Horticulture, Arts and Design, Finance, Food Processing, Technical Design, Clothing Textiles, History of Art, Nutrition, Decoration, Institutional and Hotel Management, Financial Accounting, Industrial Assets, Agriculture, Drawing and Coloring, and Animal and Agricultural Production.

2. The Examination Council and the Technical Committees for the Administration of Exams authorized the Secondary Certificate Examinations by the *Examination Regulations* and the *Student's Guide to Secondary Examination* documents in 2010.

3. The *Examination Regulations* and the *Student's Guide to Secondary Examinations* documents are not available to the public, although they are available in all schools.

4. The *Examination Regulations* and the *Student's Guide to Secondary Examinations* documents address some key aspects of the examinations, which include an outline of governance, distribution of power, responsibilities among key entities, procedures to investigate and address security breaches, cheating, or other forms of inappropriate behaviors, procedures for special or disadvantaged students, specifications for who can sit for the examination, and an explanation of the format of the examination questions. However, the policy documents do not describe the purposes of the examination or authorized uses of results, state funding sources, identify rules about preparation, or explain alignment of the examinations with curricula and standards.

5. Most stakeholder groups support the examinations. Policymakers, educators, media, and universities strongly support the examinations, while teacher unions, students, parents, think-tanks or NGOs, and employers are neutral about it.

6. Stakeholder groups make coordinated attempts to improve the Secondary Certificate Examinations.

7. Leadership in charge of the Secondary Certificate Examinations generally welcomes efforts to improve the examinations.

8. There is regular funding for the examinations allocated by the government and from student fees. The government finances 70 percent of examination costs, while students with financial means cover the remaining 30 percent. Exceptions are made for students who cannot afford examination fees.

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

10. Funding for the Secondary Certificate Examinations does not cover research and development activities.
11. The General Directorate of Examinations and Assessment, which is an office or branch within the Ministry of Education, has had primarily responsibility for running the examinations since 1950.
12. The General Directorate of Examinations and Assessment is accountable to an external board or agency, the Examinations Council. The Examinations Council includes representatives from universities, education districts, and the committees of education and services in the National Council (The Parliament).
13. Results of the examinations are officially recognized by certification and selection systems in Sudan, and in other countries, including the Kingdom of Saudi Arabia, Egypt, the UAE, and India.
14. The General Directorate of Examinations and Assessment 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 General Directorate of Examinations and Assessment is adequately staffed with permanent and full-time staff to carry out the examination effectively with minimal issues. There have been delays in scoring (due to the fact that the examination is scored mostly by hand) and in reporting results (given needs for their review). In addition, there are some issues in the scoring of essay questions, as they may be assessed differently depending on the scorer. However, none of the following issues related to the performance of human resources responsible for the examinations have been identified: frequent errors in data processing and in examination questions, omission of curricular topics, weaknesses in test design, delays in administering the examination due to issues with the design of examination questions, poor training of test administrators, unclear instruction and guidelines in administering the examinations, or errors in scoring leading to delays in results being reported.
16. Sudan offers a wide range of opportunities that prepare for work on the Secondary Certificate Examinations. For example, there are university graduate programs and university courses on educational measurement and evaluation at Khartoum University, Zaeem Al Azhari University, Oumdurman Islamic University, Al Jazeera University, and Kasiala University (there are also university courses at federal and state level training institutes). Non-university training courses or workshops on educational measurement and evaluation and funding for attending international programs, courses, or workshops on educational measurement and evaluation are made available. However, there are no internships in the examination office.
17. There is a clear understanding that the examinations measure the national school curriculum guidelines or standards.
18. Stakeholder groups largely accept what is measured by the Secondary Certificate Examinations.
19. Comprehensive material to prepare for the examinations is available and accessible to all students. For example, information on how to prepare for the examinations, examples of the types of questions that are on the examinations, and the framework document explaining what is measured on the examinations, are available. However, the report on the strengths and weaknesses in student performance is not available.

20. Up-to-date courses or workshops on examinations are not available to teachers. However, there are some workshops available on an irregular basis.
21. Teachers are involved in very few examination-related tasks. For example, although they are involved in selecting or creating examination questions, administering the examination, and supervising examination procedures, teachers are not involved in selecting or creating examination scoring guides, scoring the examinations, acting as a judge, or resolving inconsistencies between examination scores and school grades.
22. There is a comprehensive technical report but with restricted circulation.
23. There is only one systematic mechanism, internal review or observers, in place to ensure the quality of the examination. External review or observers, external certification or audits, pilot or field testing, or translation verification are not in place.
24. Inappropriate behavior surrounding the examination process is low. While impersonation (when an individual other than the registered candidate takes the examination and copying from other candidates) does occur, leakage of the content of an examination paper or part of a paper prior to the examination, using unauthorized materials such as prepared answers and notes, collusion among candidates via mobile phones or passing of paper, intimidation of examination supervisors, markers, or officials, issuing forged certificates or altering results information, and provision of external assistance via the supervisor or mobile phone do not occur during the examination process. In order to address instances when there is impersonation, students' photographs are affixed to place cards during the examination process.
25. Results of the Secondary Certificate Examinations are credible for all stakeholder groups.
26. All students can take the Secondary Certificate Examinations; there are no language, gender, or other equivalent barriers.
27. All stakeholder groups use examination results in a proper way.
28. Student results are confidential; only students and persons with a legitimate, professional interest in the test taker can know the results.
29. There are very limited options for students who do not perform well on the examinations. While they may retake the examinations or repeat the grade, there are no options for students to attend remedial or preparatory courses in order to prepare to retake the examinations, or to opt for less selective schools, universities, or tracks.
30. There is only one mechanism, expert review groups, in place to monitor the consequences of the examinations. In addition to there being no funding for independent research on the examination, and no permanent oversight committee, there are also no regularly updated studies or regular focus groups or surveys of key stakeholders.

SUDAN

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. ³ *	The policy document is available to the public.	This option does not apply to this dimension.
There is no plan for NLSA activity.	This option does not apply to this dimension.	There is a general understanding that the NLSA will take place.	There is a written NLSA plan for the coming years. ⁴ *
ENABLING CONTEXT 2: <i>Having strong public engagement for NLSA</i>			
All stakeholder groups strongly oppose the NLSA or are indifferent to it.	Some stakeholder groups oppose the NLSA.	Most stakeholders groups support the NLSA.	All stakeholder groups support the NLSA. ⁵ *

(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 3: <i>Having regular funding for NLSA</i>			
There is no funding allocated to the NLSA.	There is irregular funding allocated to the NLSA. ⁶ *	There is regular funding allocated to the NLSA.	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding covers some core NLSA activities: design, administration, analysis and reporting.	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. ⁸ *	This option does not apply to this dimension.	Funding covers research and development activities.
ENABLING CONTEXT 4: <i>Having strong organizational structures for NLSA</i>			
There is no NLSA office, ad hoc unit or team.	The NLSA office is a temporary agency or group of people. ⁹ *	The NLSA office is a permanent agency, institution or unit.	This option does not apply to this dimension.
This option does not apply to this dimension.	Political considerations regularly hamper technical considerations.	Political considerations sometimes hamper technical considerations.	Political considerations never hamper technical considerations. ¹⁰ *
This option does not apply to this dimension.	The NLSA office is not accountable to a clearly recognized body.	The NLSA office is accountable to a clearly recognized body. ¹¹ *	This option does not apply to this dimension.

(CONTINUED)

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

SYSTEM ALIGNMENT

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

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

ASSESSMENT QUALITY

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

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ASSESSMENT QUALITY 1: <i>Ensuring the quality of the NLSA</i>			
No options are offered to include all groups of students in the NLSA. ¹⁸ *	This option does not apply to this dimension.	At least one option is offered to include all groups of students in the NLSA.	Different options are offered to include all groups of students in the NLSA.
There are no mechanisms in place to ensure the quality of the NLSA.	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.	NLSA results are poorly disseminated.	NLSA results are disseminated in an effective way. ²¹ *	This option does not apply to this dimension.
NLSA information is not used or is used in ways inconsistent with the purposes or the technical characteristics of the assessment.	This option does not apply to this dimension.	NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment. ²²	NLSA information is used by all stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment. ²²
There are no mechanisms in place to monitor the consequences of the NLSA. ²³ *	This option does not apply to this dimension.	There are some mechanisms in place to monitor the consequences of the NLSA.	There are a variety of mechanisms in place to monitor the consequences of the NLSA.

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

1. The Survey of the State of Education in States (provinces) in Sudan was conducted twice, with the main purposes of monitoring education quality at the system level, supporting schools and teachers and policy design, evaluation, or decision making. These large-scale assessments were part of Service Delivery Surveys that sought to assess resource availability and utilization in schools, as well as learning levels. They were administered to a non-representative sample of students in grades 4 and 5 in three states in the 2009-10 school year and to a representative sample of students in grades 4 and 5 in four states (including three conflict-affected states) in the 2010-11 school year. The states included were North Kordofan, Nile River, Kessla, North Darfur, South Kordofan, Red Sea, and Blue Nile, and the subjects covered were Mathematics, English language, and Arabic language.
2. The Directorate of Educational Planning in the Ministry of Education authorized the document, *An Evaluation Study for Basic Education in the States of Northern Darfur, the Red Sea, Southern Kordofan and the Blue Nile*, in 2010.
3. The policy document, *An Evaluation Study for Basic Education in the States of Northern Darfur, the Red Sea, Southern Kordofan, and the Blue Nile*, is not available to the public.
4. The *Project to the Design the National Learning Assessment* document is a written NLSA plan for the coming years.
5. All stakeholder groups, including policymakers, teacher unions, educators, students, parents, media, think tanks and NGOs, universities, and employers, support the assessment.
6. There is irregular funding allocated to the assessment from grants, supported by the World Bank Sudan multi-donor trust fund and the Education Program Development Fund.
7. Funding covers all core assessment activities including assessment design and administration, data analysis and reporting, long- or medium-term planning of program milestones, and staff training.
8. Funding for the assessment does not cover research and development activities.
9. There is a group of people temporarily assigned to carry out the assessment exercise under the General Directorate of Planning and General Directorate of Teacher Training.
10. Political considerations never hamper technical considerations, and large-scale assessment results have never been withheld from publication due to political reasons.
11. The group responsible for carrying out the large-scale assessment is accountable to a clearly recognized body.

12. While no issues have been identified with the performance of the human resources that are responsible for the large-scale assessment, mainly temporary or part-time staff work on the NLSA.
13. Sudan offers some opportunities to prepare individuals for work on the assessment. For example, university graduate programs and courses on educational measurement and evaluation are available at all universities with faculties of education, which include Khartoum University, Zaeem Al Azhari University, Oumdurman University, Al Jazeera University, and Kassala University. In addition, non-university training courses or workshops on educational measurement and evaluation, as well as funding for attending international programs, courses, or workshops on educational measurement and evaluation are available. However, many of the available opportunities are limited and not of high quality. At the same time, there are no internships or short-term employment opportunities available in the large-scale assessment office.
14. The assessment measures performance against nationally and internationally recognized curriculum guidelines or learning standards.
15. Stakeholder groups largely accept what the assessment measures.
16. There are no mechanisms in place to ensure that the assessment accurately measures what it is supposed to measure. For example, there are no regular (independent or internal) or ad-hoc reviews of the alignment between the assessment instrument and what it is supposed to measure.
17. Courses or workshops on the assessment are not available to teachers.
18. No options are offered to include all groups of students in the assessment. For example, the assessment is not offered in the language of instruction for all student groups, no accommodations or alternative assessments are provided for students with disabilities, and no special plans are made to ensure that the large-scale assessment is administered to students in hard-to-reach areas.
19. There are some mechanisms in place to ensure the quality of the assessment. For example, all proctors or administrators are trained according to a protocol, 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, scorers are trained to ensure high inter-rater reliability, and there is external certification or audit. However, discrepancies are not required to be recorded on a standard sheet, there is no double scoring or processing of data, and there are no internal or external reviewers or observers.
20. There is a comprehensive technical report on the assessment, which is publicly available.
21. Assessment results are disseminated effectively. Although there are workshops or presentations for key stakeholders on the results, results are featured in newspapers, and media briefings are organized to discuss results, results are not disseminated within twelve months after the large-scale assessment is administered, and reports with results are not made available to all stakeholder groups. The main reports on the results do not contain information on the overall achievement levels and by subgroups, trends over time overall and for subgroups, or standard errors.
22. All stakeholders use the assessment information in a way that is consistent with the purposes and technical characteristics of the assessment.

23. There are no mechanisms in place to monitor the consequences of the assessment. For example, there are no expert review groups, themed conferences that provide a forum to discuss research and other data on the consequences of the large-scale assessment, regular focus groups or surveys of key stakeholders, a permanent oversight committee, or funding for independent research on the impact of the large-scale assessment.

SUDAN

International Large-Scale Assessment (ILSA)

ENABLING CONTEXT

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

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

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

SYSTEM ALIGNMENT

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

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

ASSESSMENT QUALITY

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

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ASSESSMENT QUALITY 1: <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.	This option does not apply to this dimension.
The country/system has not contributed new knowledge on ILSA.	This option does not apply to this dimension.	This option does not apply to this dimension.	The country/system has contributed new knowledge on ILSA.
ASSESSMENT QUALITY 2: <i>Ensuring effective uses of ILSA</i>			
If any, country/system-specific results and information are not disseminated in the country/system.	Country/system-specific results and information are disseminated irregularly in the country/system.	Country/system-specific results and information are regularly disseminated in the country/system.	Country/system-specific results and information are regularly and widely disseminated in the country/system.
Products to provide feedback to schools and educators about the ILSA results are not made available.	This option does not apply to this dimension.	Products to provide feedback to schools and educators about the ILSA results are sometimes made available.	Products to provide feedback to schools and educators about ILSA results are systematically made available.
There is no media coverage of the ILSA results.	There is limited media coverage of the ILSA results.	There is some media coverage of the ILSA results.	There is wide media coverage of the ILSA results.
If any, country/system-specific results and information from the ILSA are not used to inform decision making in the country/system.	Results from the ILSA are used in a limited way to inform decision making in the country/system.	Results from the ILSA are used in some ways to inform decision making in the country/system.	Results from the ILSA are used in a variety of ways to inform decision making in the country/system.
It is not clear that decisions based on ILSA results have had a positive impact on students' achievement levels.	This option does not apply to this dimension.	This option does not apply to this dimension.	Decisions based on the ILSA results have had a positive impact on students' achievement levels.

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

1. Sudan has never participated in an international large-scale assessment and has not taken concrete steps to participate in one in the next five years. This set of facts provides us with sufficient information to determine the development level of ILSA in Sudan, and the remainder of the ILSA rubric is intentionally left blank as a result.
2. Sudan has not taken concrete steps to participate in an ILSA in the next 5 years.
3. No policy document addresses participation in ILSA in Sudan.
4. This option does not apply to this dimension.

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 UNESCO which led data collection efforts. It benefited from feedback and review from Elizabeth Ninan, Senior Education Specialist and Tazeen Fasih, Senior Human Development Economist and Task Team Leader for education projects in Sudan in the World Bank's Africa region, as well as participants of the national validation seminar organized by UNESCO in Sudan.

References

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

United Nations Statistics Division. "Country Profile: Sudan." Data retrieved from <http://data.un.org> on September 11, 2013.

World Bank. 2011. *Main report*. Vol. 1 of *Sudan - Basic Education Project: restructuring*. Washington D.C.: The World Bank.

-----, 2012. *The Status of the Education Sector in Sudan*. Washington, DC: World Bank.

-----, 2013. *Sudan - Basic Education Recovery Project*. Washington DC: World Bank.

-----, 2013. World Bank Development Indicators: Sudan Country Indicator Data. Washington, DC: World Bank. Data retrieved from <http://databank.worldbank.org/data> on September 11, 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.



THE WORLD BANK