



Key Policy Areas for Student Assessment

Status

1. Classroom Assessment

Guidelines for classroom assessment practices are provided in formal, system-level documents that are publicly available. These documents provide information on the type and number of assessments to be conducted by classroom teachers during the school year. Teachers primarily develop their classroom assessment expertise on the job. There has been a concerted effort to improve teachers' skills in classroom assessment, in part through special projects funded by donors. There are insufficient resources currently available to more systematically support teachers' development of skills in classroom assessment. Only ad-hoc mechanisms are being used to monitor the quality of classroom assessment practices, which are generally considered to be weak.



2. Examinations

The examination system has been operating regularly under the guidance of a formal, publicly-available policy document. The Examinations office is part of the Department of Testing and Accreditation, which is accountable to the Ministry of Education. Regular funding for core examination activities is allocated by the government, although the quality of implementation of those activities varies. There are no mechanisms in place to ensure the quality of the examinations.



3. National Large-Scale Assessment (NLSA)

The NLSA program has been operating on an irregular basis. Although there is no policy document that authorizes the NLSA, there is a general understanding that it will take place. The program only has irregular funding, provided by donors, that covers some core NLSA activities. The NLSA office must draw on staff from the National Institute of Educational Sciences and universities to carry out (with the support of international consultants) NLSA activities, and thus is inadequately staffed to effectively carry out these activities on a long-term basis.



4. International Large-Scale Assessment (ILSA)

To date, Vietnam has never participated in an ILSA, but has taken concrete steps to participate in one in the near future by registering for PISA 2012.



Introduction

In 2009, Vietnam joined the Russia Education Aid for Development (READ) Trust Fund program, the goal of which is to help countries improve their capacity to design, carry out, analyze, and use assessments for improved student learning. As part of the READ Trust Fund program, and in order to gain a better understanding of the strengths and weaknesses of its existing assessment system, Vietnam participated in a formal exercise to benchmark this system 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 Vietnam

Vietnam is a lower middle income country in East Asia. GDP per capita (current US\$, 2009) is \$1130, with annual growth of approximately 4.2 percent. At the primary level, gross school enrollment is 104 percent; at the secondary level, it is 76 percent. Gross enrollment for boys and girls is approximately equal at both levels. The percentage of

trained teachers in the country is high at 97 percent, and the student to teacher ratio is relatively low at 20 students per teacher.

Vietnam spent about 5.3 percent of GDP on education in 2008 (latest available data) and approximately 20 percent of the national budget was allocated to education.

Vietnam has made important progress in improving its education system since the 1990s, including increasing the educational attainment of the Vietnamese population, introducing universal enrollment in primary education (the Law on Universal Primary Education that was adopted in 1991 required compulsory basic education (grades 1-5) for all children), significantly increasing completion rates, and improving student achievement levels.

At the same time, Vietnam has identified a number of areas of focus for further development, including improving access to and the quality of education for lower-income students and ethnic minorities; improving learning outcomes, especially for children from disadvantaged groups; improving pedagogical practices; expanding access to school resources; and improving school governance.

Detailed information on Vietnam's student assessment system was obtained from the country's self-diagnosis report that was prepared as part of the READ Trust Fund program in 2009. At the time of the review, SABER-Student Assessment questionnaires had not been developed. Self-diagnosis information was analyzed using the SABER-Student Assessment rubrics. It is important to remember that the SABER-Student Assessment rubrics 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 Vietnam, 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 Vietnam's immediate interests and needs. Detailed, completed rubrics for each assessment type are provided in Appendix 5.

Classroom Assessment

Level of development

Emerging



In Vietnam, the official curriculum sets out assessment policies for teachers to apply in their classrooms. The Ministry of Education and Training provides regulations governing the type and number of assessments to be conducted by teachers during the school year. This document is in the form of a Ministry of Education and Training decision, which has been communicated to all schools.

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. Specifically, the curriculum and outcome standards are listed, but tend to be general in nature and hence rarely used by schools for devising assessment tasks.

There are scarce system-wide resources available to teachers to help them engage in classroom assessment activities. This is because the education budget is primarily built around teacher salaries, and therefore limited resources are available for other things.

There are some system-level mechanisms to ensure that teachers develop skills and expertise in classroom assessment. Limited training in assessment is provided during pre-service teacher training. In general, teachers develop most of their expertise on the job. In the last decade, there has been a concerted campaign to upgrade teachers' skills in assessment, which has been achieved through a set of special projects funded by donors such as Asian Development Bank and the World Bank. Short courses of two to three days have been offered to thousands of teachers in the last three years.

In general, classroom assessment practices are considered to be weak. While there are clear guidelines for conducting classroom assessment activities, primary school teachers in particular need additional training in how to assess students in all subject areas. Secondary school teachers need assistance in how to assess their own subject areas. Donor-supported projects provide assistance in monitoring and building teacher capacity to improve the quality of classroom assessment practices.

Mechanisms are in place to monitor the quality of classroom assessment practices, such as moderation efforts. However, these tend to be conducted on an ad hoc basis.

Classroom assessment information is required to be disseminated to key stakeholders, such as school administrators, education managers, and parents. At the same time, there are limited required uses of classroom assessment information to support student learning.

Suggested policy options:

1. Make available system-level resources for teachers to enable them to engage in quality classroom assessment practices.
2. Introduce more systematic mechanisms to monitor and provide feedback on the quality of classroom assessment practices.

Examinations

Level of development

Emerging



The examinations program in Vietnam includes grade 9 examinations (used for school cycle certification), the *School Leaving Examination* at grade 12 and the *University Entrance Examination*. The *School Leaving Examination* and the *University Entrance Examination* are required by national law. Grade 9 examinations are set by each province. Not all provinces are required to have grade 9 examinations as some make their decisions based on students' achievement records in grades six through nine. The results of the *School Leaving Examination* are used mainly for entry to vocational/professional colleges. They are not used extensively by employers, and are not used by universities. The *School Leaving Examination* and the *University Entrance Examination* guidelines are disseminated every year before the examination.

Regular funding is allocated for the examinations by the government. This funding covers all core examination activities, although not all activities are actually carried out to specification or standards. Funding does not cover research and development, and no systematic research on the validity and reliability of the examination results has been done.

The Examinations office, GDETA, belongs to the Department of Testing and Accreditation and is accountable to the Ministry of Education and Training. The office has 35 full-time professional staff whose main functions with regard to examinations are organizational and supervisory. Vietnam does not offer opportunities that prepare individuals for work on the examinations. Very few of the GDETA staff have had any special training in assessment policies and practices, and they do not actively participate in setting or monitoring the examinations. On-the-job training is the primary way of gaining expertise in assessment. Teachers develop and review the examinations. However, there are no up-to-date courses or workshops on examinations available to teachers. Although teachers receive some training in assessment in their pre-service years, the content is more general and theoretical, and most contemporary assessment

practices are learned on the job. Teams of teachers who undertake the marking of test papers receive special training to standardize procedures.

There is a clear understanding of what the examinations measure. The examination questions are based on an examination framework for each subject that tends to follow the relevant textbook in use in schools rather than the official curriculum. The *School Leaving Examination* is comprised only of multiple choice questions. The *University Entrance Examination* includes a variety of item types, including open ended questions. The questions, however, are not pilot tested. The *University Entrance Examination* is perceived as more rigorous and valid than the grade 12 test and hence is used by universities for their admission decisions. All students can take the examinations; there are no language, gender, or other equivalent barriers.

There are no mechanisms in place, such as pilot testing or pre-testing, to ensure the quality of the examinations. Most of the work in developing the examinations is undertaken by a team of teachers who are solely responsible for ensuring the validity and quality of the examination results. Some technical documentation on the examinations is prepared, but it is not in a formal report format.

Currently, there are no mechanisms in place to monitor the consequences of the examinations.

Suggested policy options:

1. Enhance public confidence in the quality of the examinations by making publicly available a high-quality, independent technical report, and introducing systematic and transparent mechanisms to ensure quality at key stages of the examinations process.
2. Monitor the consequences of the examinations by funding independent research on their impact, creating a permanent oversight committee, or conducting regular quality reviews.

National Large-Scale Assessment (NLSA)

Level of development



National assessments of student achievement levels at grade 5 were conducted in 2001 and 2007. The Asian Development Bank recently funded a project to carry out a national assessment of grade 6 and grade 9 students. While there is no policy document pertaining to NLSA, there is a general understanding that the NLSA will take place.

The NLSA office is a temporary agency or group of people. Specifically, the surveys are done by teams from the Vietnam National Institute of Educational Sciences (VNIES) and local universities, with support from international consultants.

There is irregular funding allocated to the NLSA. The VNIES staff is funded through the government's education budget, while expenses pertaining to the national assessments of student achievement levels have been covered by external donors. The funding covers core NLSA activities (design, administration, analysis, and reporting), as well as research and development activities.

Almost all student groups are included in the NLSA sample. Only very hard to reach areas of the country are excluded. In addition, student with disabilities are not provided any special accommodations.

Vietnam offers some opportunities to prepare individuals for work on the NLSA. More than 20 NLSA team members have attended training courses on national large-scale survey projects organized by the World Bank and the Asian Development Bank. These team members are now the Vietnamese experts in the areas of test design, sampling, data processing, and other technical matters. The training of NLSA team members in the implementation of NLSAs has been mainly provided by donors and foreign experts.

The NLSAs are developed on the basis of the national curriculum. Anchor items were selected from the first NLSA exercise and used in the second NLSA exercise for purposes of comparison. The tests results, when presented publicly, were not opposed, although some questioned the difficulty levels. There are also regular internal reviews of the NLSA to ensure that it measures what it is intended to measure.

There are some mechanisms in place to ensure the quality of the NLSA. In regards to the 2001 and 2007 assessments, the anchor items were pretested, and questions appear to have been given careful scrutiny to eliminate ambiguities and other potential weaknesses. The NLSAs included items assessing students' ability to make creative applications, and items measuring problem-solving skills. The tests and questionnaires included questions from recent international surveys to provide a basis for comparing results with other countries. The tests were trialed and revised. However, the final tests omitted several key areas in the curriculum, and appeared to include many relatively easy questions.

There is a comprehensive, high-quality technical report that was made available to the public. However, NLSA results are poorly disseminated.

Some mechanisms are in place to monitor the consequences of the NLSA, such as workshops and expert reviews.

Suggested policy options:

1. Institutionalize the NLSA by developing a formal policy document that authorizes the NLSA program, and by ensuring regular government funding to cover all core NLSA activities.
2. Create a permanent, adequately-staffed office responsible for NLSA activities and accountable to an external body.

International Large-Scale Assessment (ILSA)

Level of development



Vietnam has not participated in an ILSA in the last 10 years. However, it has registered for PISA 2012. At the moment, there is no policy document that addresses participation in ILSAs, and there is no funding for participation in ILSAs.

A team and a national coordinator have been identified to carry out ILSA activities. The ILSA office is newly established and thus has only a few permanent staff. The ILSA team has attended some international workshops and meetings. Some opportunities to learn about ILSA have been made available in the country to ILSA team members only.

Suggested policy options:

1. Provide a stable organizational structure within Vietnam for the ILSA program, and introduce a policy document that addresses participation in, and the allocation of government funds for, ILSA activities.
2. Build capacity within Vietnam to carry out ILSA activities, including ensuring that the national coordinator is fluent in the international language of the assessment (English, in the case of PISA) and providing the necessary training to ILSA staff so that they can carry out ILSA activities effectively and with minimal issues.

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				Justification
	LATENT (Absence of, or deviation from, attribute)	EMERGING (On way to meeting minimum standard)	ESTABLISHED (Acceptable minimum standard)	ADVANCED (Best practice)	
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

	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 Vietnam

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

VIETNAM
Classroom Assessment

ENABLING CONTEXT AND SYSTEM ALIGNMENT

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

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

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

Classroom Assessment: Development level rating justifications

1. The official curriculum sets out assessment policies for teachers to apply in their classrooms. MOET has a set of regulations governing the type and number of assessments to be made by teachers in the school year.
2. The document is in the form of MOET's decision which has been communicated to all schools.
3. The budget for education is primarily to cover teacher salaries; hence, little is left for professional development and teacher support.
4. The curriculum and outcome standards are specified, but tend to be general and hence rarely used by schools for devising assessment tasks.
5. While limited training in assessment is provided during pre-service teacher training, teachers develop their expertise mostly on the job. In the last decade, there has been a concerted campaign to upgrade teachers' skills in assessment, which has been achieved by a set of special projects funded by donors such as ADB or the World Bank. Short courses of 2-3 days in duration have been offered to thousands of teachers in the last three years.
6. While there are clear guidelines on conducting classroom assessment, primary school teachers in particular need additional training in how to assess all subject areas. Secondary teachers need assistance in how to assess their own subject areas. Donor-supported projects provide assistance in monitoring and building teacher capacity to improve the quality of classroom assessment.
7. There are some moderation efforts but they not systematic.
8. School administrators, education managers and parents receive students' grades.
9. There are clear guidelines on the processes of conducting classroom assessment and on using the assessment results to support student learning. Classroom assessment, however, is not an input for external examination results at the Grade 12 level, while it is for the Grade 9 examinations.

VIETNAM *Examinations*

ENABLING CONTEXT

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

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

(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 regular funding</i></p>			
<p>There is no funding allocated for the examination.</p>	<p>There is irregular funding allocated for the examination.</p>	<p>There is regular funding allocated for the examination.⁴</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Funding covers some core examination activities: design, administration, data processing or reporting.</p>	<p>Funding covers all core examination activities: design, administration, data processing and reporting.⁵</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Funding does not cover research and development.⁶</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>	<p>Funding covers research and development.</p>
<p>ENABLING CONTEXT 4: <i>Having strong organizational structures</i></p>			
<p>The examination office does not exist or is newly established.</p>	<p>The examination office is newly established.</p>	<p>The examination office is a stable organization.⁷</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>The examination office is not accountable to an external board or agency.⁸</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>	<p>The examination office is accountable to an external board or agency.</p>	<p>This option does not apply to this dimension.</p>
<p>Examination results are not recognized by any certification or selection system.</p>	<p>Examination results are recognized by certification or selection system in the country.⁹</p> <p style="text-align: right;">*</p>	<p>Examination results are recognized by one certification or selection system in another country.</p>	<p>Examination results are recognized by two or more certification or selection system in another country.</p>
<p>The examination office does not have the required facilities to carry out the examination.</p>	<p>The examination office has some of the required facilities to carry out the examination.¹⁰</p> <p style="text-align: right;">*</p>	<p>The examination office has all of the required facilities to carry out the examination.</p>	<p>The examination office has state of the art facilities to carry out the examination.</p>

(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.¹¹ *</p>	<p>The examination office is adequately staffed to carry out the assessment effectively, with no issues.</p>
<p>The country does not offer opportunities that prepare for work on the examination.¹² *</p>	<p>This option does not apply to this dimension.</p>	<p>The country offers some opportunities that prepare for work on the examination.</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.

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>SYSTEM ALIGNMENT 1: <i>Aligning examinations with learning goals and opportunities to learn</i></p>			
<p>It is not clear what the examination measures.</p>	<p>This option does not apply to this dimension.</p>	<p>There is a clear understanding of what the examination measures.¹³</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>What the examination measures is questioned by some stakeholder groups.</p>	<p>This option does not apply to this dimension.</p>	<p>What is measured by the examination is largely accepted by stakeholder groups.¹⁴</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>
<p>Material to prepare for the examination is minimal and it is only accessible to very few students.</p>	<p>There is some material to prepare for the examination that is accessible to some students.</p>	<p>There is comprehensive material to prepare for the examination that is accessible to most students.¹⁵</p> <p style="text-align: right;">*</p>	<p>There is comprehensive material to prepare for the examination that is accessible to all students.</p>
<p>SYSTEM ALIGNMENT 2: <i>Providing teachers with opportunities to learn about the examination</i></p>			
<p>There are no courses or workshops on examinations available to teachers.</p>	<p>There are no up-to-date courses or workshops on examinations available to teachers.¹⁶</p> <p style="text-align: right;">*</p>	<p>There are up-to-date voluntary courses or workshops on examinations available to teachers.</p>	<p>There are up-to-date compulsory courses or workshops on examinations for teachers.</p>
<p>Teachers are excluded from all examination-related tasks.</p>	<p>Teachers are involved in very few examination-related tasks.</p> <p style="text-align: right;">*</p>	<p>Teachers are involved in some examination-related tasks.</p>	<p>Teachers are involved in most examination-related tasks.</p>

ASSESSMENT QUALITY

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

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ASSESSMENT QUALITY 1: <i>Ensuring quality</i></p>			
<p>There is no technical report or other documentation.</p>	<p>There is some documentation on the examination, but it is not in a formal report format. *</p>	<p>There is a comprehensive technical report but with restricted circulation.</p>	<p>There is a comprehensive, high quality technical report available to the general public.</p>
<p>There are no mechanisms in place to ensure the quality of the examination.¹⁷ *</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.</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.¹⁸ *</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.¹⁹ *</p>	<p>The examination results are credible for all stakeholder groups.</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.²⁰ *</p>

(CONTINUED)

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ASSESSMENT QUALITY 3: <i>Using examination information in a fair way</i></p>			
<p>Examination results are not used in a proper way by all stakeholder groups.</p>	<p>Examination results are used by some stakeholder groups in a proper way.</p>	<p>Examination results are used by most stakeholder groups in a proper way.²¹</p> <p style="text-align: right;">*</p>	<p>Examination results are used by all stakeholder groups in a proper way.</p>
<p>Student names and results are public.²²</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>	<p>Students' results are confidential.</p>	<p>This option does not apply to this dimension.</p>
<p>ASSESSMENT QUALITY 4: <i>Ensuring positive consequences of the examination</i></p>			
<p>There are no options for students who do not perform well on the examination, or students must leave the education system.</p>	<p>There are very limited options for students who do not perform well on the examination.²³</p> <p style="text-align: right;">*</p>	<p>There are some options for students who do not perform well on the examination.</p>	<p>There is a variety of options for students who do not perform well on the examination.</p>
<p>There are no mechanisms in place to monitor the consequences of the examination.</p> <p style="text-align: right;">*</p>	<p>This option does not apply to this dimension.</p>	<p>There are some mechanisms in place to monitor the consequences of the examination.</p>	<p>There is a variety of mechanisms in place to monitor the consequences of the examination.</p>

Examinations: Development level rating justifications

1. The Grade 12 examinations are required by law, and are therefore part of the legal framework. Grade 9 examinations, however, are set by each province, and not all provinces are required to have the grade 9 examinations (as some provinces make their selections entirely based on the students' achievement record in grades six through nine).
2. Grade 12 and university entrance exam guidelines are produced and disseminated every year before the exam date.
3. There are independent attempts to improve the examinations, including by policy makers and the media. There is a lot of debate on how best to improve the examinations.
4. Regular funding is allocated for the examination by the government.
5. Funding covers all core examination activities, although not all activities are actually done to specification/standards.
6. No systematic research on the validity and reliability of the exam results has been done.
7. The Examination office belongs to the Department of Testing and Accreditation.
8. GDETA is part of MOET, to which it is accountable. Therefore, GDETA is not accountable to an external board or agency.
9. In Vietnam, the results of the Grade 12 School Leaving Examination are used mainly for entry to vocational/professional colleges, are not used extensively by employers, and are not used at all by universities. It is unclear whether the examination results are recognized in another country.
10. The Examination office has scanning and computing capacity for administering and recording exam results.
11. The examination office has 35 full-time professional staff whose main functions with regard to examinations are organizational and supervisory. Staffing is stable, with only two staff leaving in the last ten years.
12. "Very few of the staff of GDETA, however, have had any special training in assessment policies and practices, and they do not participate actively in setting or monitoring the examinations. On-the-job training is the primary way of gaining expertise in assessment and teams of teachers are depended upon to develop and review the examinations. "
13. The examination questions are based on an examination framework for each subject that tends to follow the relevant textbook in use in schools rather than the official curriculum. Additionally, the Grade 12 School Leaving Examination is comprised of only multiple choice questions, and topics not readily examined

through multiple-choice questions, such as oral expression, listening, writing, practical skills, and project work, are not included in the examination. The University Entrance Examination, however, includes a variety of item types, including open ended questions. The questions, however, are not pilot tested and time limits are not tested.

14. Few stakeholders question the exam content.

15. While nearly all schools have access to the required textbooks, many students from minority ethnic groups (of which there are 54) are disadvantaged because the examinations have to be taken in the Vietnamese language. These students are also handicapped because their teachers tend to be less qualified and their resources are leaner.

16. Although teachers do receive some training in assessment in their pre-service years, the content is more general and theoretical, and most current assessment practices are learned on the job. Additionally, teams of teachers who undertake the marking of test papers receive special training to standardize procedures. The training varies from three days to a week.

17. Most of the work in developing the examinations is undertaken by a team of teachers who are solely responsible for ensuring the validity and quality of the examination. No trial testing or pre-testing is done.

18. While the examinations are carefully supervised, cheating does occur and many staff have minimal faith in the security policies in place for the Grade 12 examinations.

19. The University entrance exam is perceived as more rigorous and valid than Grade 12 graduation and hence is used by universities for their admission decisions.

20. All students study in Vietnamese from Lower secondary onwards; the gender ratio seems to be balanced amongst exam takers.

21. Examinations results are used for certifying high school leaving and making university admission decisions.

22. Student names and results are published on the school bulletin board.

23. There is a provision for recounts where students feel that they deserve a better score on the examination. Additionally, students may retake the examination the following year.

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

ENABLING CONTEXT

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

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

(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 regular funding for NLSA</i></p>			
<p>There is no funding allocated to the NLSA.</p>	<p>There is irregular funding allocated to the NLSA.³</p> <p style="text-align: right;">*</p>	<p>There is regular funding allocated to the NLSA.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Funding covers some core NLSA activities: design, administration, analysis and reporting.⁴</p> <p style="text-align: right;">*</p>	<p>Funding covers all core NLSA activities: design, administration, analysis and reporting.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Funding does not cover research and development activities.</p>	<p>This option does not apply to this dimension.</p>	<p>Funding covers research and development activities.⁵</p> <p style="text-align: right;">*</p>
<p>ENABLING CONTEXT 4: <i>Having strong organizational structures for NLSA</i></p>			
<p>There is no NLSA office, ad hoc unit or team.</p>	<p>The NLSA office is a temporary agency or group of people.⁶</p> <p style="text-align: right;">*</p>	<p>The NLSA office is a permanent agency, institution or unit.</p>	<p>This option does not apply to this dimension.</p>
<p>This option does not apply to this dimension.</p>	<p>Political considerations regularly hamper technical considerations.</p>	<p>Political considerations sometimes hamper technical considerations.</p>	<p>Political considerations never hamper technical considerations.⁷</p> <p style="text-align: right;">*</p>
<p>This option does not apply to this dimension.</p>	<p>The NLSA office is not accountable to a clearly recognized body.⁸</p> <p style="text-align: right;">*</p>	<p>The NLSA office is accountable to a clearly recognized body.</p>	<p>This option does not apply to this dimension.</p>

(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.⁹</p> <p style="text-align: right;">*</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.⁹</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.¹⁰</p> <p style="text-align: right;">*</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>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</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.¹¹</p> <p style="text-align: right;">*</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.</p>	<p>What the NLSA measures is largely accepted by stakeholder groups.¹²</p> <p style="text-align: right;">*</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.</p> <p style="text-align: right;">*</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.</p>	<p>There are some courses or workshops on the NLSA offered on a regular basis.¹³</p> <p style="text-align: right;">*</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.

<p>LATENT</p> <p>● ○ ○ ○</p>	<p>EMERGING</p> <p>● ● ○ ○</p>	<p>ESTABLISHED</p> <p>● ● ● ○</p>	<p>ADVANCED</p> <p>● ● ● ●</p>
<p>ASSESSMENT QUALITY 1: <i>Ensuring the quality of the NLSA</i></p>			
<p>No options are offered to include all groups of students in the NLSA.</p>	<p>This option does not apply to this dimension.</p>	<p>At least one option is offered to include all groups of students in the NLSA.¹⁴</p> <p style="text-align: right;">*</p>	<p>Different options are offered to include all groups of students in the NLSA.</p>
<p>There are no mechanisms in place to ensure the quality of the NLSA.</p>	<p>This option does not apply to this dimension.</p>	<p>There are some mechanisms in place to ensure the quality of the NLSA.¹⁵</p> <p style="text-align: right;">*</p>	<p>There are a variety of mechanisms in place to ensure the quality of the NLSA.</p>
<p>There is no technical report or other documentation about the NLSA.</p>	<p>There is some documentation about the technical aspects of the NLSA, but it is not in a formal report format.</p>	<p>There is a comprehensive technical report but with restricted circulation.</p>	<p>There is a comprehensive, high quality technical report available to the general public.¹⁶</p> <p style="text-align: right;">*</p>
<p>ASSESSMENT QUALITY 2: <i>Ensuring effective uses of the NLSA</i></p>			
<p>NLSA results are not disseminated.</p>	<p>NLSA results are poorly disseminated.</p> <p style="text-align: right;">*</p>	<p>NLSA results are disseminated in an effective way.</p>	<p>This option does not apply to this dimension.</p>
<p>NLSA information is not used or is used in ways inconsistent with the purposes or the technical characteristics of the assessment.</p>	<p>This option does not apply to this dimension.</p>	<p>NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.</p> <p style="text-align: right;">*</p>	<p>NLSA information is used by all stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.</p>
<p>There are no mechanisms in place to monitor the consequences of the NLSA.</p>	<p>This option does not apply to this dimension.</p>	<p>There are some mechanisms in place to monitor the consequences of the NLSA.¹⁷</p> <p style="text-align: right;">*</p>	<p>There are a variety of mechanisms in place to monitor the consequences of the NLSA.</p>

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

1. National assessments of student achievement levels at Grade 5 were conducted in 2001 and in 2007. The Asian Development Bank recently funded a project to carry out a national assessment of Grade 6 and Grade 9 students.
2. There is no opposition to date although some provinces were anxious when seeing their results below the national average.
3. The VNIES staff are funded through the Education budget (paid for by the government), while expenses pertaining to the national assessments of student achievement levels have been covered by external donors, such as the World Bank, ADB, and Inrec/Mekong Economics.
4. The funding from donors covers some core NLSA activities.
5. Funding covers research and development. For example, funding covers design and anchoring between surveys and data analysis.
6. The surveys are done by teams from the National Institute of Educational Sciences and universities, with support from international consultants.
7. Political considerations never hamper technical considerations. For example, for the Grade 5 survey, the sample was constructed so that it was representative at both the national and provincial levels.
8. The NLSA office has not been officially established.
9. Internally, assistance for conducting the NLSA was provided by GDETA staff, Primary Education Department officials, Department staff in the provinces, and foreign consultants. The tests and questionnaires used in the survey were developed by teams of local educators, assisted by foreign consultants. After marking, recording and cleaning, the merging of results and further analyses were undertaken largely by foreign consultants. Some of the VNIES staff were trained in Europe to assist with the analyses. While the main implementers of the project were VNIES staff, the final reports were largely written by World Bank staff.
10. More than 20 of the staff have attended training courses organized by the World Bank and the ADB, especially for national large-scale survey projects, and these staff are now the Vietnamese experts in the areas of test design, sampling, data processing, and other technical matters. The training of human resources for implementation of national large-scale assessments has been primarily provided externally by donors and foreign experts.
11. The NLSA was developed on the basis of the National Curriculum. Anchor items were selected from the first NLSA exercise and used in the second NLSA exercise for purposes of comparison. The tests results, when presented, were not opposed though some questioned the difficulty levels.
12. The World Bank, IIEP, University of Melbourne and others provided training to Vietnam on large-scale surveys.

13. Almost all student groups are included in the NLSA sample. Only very hard to reach areas are not included and student with disabilities are not provided any special accommodations.
14. The anchor items were pretested, and questions appear to have been given careful scrutiny to eliminate ambiguities and other potential weaknesses. The
15. NLSA included items assessing students' ability to make creative applications, and items measuring problem-solving skills. The tests and questionnaires included questions from recent international surveys to provide a basis for comparing results with other countries. The tests were trialed and revised. However, the final test omitted several key areas in the curriculum, and appeared to include many relatively easy questions.
16. A volume on technical aspects of the surveys was compiled and made public.
17. Workshops and expert reviews are held.

VIETNAM
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.
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. ¹ *	This option does not apply to this dimension.
There is no policy document that addresses participation in ILSA. *	There is an informal or draft policy document that addresses participation in ILSA.	There is a formal policy document that addresses participation in ILSA.	This option does not apply to this dimension.
This option does not apply to this dimension.	The policy document is not available to the public.	The policy document is available to the public.	This option does not apply to this dimension.
ENABLING CONTEXT 2: <i>Having regular funding for ILSA</i>			
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.	Funding covers all core activities of the ILSA.
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. ² *	The ILSA office is adequately staffed or trained to carry out the ILSA effectively, with minimal issues.	The ILSA office is adequately staffed and trained to carry out the ILSA effectively, with no issues.

SYSTEM ALIGNMENT

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

LATENT	EMERGING	ESTABLISHED	ADVANCED
SYSTEM ALIGNMENT 1: <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. ³ *	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. ⁴ *	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.	The country/system met all technical standards required to have its data presented in the main displays of the international report.
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. Vietnam has registered for PISA 2012.
2. The office is newly established with few permanent staff.
3. There are some opportunities to learn about the ILSA.
4. The opportunities to learn about the ILSA are available only to the ILSA team members.

Acknowledgements

This report was prepared by The World Bank SABER-Student Assessment team in collaboration with Thanh Thi Mai, World Bank Senior Education Specialist and Task Team Leader for education projects in Vietnam.

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 Children's Fund (UNICEF). "Vietnam Education Overview." Data retrieved from http://www.unicef.org/vietnam/girls_education.html on April 11, 2013.

United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics. Vietnam Country Indicator Data. Montreal, QC: UNESCO. Data retrieved from <http://stats.uis.unesco.org/unesco> on March 12, 2013.

World Bank. 2013. "Vietnam at a Glance." Available online: Data retrieved from http://devdata.worldbank.org/AAG/vnm_aag.pdf on April 11, 2013.

———. 2011. "Vietnam: High Quality Education for All by 2020." Washington, DC: World Bank.

World Bank Development Indicators. Vietnam Country Indicator Data. Washington, DC: World Bank. Data retrieved from <http://databank.worldbank.org/data> on March 12, 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.

