



The Former Yugoslav Republic of Macedonia

STUDENT ASSESSMENT

SABER Country Report
2012

Key Policy Areas for Student Assessment

Status

1. Classroom Assessment

In FYR Macedonia, there are several formal system-level documents that provide guidelines for classroom assessment. While there are some system-wide resources and materials available to teachers for carrying out classroom assessment activities, there are limited opportunities available to them for learning about or developing more effective classroom assessment practices. In general, classroom assessment practices are considered to be weak, and there are limited systematic mechanisms in place to monitor their quality.

Emerging
●●○○

2. Examinations

The State Matura has been administered every year since 2008 to grade 12 students. Results of the State Matura are used for certifying grade completion, determining admission to university and other higher education institutions, monitoring education quality levels, and planning education policy reforms. Funding for the State Matura is provided by the government to the National Examination Center. Currently, there are no up-to-date courses or workshops on the State Matura available to classroom teachers.

Emerging
●●○○

3. National Large-Scale Assessment (NLSA)

The External Assessment of Students' Achievement in Primary and Secondary Education ("External Assessment") was piloted in different grades and subjects in 2010 and 2011. The Ministry of Education and Science announced that the External Assessment would be formally launched in the 2012-2013 school year. The official purpose of the External Assessment is to evaluate the objectivity of teachers' grading.

Latent
●○○○

4. International Large-Scale Assessment (ILSA)

FYR Macedonia has participated in a number of ILSA exercises, including PIRLS (2001, 2006), TIMSS (1999, 2003, 2011), PISA (2000), and PISA Plus (2001). However, there is no policy document that systematically addresses the country's participation in international assessments. Most of the funding for FYR Macedonia's participation in ILSA exercises has been provided by donors, including the World Bank. Opportunities to learn about ILSA are available only to individuals working directly on a specific ILSA exercise.

Emerging
●●○○



Introduction

In 2001, FYR Macedonia embarked on a new education policy agenda. One of the goals of this agenda was to improve the country's student assessment system. In order to gain a better understanding of the strengths and weaknesses of its existing assessment system, FYR Macedonia decided to benchmark this system using standardized tools developed under The World Bank's Systems Approach for Better Education Results (SABER) program. SABER is an evidence-based program to help countries systematically examine and strengthen the performance of different aspects of their education systems.

What is SABER-Student Assessment?

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

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

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

SABER-Student Assessment methodology

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

Assessment types and purposes

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

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

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

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

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

Quality drivers of an assessment system

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

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

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

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

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

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

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

Source: World Bank.

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

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

Levels of development

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

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

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

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

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

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

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

Education in FYR Macedonia

FYR Macedonia is an upper-middle-income country in Eastern Europe. GDP per capita (current US\$) is \$4,925, with annual growth of about 2.9 percent. While unemployment is very high at 31.4 percent (2011), education is one of the main sectors experiencing employment growth, particularly in the second quarter of 2012.

Primary and secondary education is compulsory in FYR Macedonia. Beginning in the 2007-2008 school year, children start school at the age of 6. Also since the 2007-2008 school year, the length of the primary education cycle increased from eight to nine years. As of 2009, school enrollment at the primary level is 90 percent and the completion rate is 92 percent.

Investments have been made to improve the physical infrastructure for education, to increase the number of teachers (particularly in primary and lower-secondary schools), and to modernize the curricula. Additionally, FYR Macedonia has focused on improving education outcomes. FYR Macedonia's performance on international large-scale assessments (TMSS, PISA, and PIRLS) in 2000 and 2001 prompted several education reforms in this regard. As a result, between 2004 and 2008, enrollment in secondary schools increased from 85 to 95 percent, and dropout rates in primary school decreased to below 2 percent. To improve teacher performance, training opportunities have been made available, and specifications for the accreditation and monitoring of teacher training have been introduced. These and other education reforms have led to school managers and teachers reporting improvements in key areas of education. Specifically, from 2004 to 2007, almost half of school managers and teachers reported improvements in student achievement while nearly two thirds observed improvements in planning and assessment processes.

Detailed information was collected on FYR Macedonia's student assessment system using the SABER-Student Assessment questionnaires and rubrics. It is important to remember that these tools primarily focus on benchmarking a country's policies and arrangements for assessment activities at the system or macro level. Additional data would need to be collected to determine actual, on-the-ground practices in FYR Macedonia, 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 FYR Macedonia's immediate interests and needs. Detailed, completed rubrics for each assessment type in FYR Macedonia are provided in Appendix 5.

Classroom Assessment

Level of development

Emerging



There are several formal system-level documents that provide guidelines for classroom assessment. The *Teaching Curricula by Subject and Grades* (2001-2010) series are official documents authorized by the Bureau for Development of Education in the Ministry of Education that specify what students are expected to learn. However, these and other available documents do not specify the level of performance that students are expected to demonstrate in relation to what they have learned.

There are some system-wide resources/materials available to teachers for carrying out classroom assessment activities. For example, the *Teaching Curricula by Subject and Grades* documents outline what students are expected to learn in different subject areas at different grade/age levels and the *Assessment Criteria* document provides the level(s) of performance that students are expected to reach in different subject areas at different grade/age levels. The Ministry of Education website provides examples of good assessment practices and case studies on classroom assessment. An on-line item bank, containing some items that were previously administered on the *State Matura* and on the *National Assessment*, was recently added to this website (posted in June 2011), although the number of items in the item bank is very limited.

There are limited opportunities available to teachers for learning about or developing more effective classroom assessment practices. A small number of courses on classroom assessment are offered during pre-service teacher training. In the case of in-service teacher training, some courses on classroom assessment have been offered on an ad-hoc basis.

Classroom assessment practices are generally considered weak, with an overemphasis on their use as an administrative tool than as a pedagogical resource. It is apparently very common to observe uneven application of standards for grading students' work, and grade inflation is considered a serious problem. Reports also indicate that classroom assessment activities tend to be mainly about recalling information and that teachers do not use explicit or a priori criteria for scoring or grading students' work. Additionally, it is reportedly common to observe errors in the scoring or grading of students' work. On the positive side, it appears that parents are well informed about students' grades, classroom assessment information is seen as providing useful feedback to students, and classroom assessment practices are viewed as being aligned with the pedagogical and curricular framework.

Apart from classroom assessment being a required component of a teacher's performance evaluation and of school inspection, there are limited systematic mechanisms in place to monitor the quality of classroom assessment practices.

Suggested policy options:

1. Introduce required courses on classroom assessment for pre- and in-service teachers; develop and make available additional resources for teachers to carry out classroom assessment.
2. Introduce more systematic mechanisms to monitor the quality of classroom assessment practices.
3. Develop additional system-level documents to guide classroom assessment in secondary education.
4. Emphasize the quality of classroom assessment during external evaluation of schools, and introduce regular mechanisms to support teachers in improving assessment practices.

Examinations

Level of development



The *State Matura* has been administered every year since 2008 to grade 12 students. Results are used for certifying secondary education completion, determining admission to university and other higher education institutions, monitoring education quality levels, and planning education policy reforms.

Almost all stakeholders support the *State Matura*. Some stakeholders have made efforts to further improve the examination. For example, policy makers and educators have encouraged the development of a *State Matura* exam for art schools, the introduction of two levels of the mathematics exam, and the introduction of a second examination session for students who failed the examination during the first session.

The National Examination Centre (NEC) is responsible for the *State Matura*. Funding for the *State Matura* is provided to the NEC by the government, and covers all core examination activities (design, administration, data processing, and reporting), as well as staff training, but does not cover research and development.

The NEC has state-of-the-art 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.

The *State Matura* results are officially recognized both in FYR Macedonia and by certification and selection systems abroad, including in Slovenia, the UK, Bulgaria, Albania, and Serbia.

Some documentation about the technical aspects of the *State Matura* exists, but it is not in a formal report format. Comprehensive technical reports were prepared only during the piloting of the *State Matura*.

While there is a general understanding of what the *State Matura* measures, there are no regular, up-to-date courses or workshops on the exam available to classroom teachers. Focused workshops on administration and marking procedures for the exam are organized every year, but these are available only to those serving as test administrators and markers. Workshops related to specific subject areas (content, requirements) are organized from time to time. University professors and certified training providers carry out some of these workshops. Schools have to pay for their teachers' participation in these workshops. A limited number of schools are able to provide access to these workshops for their teachers. Courses for teachers of centrally-developed Matura subjects are offered for free by NEC and the Bureau for Development of Education (BDE).

Suggested policy options:

1. Introduce mechanisms for monitoring the impact of the *State Matura* on the quality of teaching and learning.
2. Improve the quality of documentation on the technical aspects of the *State Matura*.
3. Introduce regular funding for research and development, and for improving the technical quality of the *State Matura*.
4. Introduce workshops or courses for teachers on the *State Matura* that would address, for example, test item development, data analysis, and use of data to improve teaching and learning.

National Large-Scale Assessment (NLSA)

Level of development



The *External Assessment of Students' Achievement in Primary and Secondary Education (External Assessment)* was piloted in 2010 (grade 9 Mathematics) and 2011 (grade 4 Social Sciences, grade 7 Chemistry, grade 10 Mother Tongue, grade 11 Business). Laws for primary and secondary education regulate the *External Assessment*. At the time of this review, the *External Assessment* was still in the pilot phase and the Ministry of Education and Science had announced that it would be formally launched in the 2012-2013 school year.

From 1998 to 2006, the *National Assessment*, the main purposes of which were to monitor education quality at the system level; support schools and teachers; and support policy design, evaluation, and decision making, was considered the most important national large-scale assessment in FYR Macedonia. The *National Assessment* is no longer administered and, according to education policy, the *External Assessment* is now viewed as the most important national large-scale assessment in the country.

The main purpose of the *External Assessment* is to assess the objectivity of teachers' grading. Although not provided for in official documents, there also has been some discussion about formally reporting *External Assessment* results for individual students and using them to make decisions about their selection to the next level in the education cycle.

While policymakers strongly support the *External Assessment*, educators, students, and parents, as well as some donors, oppose it. Stakeholders have challenged the validity of the *External Assessment* given its intended use, and expressed concern that its high stakes nature, as well as the format of the assessment instrument (only multiple-choice items), will result in teaching to the test.

As a result of stakeholders' independent efforts to shape the *External Assessment*, regulation of the *External Assessment* shifted from the Bureau for Development of Education, which is responsible for

curriculum development and supporting teachers, to the National Examination Centre, an institution established for the purpose of administering the *External Assessment*. Additionally, while it was initially intended that students' results on the *External Assessment* would influence their final annual (summative) grades, given that the *Law on Education* provides that only teachers are responsible for grading students, a new bylaw document has been introduced which specifies that *External Assessment* results cannot influence students' final annual (summative) grades.

Currently, no teacher training courses, workshops, or presentations on the *External Assessment* are offered in FYR Macedonia.

While some documentation on the technical aspects of the *External Assessment* exists, it is not in the form of a formal report and it is not publicly available.

There are no mechanisms in place to ensure that the *External Assessment* is used in a way that is consistent with its intended purposes and technical characteristics, or to monitor its consequences.

Suggested policy options:

1. Clearly identify and communicate to key stakeholders the purposes, intended uses, and characteristics of the *External Assessment*.
2. Introduce regular training courses for NEC staff on the development and administration of the *External Assessment*.
3. Introduce in- and pre-service teacher training courses, workshops, or presentations on the *External Assessment*.
4. Introduce mechanisms to (1) ensure that the *External Assessment* is used consistently with its purposes and technical characteristics, and to (2) monitor its consequences or impact.

International Large-Scale Assessment (ILSA)

Level of development

Emerging



FYR Macedonia has participated in a number of ILSA exercises, including PIRLS (2001, 2006), TIMSS (1999, 2003, 2011), and PISA (PISA Plus in 2001). However, there is no policy document that addresses, in a systematic manner, the country's participation in international assessments. Although there is no official document or plan for future participation in an ILSA, FYR Macedonia has taken concrete steps to participate in PISA 2015 as the Ministry of Education and Science sent an official letter of interest to the OECD. The Ministry also has informally expressed its intention to participate in TIMSS 2015.

Most of the funding for FYR Macedonia's participation in ILSA exercises, including the recently concluded TIMSS 2011, has been provided by donors, including The World Bank, USAID, and UNICEF. Funding covered international participation fees, implementation of the assessment exercise in the country, and attendance at international expert meetings. Funding did not cover research and development activities.

The task of processing and analyzing data from the TIMSS 2011 exercise, as well as reporting and disseminating the results in the country, was carried out by NEC staff (as part of their official responsibilities).

Opportunities to learn about ILSA are available only to individuals working directly on a specific ILSA, and are provided by the Ministry of Education and the donor community (e.g., UNICEF, Dutch embassy).

Results from PIRLS 2001, TIMSS 1999, TIMSS 2003, and PISA Plus were printed as national reports, disseminated to stakeholders (schools, universities, educational authorities), and publicly presented at conferences. Reports are available in print and on line.

PIRLS 2006 results were presented only to representatives from the MoES and educational agencies within the MoES, primarily due to a lack of financial resources. There is a plan for the results from TIMSS 2011 to be presented in a country report.

The results of the ILSA exercises have been used by policy makers and education leaders to improve education quality in the country. For example, PIRLS and TIMSS results have been used as a basis for developing the concept for the nine-year primary education cycle, teacher training programs (programs initiated by donors have been designed to reduce gaps in student achievement that were identified by the national reports for the various ILSA exercises), and other assessment activities in the country. TIMSS and PIRLS frameworks and technical standards have been used to inform the development of the *External Assessment* and disclosed items have been used as materials for teacher training on classroom assessment.

Suggested policy options:

1. Prepare a formal policy document that addresses FYR Macedonia's participation in ILSA.
2. Introduce more regular government funding, in combination with donor funding, for carrying out ILSA activities.
3. Introduce mechanisms to capture and evaluate the impact of decisions based on ILSA results on student achievement levels.
4. Use ILSA results to inform decision making on adjusting existing mechanisms or introducing new mechanisms to improve teaching and learning.

Appendix 1: Assessment Types and Their Key Differences

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

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

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

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

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

Appendix 4: Methodology for Assigning Development Levels

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

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

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

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

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

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

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

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

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





Appendix 5: SABER-Student Assessment Rubrics for the Former Yugoslav Republic of Macedonia

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





THE FORMER YUGOSLAV REPUBLIC OF MACEDONIA
Classroom Assessment

ENABLING CONTEXT AND SYSTEM ALIGNMENT

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

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

ASSESSMENT QUALITY*Quality of classroom assessment design, administration, analysis, and use.*

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.	There are adequate required uses of classroom assessment to support student learning, excluding its use as an input for external examination results. ⁹ *	There are adequate required uses of classroom assessment to support student learning, including its use as an input for external examination results.

Classroom Assessment: Development-level rating justifications

1. System-level documents providing guidelines for classroom assessment include:
 - a. Assessment Standards document (authorized by the Bureau for Development of Education (BDE) Ministry of Education, 2008)
 - b. Assessment Criteria document (authorized by the Bureau for Development of Education (BDE) Ministry of Education, 2008). This document consists of a general description of marks (from 1-5) based on the percentage of outcomes expected to be achieved and the level of their complexity based on Bloom's taxonomy.
 - c. Teaching curricula documents by subject and grades (authorized by the Bureau for Development of Education (BDE), Ministry of Education, 2001-2010). These documents include teaching syllabi (for a subject and a grade) that provides short (1/2 -1 page) guidelines related to assessment format, principles, and methods. In the documents developed after 2007 (syllabi for grades 1-9 of primary education) guidelines are more detailed and address formative and summative assessment.
 - d. Guidelines for assessment in primary schools (authorized by Bureau for Development of Education (BDE), Ministry of Education, 2008). These guidelines contain information on school-based assessment principles, standards for classroom assessment, and requirements and expectations of teacher competences.
 - e. Indicators for assessing quality of schools work (authorized by the State educational Inspectorate (SEI), Ministry of Education, 2009).
2. The documents are available in the public libraries, in teacher training colleges, and in in-service courses for teachers. The documents are also publically available on the Ministry of Education websites.
 - a. Assessment Standards - <http://www.bro.gov.mk/?q=standardi>
 - b. Assessment Criteria - http://app.bro.gov.mk/dokumenti/kriteriumi/Kriteriumi_zao_cenuvanje.pdf
 - c. Teaching curricula by subject and grades - <http://www.bro.gov.mk/?q=nastavni-programi>
 - d. Guidelines for assessment in primary schools - <http://toolbox.pep.org.mk/Files/ASSESSMENT%20STANDARDS%20-%20Guide.pdf> (English version)
 - e. Indicators for assessing quality of schools work - http://www.mon.gov.mk/DPI/download/Indikatori_mk.pdf

Schools were informed when the documents were made available by letters and/or during professional meetings and workshops.
3. Resources teachers for classroom assessment include:
 - a. A document that outlines what students are expected to learn in different subject areas at different grade/age levels (for example, the "Teaching curricula by subject and grades" document)
 - b. A document that outlines the level(s) of performance that students are expected to reach in different subject areas at different grade/age levels (for example, the "Assessment Criteria" document)
 - c. Some textbooks or workbooks that provide support for classroom assessment that were provided to teachers during a workshop that was part of a project to improve school-based assessment (2007-2011).
 - d. Scoring criteria or rubrics for students' work (which are available for some subjects or topics and have been developed as a result of projects related to improving classroom assessment).
 - e. Item banks or pools with examples of questions - there is limited number of released items from different types of assessments (national or international) available in printed collections published by NEC. An on line item bank was recently developed (posted in June 2011). Number of items

in this item bank is very limited. At the moment, there are 390 items in 12 primary subjects; 273 items from exams in 8 subjects in State Matura 2010, 86 released items from National assessment 2001 in two subjects (mother tongue and mathematics)).












- f. Online assessment resources - there are some examples of good assessment practices or case studies on classroom assessment available on the Ministry of Education website.
4. "Teaching curricula by subject and grades" documents specify what students are expected to learn, but does not specify to what performance level.
 5. During pre-service teacher training, future teachers have minimal courses on classroom assessment. In-service teacher training on classroom assessment has been administered on an ad-hoc basis. These trainings were organized by certified training providers, and schools and/or teachers could select a particular training from the training catalogue in an area of interest. Assessment has been frequently chosen, with 4,139 out of 26,938 teachers selecting classroom assessment as a topic. Ten training topics have been offered with classroom assessment being one of the four most popular topics selected. From 2003 – 2005 most of the lower primary teachers (grade 1-3) were trained in classroom assessment initiated by the new law regulation on descriptive assessment (students' achievements is described using verbal/ written description instead the numerical marks that were used before). From 2007 -2011, as part of a project for improving SBA, the majority (85 percent) of subject teachers (teachers that teach separate subjects in upper primary and secondary schools) participated in series of workshops on classroom assessment that lasted four to six days. Trainings were organized once from 2007-2011. There were two workshops: four days on formative assessment and two days on teachers' summative assessment.
 6. It is very common to observe uneven application of standards for grading students' work and grade inflation is a serious problem; it is common to observe the classroom assessment activities are mainly about recalling information, teachers do not use explicit or a priori criteria for scoring or grading students' work, it is common to observe errors in the scoring or grading of students' work, and classroom assessment is mainly used as administrative or control tool rather than as pedagogical resource; it is not common for classroom assessment activities to rely mainly on multiple-choice, selection-type questions, parents are well informed of students' grades, classroom assessment information provides useful feedback to students and classroom assessment practices are aligned with the pedagogical and curricular framework.
 7. Mechanisms include that classroom assessment is a required component of a teacher's performance evaluation and of school inspection or teacher supervision.
 8. Teachers are obligated to report on individual student's performance to parents and student at least four times during the school year. Schools are obligated to report aggregated data about students' performance to school district and the Ministry of Education at the end of school year.
 9. The required uses of classroom assessment activities to promote and inform student learning include: diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child's learning, planning next steps in instruction, and grading students for internal classroom uses.

THE FORMER YUGOSLAV REPUBLIC OF MACEDONIA

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. ¹ 	The examination is a stable program that has been operating regularly.	This option does not apply to this dimension
There is no policy document that authorizes the examination.	There is an informal or draft policy document that authorizes the examination.	There is a formal policy document that authorizes the examination. ² 	This option does not apply to this dimension.
This option does not apply to this dimension.	The policy document is not available to the public	The policy document is available to the public. ³ 	This option does not apply to this dimension.
This option does not apply to this dimension.	This option does not apply to this dimension.	The policy document addresses some key aspects of the examination.	The policy document addresses all key aspects of the examination. ⁴ 
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.
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. ¹¹ *	The examination office is a stable organization.	This option does not apply to this dimension.
The examination office is not accountable to an external board or agency.	This option does not apply to this dimension.	The examination office is accountable to an external board or agency. ¹² *	This option does not apply to this dimension.
Examination results are not recognized by any certification or selection system.	Examination results are recognized by certification or selection system in the country.	Examination results are recognized by one certification or selection system in another country.	Examination results are recognized by two or more certification or selection system in another country. ¹³ *
The examination office does not have the required facilities to carry out the examination.	The examination office has some of the required facilities to carry out the examination.	The examination office has all of the required facilities to carry out the examination.	The examination office has state of the art facilities to carry out the examination. ¹⁴ *





(CONTINUED)

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





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

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. ²² *	There is a comprehensive technical report but with restricted circulation.	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. ²⁸ *	This option does not apply to this dimension.	Students' results are confidential.	This option does not apply to this dimension.
ASSESSMENT QUALITY 4: <i>Ensuring positive consequences of the examination</i>			
There are no options for students who do not perform well on the examination, or students must leave the education system.	There are very limited options for students who do not perform well on the examination.	There are some options for students who do not perform well on the examination. ²⁹ *	There is a variety of options for students who do not perform well on the examination.
There are no mechanisms in place to monitor the consequences of the examination. ³⁰ *	This option does not apply to this dimension.	There are some mechanisms in place to monitor the consequences of the examination.	There is a variety of mechanisms in place to monitor the consequences of the examination.

Examinations: Development-level rating justifications

1. The State Matura has been administered every year since 2008 (four times). It is administered at grade 12 and is used for student certification for grade completion, selection to university or higher education institution, for monitoring education quality levels, and for planning education policy reforms. Starting in 2007, students have had to take four exams: mother tongue and three subjects selected from a list of 12 subjects. Two of the exams are administered at the central level, and two are administered by a student's school. Currently, five subjects in total are administered at the central level, and seven are administered by schools. Under the next phase of the Matura Exam, which was expected to begin in 2010 but has been postponed to 2013, all of the exams on the list of 12 optional subjects will be administered at the central level.
2. Documents include:
 - a. Rulebook for taking Matura exams and assessment of students' results on the exams on Matura for Gymnasium and Secondary Vocational Education (authorized by the Ministry of Education and Science, 2007 and 2010 (revised)).
 - b. Concept paper for Matura and Final Exam in four years Secondary Education (authorized by State Matura Committee and Ministry of Education and Science, 2005 and 2010 (revised)).
 - c. Exams' syllabi. For each teaching subject, these documents provide the purpose of the exam; exam content; types of the items; and test specification.
 - d. Documents related to exams organization and administration procedures for organization and administration of the external exams in the State Matura (internal and confidential document of NEC)
3. All documents have been published and distributed to all secondary schools. They are also available on-line. Documents include:
 - a. Rulebook for taking Matura exams and assessment of students' results on the exams on Matura for Gymnasium and Secondary Vocational Education:
<http://www.slvesnik.com.mk/Issues/58AC759457D37C428A95285484B3C4CA.pdf> p.23
 - b. Concept paper for Matura and Final Exam in four years Secondary Education:
http://www.matura.gov.mk/data_files/state_graduate/mk/5115_Koncepcija.pdf (in Macedonian language)
 - c. Exams' syllabi:
<http://www.matura.gov.mk/documents.aspx?language=MK&page=O6dtQQpiV3o=>
4. The policy documents cover various aspects of the examination:

The Law for National Examination Centre, the Rulebook for taking Matura exams and assessment of students' results on the exams on Matura for Gymnasium and Secondary Vocational Education, the Guidance for the Mature exam, the Guidance for the local coordinators, and the Guidance for the test administrator outline governance, distribution of power, responsibilities among key entities; the Concept paper for Matura and Final Exam in four years Secondary Education and the Amendments to the Law for Secondary Education (Article 27) describe the purpose of the examination; the Concept paper for Matura and Final Exam in four years Secondary Education and the Law on High Education describe the authorized uses of results; the Rulebook for taking Matura exams and assessment of students' results on the exams on Matura for Gymnasium and Secondary Vocational Education specifies who can sit for

the examination, outlines procedures for special/disadvantaged students, and outlines procedures to investigate and address security breaches, cheating, or other forms of inappropriate behavior; procedures for organization and administration of the external exams in the State Matura (internal and confidential document of NEC) identifies rules about preparation; and Exams' syllabi explain the alignment with curricula and standards and the format of the examination questions.

5. Policymakers and employers strongly support the Matura exams. Introducing Matura exams was result of the joint effort of the policymakers and professionals. Educators, students, parents, and universities support the Matura exams as well. BDE advisors have taken part in the development of the documents for Matura (Concept for Matura, Exam syllabi), development of the test items, and monitoring the Matura administration. Most teachers support the Matura exam because it motivates students to learn more. As communicated by the media, teachers and school principals support the Matura. While students initially strongly opposed the examination (there were student demonstrations against the Matura and negotiation with student associations), they are now supportive or neutral of it. The parents were initially confused about the Matura or opposed it. They now support it because the Matura is lower stakes than the university entrance exams use to be. Universities were consulted and involved in the process of developing the Concept for Matura, and they accepted to replace the entrance exams with the results of the Matura. Presidents of all subjects' Matura committees are professors from the universities. There have been no issues expressed by the universities on the validity of the Matura exam. The media is generally neutral to the Matura but is usually critical of the Matura results. Teacher unions and think tanks and NGOs have not expressed their opinion on the Matura.
6. During the development of the Concept for Matura examination, universities submitted proposals to improve the Concept for the examination. After approval of the Concept, educators and policymakers took the initiative to:
 - a. develop a State Matura exam for art schools;
 - b. introduce two levels of the exam in mathematics for the State Matura;
 - c. introduce a second examination session for those students who failed the State Matura in the first session (the first examination session is in June, the second one is in August);
 - d. harmonize the content of the Mother tongue exam in VET schools with the exam in General Secondary schools.
7. There were no attempts made to improve the Matura exam that were not well received by the leadership in charge of the exam.
8. Funding for the State Matura is provided by the government to the National Examination Centre.
9. Funding also covers staff training.
10. Funding does not cover research and development.
11. National Examination Centre (NEC) was established as a semi-independent agency in 2009. The agency use to be unit within the Bureau for Development of Education (BDE).

12. NEC is accountable to State Matura Board a external body established by the Ministry of Education. It consists of the representatives from: Ministry of Education and Science; Universities; National Examination Centre; Bureau for Development of Education and VET Centre.
13. The examination results are officially recognized in the country and by more than one certification and selection system abroad, including in Slovenia, the UK, Bulgaria, Albania, and Serbia.
14. The examination office has computers for all technical staff, a secure building, secure storage facilities, access to adequate computer servers, the ability to backup data, and adequate communication (telephone, email, internet) tools.
15. There is permanent or full-time staff, but it is insufficient to meet needs of the examination. Specifically, NEC lack subject specialist for some mother tongue (Turkish) and foreign (Russian, French) languages, but this issue has been addressed by the use of the BDE subject specialists. Subject specialists are members of the subject's Matura committee, provide technical support to test developers, test markers, and are responsible for timely and secure preparation of exam papers.

There have been some indications of poor exam monitoring by the some of the test administrators due to some test administrators not following the test administration rules and not being sanctioned for allowing students to engage in improper behavior.

16. Non-university training courses or workshops on educational measurement and evaluation were provided as part of projects related to establishing the examination system and mainly by donors. While these opportunities are still in place, they are rarely used due to limited financing. University professors and certified training providers offer courses on educational measurement and evaluation, and schools have to pay for their teachers participation or the teachers themselves have to pay. A limited number of school can provide these trainings for their teachers.
17. The Matura exam is tied to the national curricula and standards.
18. The results from the Matura exam are used by universities for admission as well as by employers.
19. The materials needed to prepare for the examination is widely accessible by all students (over 90 percent) in a variety of learning contexts. Materials needed to prepare for the Matura exams are regular textbooks used during secondary education. Textbooks, exam items and marking schemes from previous exams, and user friendly guidance for taking Matura are available for free to all students in all languages of instruction for compulsory and elective subjects. Tests from previous exams are available online, information on how to prepare for the examination is available in the form of a booklet that is distributed to students, and the framework documents explaining what is measured on the examination (also known as the "Exams' syllabi") are also made available.
20. There are voluntary courses or workshops that are not regularly updated. Workshops for the State Matura administration and test marking are organized every year and are regularly updated. These courses are available just for the test administrators (approximately 2000 people) and test markers (approximately 320 people).

Workshops related to separate subjects (content, requirements) are organized from time to time, and university professors and certified training providers carry out some of these workshops. Schools have to pay for their teachers' participation or teachers themselves must pay. A limited number of schools can provide access to the workshops for their teachers. Courses for teachers of centrally developed subjects in Matura are offered by NEC and BDE for free.

21. Teachers create examination questions for a pool for each compulsory and elective exam that is centrally developed. Test items developers are usually invited based on their previous experience in curriculum and exam syllabi development, experience as test markers and reputation as good teachers. Teachers have role of administrators in a school other than their own, by a schedule developed by NEC. Teachers also score the exam.
22. There is some documentation about the technical aspects of the examination, but it is not in a formal report format. Comprehensive technical reports were prepared only during the piloting phase.
23. Items and administration procedures were piloted during the preparatory phase of the Matura exam. Technical personnel from the NEC review item pools and select the best items. The President of the Matura committee (university professor) for each exam makes the final selection of the items that will be included in the exam. Selection is based on the quality of test items and the relevance of the items for the test specification (which is part of exam syllabi). There are attempts to maintain the difficulty levels of the Matura exams from one year to the next.
24. Inappropriate behaviors that diminish the credibility of the examination typically occur during the examination process include: copying from other candidates; collusion among candidates via mobile phones, passing of paper, or equivalent; and the provision of external assistance via the supervisor, mobile phone etc.










According to the rulebook on test administration, if someone is caught cheating, he or she will be warned and if the same person continues with cheating, that person will be disqualified. If copying is proven on open-ended and essay-type of items, those items are marked with 0 points. The new rule of conduct (2010) identifies penalties for test administrators and supervisors who engage in or support students' improper behavior.

25. All universities (state and private) use exam results as a selection criterion. Employers also accept results as valid and reliable.
26. All students may take the examination, regardless of background (e.g., gender, ethnic group), location (e.g., urban, rural), ability to pay (e.g., transportation, fees) or the like.
27. There is no systematic evidence on improper use of the results.
28. Students results on Matura are displayed at the school and everyone can see the results. Making the results public is considered part of the strategy for transparency of the examination.
29. Students who do not perform well on the examination may retake the examination once again in the same school year (in August) and next school years in two terms (June and August). Students may also opt for less selective schools/universities/tracks and can repeat the grade.
30. There are no mechanisms in place to monitor the consequences of the examination.

THE FORMER YUGOSLAV REPUBLIC OF MACEDONIA
National (or System-Level) Large-Scale Assessment (NLSA)

ENABLING CONTEXT







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

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





(CONTINUED)

LATENT ●○○○	EMERGING ●●○○	ESTABLISHED ●●●○	ADVANCED ●●●●
ENABLING CONTEXT 3: <i>Having regular funding for NLSA</i>			
There is no funding allocated to the NLSA.	There is irregular funding allocated to the NLSA.	There is regular funding allocated to the NLSA. ⁶ *	This option does not apply to this dimension.
This option does not apply to this dimension.	Funding covers some core NLSA activities: design, administration, analysis and reporting.	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.	The NLSA office is a permanent agency, institution or unit. ⁹ *	This option does not apply to this dimension.
This option does not apply to this dimension.	Political considerations regularly hamper technical considerations. ¹⁰ *	Political considerations sometimes hamper technical considerations.	Political considerations never hamper technical considerations.
This option does not apply to this dimension.	The NLSA office is not accountable to a clearly recognized body. ¹¹ *	The NLSA office is accountable to a clearly recognized body.	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. ¹⁴ *	This option does not apply to this dimension.	The NLSA measures performance against curriculum or learning standards.	This option does not apply to this dimension.
What the NLSA measures is generally questioned by stakeholder groups.	This option does not apply to this dimension.	What the NLSA measures is questioned by some stakeholder groups. ¹⁵ *	What the NLSA measures is largely accepted by stakeholder groups.
There are no mechanisms in place to ensure that the NLSA accurately measures what it is supposed to measure. ¹⁶ *	There are ad hoc reviews of the NLSA to ensure that it measures what it is intended to measure.	There are regular internal reviews of the NLSA to ensure that it measures what it is intended to measure.	This option does not apply to this dimension.
SYSTEM ALIGNMENT 2: <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.	This option does not apply to this dimension.	At least one option is offered to include all groups of students in the NLSA. ¹⁸ *	Different options are offered to include all groups of students in the NLSA.
There are no mechanisms in place to ensure the quality of the NLSA.	This option does not apply to this dimension.	There are some mechanisms in place to ensure the quality of the NLSA. ¹⁹ *	There are a variety of mechanisms in place to ensure the quality of the NLSA.
There is no technical report or other documentation about the NLSA.	There is some documentation about the technical aspects of the NLSA, but it is not in a formal report format. ²⁰ *	There is a comprehensive technical report but with restricted circulation.	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. ²¹ *	NLSA results are disseminated in an effective way.	This option does not apply to this dimension.
NLSA information is not used or is used in ways inconsistent with the purposes or the technical characteristics of the assessment. ²² *	This option does not apply to this dimension.	NLSA results are used by some stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.	NLSA information is used by all stakeholder groups in a way that is consistent with the purposes and technical characteristics of the assessment.
There are no mechanisms in place to monitor the consequences of the NLSA. 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 (or System-Level) Large Scale Assessment (NLSA): Development-level rating justifications

1. The External Assessment of Students' Achievement in Primary and Secondary Education was piloted in 2010 (in 27 out of 95 secondary schools, and assessed Mathematics at grade 9) and in 2011 (in 14 of 343 primary schools, and assessed Social sciences at grade 4 and Chemistry at grade 7; in 11 of 95 secondary schools, and assessed Mother tongue (Macedonian, Albanian, Turkish) in grade 10, and Business in grade 11.
2. Formal policy documents include:
 - a. Law on Primary Education and its amendments art. 54, 71 (authorized by the Ministry of Education and Science, 2008 (Official Gazette, No 103, 19.08.2008 p.2) and amendments 2010 (Official Gazette, No 33, 09.03.2010 p.10))
 - b. Law on Secondary education and its amendments art. 45a, 45b, 45v, 56 and 70 (authorized by the Ministry of Education and Science, 2008 (Official Gazette, No 92, 22.07.2008 p. 11) and amendments 2010 (Official Gazette, No 33, 09.03.2010, p 2))
 - c. Rule on Organization and Administration of the External Assessment of Students in Primary Schools, Establishing and Work of the School Committees, Confidentiality of the Testing Materials, Verification of the Testing Materials by the School Committee and the Template and Content of the Report (2010)
 - d. Rule on Organization and Administration of the External Assessment of Students in Secondary Schools, Establishing and Work of the School Committees, Confidentiality of the Testing Materials, Verification of the Testing Materials by the School Committee and the Template and Content of the Report (2010)
3. The laws and rules are available online and in hard-copy.
4. While there is no large-scale assessment plan for the coming years or future assessment rounds and the External Assessment of Students' Achievement in Primary and Secondary Education is still in the pilot phase, the MoES has announced that this assessment will be launched next school year (2012/13).
5. Independent efforts have been made by different stakeholder groups to reform the External Assessment of Students' Achievement in Primary and Secondary Education. Individuals have made motions to the Constitutional Court, challenging particular articles in the laws (provisions related to division of responsibilities between the education institutions and the idea that grades that student will receive on the External Assessment to influence the summative grades of the student – final / annual grades). Thus, the MoES has made revision of the laws, specifically:
 - a. A new institution – the National Examination Center, which is responsible for administration of the assessment was established
 - b. In the new bylaw document, there is no longer the option of using External Assessment results to influence students' grades

While policymakers strongly support the External Assessment of Students' Achievement in Primary and Secondary Education, educators, students, and parents oppose it, while donors strongly oppose it. For example, the World Bank team has given an opinion that Ministry of Education has to reconsider the decision of introducing massive external assessment, addressing that the data from the students' learning assessment should be used primarily to direct responses toward improvement of teaching and learning. Similar suggestions had been given by the USAID financed Primary Education Project, addressing

that the stake of the assessment for teachers and format of assessment instruments (only multiple choice items) force teaching for the test and hampers the Project's efforts to improve assessment for learning.

6. There is regular (continuous and predictable) funding allocated by the government. Test item development and administration of the pilot have been funded so far. The Ministry will allocate sufficient funds in the budget for the next fiscal year.
7. Because the assessment is still in the pilot phase, the funds that were allocated so far were for the piloting of the assessment. Funding for full implementation will cover assessment administration, data analysis, data reporting, staff training, and the development of the software and establishment of a system for on line testing.
8. Funding will not cover research and development activities.
9. According to the laws on the Primary and Secondary Education and law for the National Examination Center (NEC), the NEC is responsible for administration of the assessment, Bureau for Development in Education (BDE) and the VET Center are responsible for item development.
10. Policymakers are concerned about grades inflation and low reliability of students' grades, and intend to use the data on reliability of teacher's grading as an indicator for teacher evaluation in order to force teachers to improve the reliability of their assessment practices

Professionals have shared with the Ministry their suggestions on the purpose of the External Assessment, test format, sampling, the issue of validity, particularly given the intended use of the assessment results (for evaluating students and teachers).

Policy makers have not taken most of the suggestions given by the professionals into consideration.

11. The legal framework regarding the External Assessment does not regulate accountability of the NEC, and it is not clear to whom the NEC is accountable.
12. There is permanent or full-time staff, but it is insufficient to meet the needs of the External Assessment when it will be implemented after the pilot phase. The Ministry plans to finance part time staff during the administration and data processing periods every year.

Issues that have been identified with the performance of the human resources responsible for the External Assessment include omission of curricular topics and weaknesses in test design (only multiple choice items have been introduced on the External Assessment and therefore some domains, such as writing, practical experimentation, and oral exams were not included in the test design; the External Assessment also does not measure higher order cognitive skills).

13. No opportunities have been offered on an annual basis. NEC staff had received high quality training for item development/ test design and administration of the exams by the CITO and Anglia Assessment experts. This technical assistance was provided during the period of establishing NEC (2002 - 2006).

There have been some training organized for the item writers before the first pilot of the External Assessment. Courses, according to the participants, were of satisfactory quality but the participants expressed need for more hands-on training.





Trainings have also been made available through projects for establishing the system of external assessment and ad hoc trainings for test item developers.

14. The External Assessment measures the objectivity of teacher's grading. The main purpose of the assessment is to evaluate the objectivity of teachers' grading. It is intended that each student's results on the External Assessment will be reported in a Diploma Supplement (DC). The DS will be used as additional criteria for selection of the student in a next cycle in educational system, i.e. from primary to secondary education. This is not yet regulated in any official document.
15. Education professionals have challenged the validity of the External Assessment given its intended use. The majority of teachers expressed doubts that this assessment will contribute to the enhancement of the quality of teaching and learning.
16. There are no mechanisms in place to ensure that the External Assessment accurately measures what it is supposed to measure.
17. There are no teacher training courses, workshops, or presentations on the External Assessment (e.g., domains measured, how to read and use results) offered in the country. The External Assessment is currently in the pilot phase.
18. It is planned for the External Assessment to be offered in all four languages of instruction in FYR Macedonia.
19. Mechanisms to ensure the quality of the External Assessment include a standardized manual for assessment administrators, and a pilot is conducted before the main data collection takes place.
20. There is some documentation about the technical aspects of the assessment, but it is not in a formal report format and is not available to the public.
21. Results are featured in newspapers, magazines, radio, or television. The External Assessment is currently in the pilot phase.
22. The External Assessment is currently in the pilot phase.
23. Mechanisms to monitor the consequences of the External Assessment have not yet been developed.

THE FORMER YUGOSLAV REPUBLIC OF MACEDONIA
International Large-Scale Assessment (ILSA)

ENABLING CONTEXT

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

LATENT 	EMERGING 	ESTABLISHED 	ADVANCED 
ENABLING CONTEXT 1: <i>Setting clear policies for ILSA</i>			
The country/system has not participated in an ILSA in the last 10 years.	This option does not apply to this dimension.	The country/system has participated in at least one ILSA in the last 10 years.	The country/system has participated in two or more ILSA in the last 10 years. ¹ *
The country/system has not taken concrete steps to participate in an ILSA in the next 5 years. 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.	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. ⁴ *	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. ⁵ *	Funding covers all core activities of the ILSA.	Funding covers all core activities of the ILSA.
Funding does not cover research and development activities. ⁶ *	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. ⁷ *	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. ⁸ *	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. ⁹ *

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. ¹⁰ *	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. ¹¹ *	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. ¹² *	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. ¹³ *
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. ¹⁴ *
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. ¹⁵ *	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. ¹⁶ *	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. ¹⁷ *	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. ¹⁸ *	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. ¹⁹ *	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. FYR Macedonia has participated in PIRLS (in 2001, 2006), TIMSS (in 1999, 2003, and 2011), and PISA (in 2000 and PISA Plus in 2001). This rubric is completed with regard to FYR Macedonia's participation in TIMSS 2011.
2. There is no official document or plan for future participation in an ILSA, although the Ministry of Education and Science (MoES) has expressed intention to participate in PISA 2015 and TIMSS 2015.
3. The decision for FYR Macedonia to participate in an ILSA has been made on an ad-hoc basis.
4. Funding for the participation fee was provided for by international agencies USAID, UNICEF, and MoES using the World Bank loan proceeds (under the "Education Modernization Project"). Most of the work related to the administration and marking was done by teachers on a voluntary basis. Test desk-top preparation, data entry and data cleaning were done by the National Examination Centre (these activities are part of its official responsibilities). Translation and printing was paid for by the MoES. Participation on the international expert meetings was covered by the MoES from the Ministry discretion funding and by UNICEF.
5. Funding covered international participation fees, implementation of the assessment exercise in the country, and attendance at international expert meetings for the assessment exercise. Processing and analyzing data collected from implementation of the assessment exercise and reporting and disseminating the assessment results in the country was conducted by NEC (as part of its official responsibilities).
6. Funding does not cover research and development.
7. The team is comprised of the national coordinator, who is an employee of NEC and was assigned by the Ministry of Education and Science to be the national coordinator, and other staff from NEC. All of the team members have gained experience in working on ILSAs over the last ten years. The national coordinator is the only person officially nominated to the ILSA team. Other members of the team work on ILSA in addition to their regular jobs in NEC. Subject specialists for the sciences are teachers who are engaged when necessary, at times on a voluntary basis.
8. The national coordinator is fluent in the language of the assessment.
9. Only the national coordinator participated in meetings from the time that FYR Macedonia joined TIMSS 2011. The national coordinator is fluent in the language in which the international-level meetings are conducted and related documentation is available. Additionally, the team has previous experience working on international assessments and the necessary training and experience to carry out the required assessment activities effectively. No issues have been identified with the carrying out of the international assessment in the country.
10. Only the national coordinator attended international meetings and workshops. During previous ILSA administrations, national coordinators and members of the ILSA team participated in almost all workshops on international assessments and on using international assessment databases offered by IEA.

11. Opportunity to learn about international assessments is provided by the Ministry of Education and the donor community (UNICEF, Dutch embassy) and covers attending international workshops or training on international assessments.
12. Opportunities to learn about the ILSA are available to individuals working directly on the specific international assessment exercise.
13. FYR Macedonia met all technical standards required to have its data presented in the main displays of the international report.
14. FYR Macedonia's contribution of new knowledge on international large-scale assessment includes:
 - a. PIRLS 2001 Encyclopedia <http://timss.bc.edu/pirls2001i/pdf/encyclopedia.pdf>
 - b. PIRLS 2006 Encyclopedia <http://timssandpirls.bc.edu/PDF/P06Encyclopedia.pdf>
 - c. TIMSS 2011 Encyclopedia (in preparation)
 - d. NC Anica Aleksova was member of the Item scaling and anchoring committee for TIMSS 1999, member of the International expert panel for Mathematics and Mathematics item development task force and Mathematics item review committee for TIMSS 2003.
 - e. NRC Bojana Naceva was member of the PIRLS 2006 Questionnaires Development Group <http://timssandpirls.bc.edu/pirls2006/framework.html>
 - f. Bojana Naceva and Gorica Mickovska: Impact of PIRLS 2001 in Republic of Macedonia, in The impact of PIRLS 2001 in 13 Countries, Studies in International Comparative and Multicultural Education , Germany , 2007
http://books.google.mk/books?id=HF9Pvd7sOvAC&pg=PA193&lpg=PA193&dq=The+impact+of+PIRLS+2001+in+13+Countries&source=bl&ots=gERf4T1JDL&sig=xqUOPJ91xYShL7XB6EG5s9Cd4B0&hl=mk&redir_esc=y#
15. There is a plan for the results from TIMSS 2011 to be presented in a country report.

Results from PIRLS 2001, TIMSS 1999, TIMSS 2003 and PISA plus were printed as national reports, disseminated to stakeholders (schools, Universities, Educational authorities) and publicly presented on the conferences. Reports are available in printed versions and on line. Results from PIRLS 2006 were not presented in a national report because the national coordinator for PIRLS left NEC and the management staff of NEC was changed, therefore writing PIRLS National report was no longer a priority.

Released TIMSS's and PIRLS's items were published and distributed to schools. Released TIMSS's items are available on-line: <http://itembank.dic.edu.mk/Documents.aspx>

TIMSS 2003 and PIRLS 2006 results have been presented only to the representatives from the MoES and educational agencies within the MoES. The main reason for not shearing the results and other relevant information widely was primarily the lack of finances and interest of the BDE for using ILSA results for improving teaching practice.

Results from ILSA have been distributed through a national report online, copies of the national and international reports were distributed to key stakeholders, country results were communicated through a press release.

16. TIMSS 1999 and PIRLS 2001 results were fed-back to the schools' representatives from all primary schools in FYR Macedonia during the regional workshops, the goals of which was to share ILSA results, to familiarize schools with these assessments (it was first time that FYR Macedonia participated in TIMSS and PIRLS) and with the types of test items used in ILSA, and to communicate this information to expected outcomes in teaching practice in FYR Macedonia. TIMSS 2003 and PIRLS 2006 results have been presented only to the representatives from the MoES and educational agencies within the MoES. The main reason for not shearing the results and other relevant information widely was primarily the lack of finances and interest of the BDE for using ILSA results for improving teaching practice.
17. Media coverage is limited to a few small articles. The media has always been invited to all press conferences related to the dissemination of ILSA results. The media has covered only basic information and the focus was usually on the results.
18. The results of the international assessment exercise have been used by policy makers or education leaders to improve education quality in the country by informing curriculum improvement (PIRLS and TIMSS results have been used as a basis for developing the concept for 9 years Primary Education (2007)), teacher training programs (teacher training programs initiated by donors have been designed to reduce gaps identified in the national reports on students' achievement in ILSA.), and other assessment activities in the country (TIMSS and PIRLS framework and technical standards have been used for the development of the concept for National Assessment and disclosed items have been used as materials for teacher training on classroom assessment).
19. There has not been regular national assessment or any impact evaluation/research related to student achievement since 2006

Acknowledgements

This report was prepared by The World Bank SABER-Student Assessment team in collaboration with Bojana Naceva, World Bank Senior Education Specialist and Task Team Leader for education projects in FYR Macedonia, and Gorica Mickovska, Project Manager at the Macedonian Civic Education Centre.

References

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

European Commission. 2012. The Former Yugoslav Republic of Macedonia 2012 Progress Report. Brussels: European Commission.

International Monetary Fund. 2012. Public Information Notice (PIN) No. 12/58. IMF Executive Board Concludes 2011 Article IV Consultation with Former Yugoslav Republic of Macedonia. Washington, DC: International Monetary Fund.

Republic of Macedonia State Statistical Office. 2012. "Primary, Lower and Upper Secondary Schools at the Beginning of the School Year 2011/2012." Skopje: State statistical office of the Republic of Macedonia.

United Nations Educational, Scientific, and Cultural Organization - International Bureau of Education. 2011. World Data on Education. VII Ed. 2010/11. "The former Yugoslav Republic of Macedonia." Geneva: UNESCO-IBE.

World Bank. 2010. Country Partnership Strategy for Former Yugoslav Republic of Macedonia for the Period FY11-FY14. Report No. 54928-MK. Washington, DC: World Bank.

———. 2012. "Macedonia, FYR at a glance." Washington, DC: World Bank. Data retrieved from http://devdata.worldbank.org/AAG/mkd_aag.pdf on January 29, 2013.

———. FYR Macedonia Country Indicator Data. Washington, DC: World Bank. Data retrieved from <http://data.worldbank.org/> on January 29, 2013.

———. The World Bank's Strategy in FYR of Macedonia: 2011-2014. Washington, DC: World Bank. Data retrieved from <http://go.worldbank.org/QS6GA93YY0> on January 29, 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