

78840

# Developing the Enabling Context for Student Assessment in Russia

Viktor Bolotov, Galina Kovaleva,  
Marina Pinskaya, and Igor Valdman





# Developing the Enabling Context for Student Assessment in Russia

*Viktor Bolotov, Galina Kovaleva,  
Marina Pinskaya, and Igor Valdman*

© 2013 The International Bank for Reconstruction and Development /  
The World Bank  
1818 H Street NW  
Washington DC 20433  
Telephone: 202-473-1000  
Internet: [www.worldbank.org](http://www.worldbank.org)

1 2 3 4 15 14 13 12

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.

### **Rights and Permissions**

The material in this work is subject to copyright. Because The World Bank encourages dissemination of its knowledge, this work may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given.

Any queries on rights and licenses, including subsidiary rights, should be addressed to the Office of the Publisher, The World Bank, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2422; e-mail: [pubrights@worldbank.org](mailto:pubrights@worldbank.org).

Cover design: Patricia Hord.Graphik Design, Alexandria, VA

# Contents

About the Series .....	v
About the Authors .....	vii
Acknowledgments .....	ix
Executive Summary .....	xi
Introduction .....	1
Enabling Context .....	3
Student Assessment Programs .....	6
Drivers for Change .....	15
Lessons Learned .....	18
Bibliography .....	22



## About the Series

Building strong education systems that promote learning is fundamental to development and economic growth. Over the past few years, as developing countries have succeeded in building more classrooms, and getting millions more children into school, the education community has begun to actively embrace the vision of measurable learning for all children in school. However, learning depends not only on resources invested in the school system, but also on the quality of the policies and institutions that enable their use and on how well the policies are implemented.

In 2011, the World Bank Group launched Education Sector Strategy 2020: Learning for All, which outlines an agenda for achieving “Learning for All” in the developing world over the next decade. To support implementation of the strategy, the World Bank commenced a multi-year program to support countries in systematically examining and strengthening the performance of their education systems. This evidence-based initiative, called SABER (Systems Approach for Better Education Results), is building a toolkit of diagnostics for examining education systems and their component policy domains against global standards, best practices, and in comparison with the policies and practices of countries around the world. By leveraging this global knowledge, SABER fills a gap in the availability of data and evidence on what matters most to improve the quality of education and achievement of better results.

SABER-Student Assessment, one of the systems examined within the SABER program, has developed tools to analyze and benchmark student assessment policies and systems around the world, with the goal of promoting stronger assessment systems that contribute to improved education quality and learning for all. To help explore the state of knowledge in the area, the SABER-Student Assessment team invited leading academics, assessment experts, and practitioners from developing and industrialized countries to come together to discuss assessment issues relevant for improving education quality and learning outcomes. The papers and case studies on student assessment in this series are the result of those conversations and the underlying research. Prior to publication, all of the papers benefited from a rigorous review

process, which included comments from World Bank staff, academics, development practitioners, and country assessment experts.

All SABER-Student Assessment papers in this series were made possible by support from the Russia Education Aid for Development Trust Fund (READ TF). READ TF is a collaboration between the Russian Federation and the World Bank that supports the improvement of student learning outcomes in low-income countries through the development of robust student assessment systems.

The SABER working paper series was produced under the general guidance of Elizabeth King, Education Director, and Harry Anthony Patrinos, Education Manager in the Human Development Network of the World Bank. The Student Assessment papers in the series were produced under the technical leadership of Marguerite Clarke, Senior Education Specialist and SABER-Student Assessment Team Coordinator in the Human Development Network of the World Bank. Papers in this series represent the independent views of the authors.



## About the Authors

**Viktor Bolotov** has been vice president of the Russian Academy of Education since 2008. A leader in the modernization of Russian education, he has contributed to the development of Russia's education reform strategy and coordinated the reform of the country's unified state examination. He has had several leading positions in the Russian education system, including head of the Department of Teacher Training Education, first deputy minister, and head of the Federal Service of Supervision in Education and Science. He has a Ph.D. in mathematics and physics and is also a doctor of pedagogy. Dr. Bolotov has over 100 publications in the area of mathematics, teacher training, and education reform in Russia.

**Galina Kovaleva** has headed the Center for Evaluating the Quality of Education, Institute of Content and Methods of Education, Russian Academy of Education since 1996. After starting her career as a secondary school physics teacher, she became a member of the national team for developing national standards and assessments in Russia. Dr. Kovaleva was involved in organizing and conducting international comparative student assessments in the country (i.e., TIMSS, PIRLS, PISA), developing the methodology of the Unified State Examination in Russia, developing a national system for evaluating the quality of education, and implementing assessment projects at the national and regional level. Dr. Kovaleva has a Ph.D. in pedagogy and is the author of over 90 publications in the areas of curriculum, national standards, instructional materials, and educational assessment.

**Marina Pinskaya** is lead research associate, Institute for Educational Studies, National Research University Higher School of Economics. Before starting to work in the field of educational studies, she gained extensive pedagogical experience working in the Russian school system for many years. She earned her Ph.D. in 2007, defending a thesis entitled "The Portfolio as a Tool for Assessment of Individual Educational Achievements." She has taken part in major projects related to the modernization of the education system in Russia carried out by the national Ministry of Education, as well as in international comparative

projects and research related to school effectiveness and school improvement that were conducted by the National Research University Higher School of Economics. Dr. Pinskaya is the author of more than 40 scientific publications. Her professional interests include educational policy, assessment of the quality of education, and school effectiveness.

**Igor Valdman** has been director of the Russian Training Center and head of the Laboratory for Educational Monitoring, Institute of Management in Education, Russian Academy of Education, since 2010. He has been a member of the Russian G8 expert group on education (St. Petersburg, 2006) and the Quality Task Team of the Education For All Fast Track Initiative partnership (2007–08). In addition, he was been involved in several national and international education projects led by the World Bank, the European Union, the Finland National Board of Education, and the Soros Foundation. He has expertise in the area of development aid for education and previously worked as director of the Center for International Cooperation in Education Development (2008–10). Dr. Valdman has a master's degree in education from Manchester University and a Ph.D. in pedagogy. He has published over 50 works in the areas of educational assessment, educational policy, school accountability, and community involvement in education.

## **Acknowledgments**

The authors are grateful to the READ Trust Fund team for the opportunity to present Russia's experience in assessing student learning and achievement as part of the SABER series. Special thanks are extended to the World Bank experts who made helpful remarks and suggestions on improving the case study and who oversaw the editing. Our thanks in particular go to Marguerite Clarke, Maria Jose Ramirez, Robin Horn, Julia Liberman, Emily Gardner, Manorama Gotur, and Diana Manevskaya.



## Executive Summary

This case study has two goals: (1) to present Russia’s experience in strengthening its student assessment system, and (2) to share lessons learned for the benefit of other countries that may be interested in strengthening their own student assessment systems. The paper examines reforms to the enabling context that supports educational assessment in Russia—that is, reforms that affected the policy framework and institutions, the development of human capacity, and funding sources. It then analyzes reforms to large-scale assessments, examinations, and classroom assessment activities; identifies the driving forces that contributed to the reforms; and extracts key lessons about strengthening an educational assessment system.

Russia has put in place laws, policies, and regulations that mandate assessment activities, which are implemented by a network of federal and regional institutions. These organizations have permanent trained staff to work on assessment activities. Their capacity in this field has been developed thanks to professional opportunities to work on large-scale assessment programs and hands-on training provided by international and national agencies. Consistent funding for assessment activities also has been provided by federal, regional, and municipal governments.

The main reform introduced by Russia was the introduction of a country-level, standardized, Unified State Examination. This examination came to replace institutional-level examinations—a major reform that enabled the country to improve the quality and fairness of examinations, while reducing corruption and unprofessional practices. Since the 1990s, Russia has been an avid participant in all major international large-scale assessment programs. These assessments have provided a comprehensive view of the Russian education system from a global perspective.

Several factors drove the strengthening of the student assessment system in Russia, the most important of which were: (1) concerns about educational quality; (2) the need to monitor learning and educational quality in a context of decentralization and accountability; (3) stable political leadership; (4) financial support to the regions; (5) the need for more valid, fair, efficient, and corruption-free university entrance

examinations; (6) the need to align assessments with newly introduced federal educational standards; (7) the transfer of international technology to national and regional assessment programs; and (8) concerns that classroom assessments were not being used to support pedagogy.

The following lessons can be drawn from this case study. Implementing reforms to a student assessment system in a large, federally-governed country is not an easy task. Nevertheless, success seems more likely when reforms have strong and stable political support, there is a clear focus or priority area for reform, and reforms are implemented using a step-by-step approach. Reformers should be prepared to face resistance to change, which will be strong when reforms threaten the interests of some stakeholders, as in the case of corruption and unprofessional practices. Leading public discussions with stakeholders is important in order to overcome resistance and gain support for reforms.

# Developing the Enabling Context for Student Assessment in Russia

*Viktor Bolotov, Galina Kovaleva,  
Marina Pinskaya, and Igor Valdman*

## Introduction

In the first decade of the twenty-first century, Russia launched a comprehensive reform of its student assessment system. This task was part of a broader reform of the education system, which aimed to ensure that students developed the skills required for the country's social and economic development. The reform was also part of a greater decentralization of educational management and accountability in Russia. What drivers allowed these reforms to take place? What were the main obstacles to the reforms? What were the main features of the assessment system after the reforms were introduced? What kind of policies, institutions, and human and fiscal resources support the current student assessment system?

This case study focuses on the main types of assessments used in education: large-scale assessments for monitoring educational quality (international, national, and regional); examinations for secondary school certification and tertiary education selection; and classroom assessments in support of teaching and learning. The paper also examines the enabling context for assessment activities in Russia, which encompasses the policy framework, institutional arrangements, and the human and fiscal resources needed to carry out assessment programs and activities.<sup>1</sup>

Russia is an interesting case study for several reasons. The country has already completed the journey from institutional-level examinations to a national standardized examination program for secondary school certification and tertiary education selection. Many developing countries

---

<sup>1</sup> For a general discussion of the enabling context for student assessment systems, see Clarke (2012).

are planning to introduce similar reforms. Russia is also interesting because, on the one hand, the country has consistently participated in international large-scale assessments but, on the other, has not yet put in place a permanent national large-scale assessment program.

To learn from the Russian experience, it is important to bear in mind the broader context in which the reforms took place. Russia is the largest country in the world, occupying close to one-third of Eurasia. It has a large multiethnic population of 142 million people, three quarters of which reside in urban areas. The country has a high Human Development Index (0.788) and an average life span of 69 years. Russia is an upper middle-income country with a per capita GDP of US\$21,000 in PPP and a federal government comprised of 83 constituent “regions.” These regions are relatively autonomous and responsible for the provision of educational services in their jurisdictions, but they are accountable to the federal government for the quality of education.

Russia has a strong education system, but there are concerns about its quality. The country inherited a strong tradition of universal high-quality education from the Soviet era. Some 13 million students presently attend 53,000 schools and are served by 1.3 million teachers. The education system is divided into preschool, primary (grades 1–4), basic (grades 5–9), secondary (grades 10–11), and tertiary education. Primary and basic education are free and compulsory according to the Russian constitution. Public and private expenditures on education are 5.1 percent of GDP. In terms of educational quality, the results of international assessments show that Russian students perform relatively well compared to developed countries. However, their performance is weaker when they are asked to apply their knowledge to solve real-life problems. These results raised concerns and drove a subsequent push to reform the education system.

Important lessons can be drawn from this case study. The Russian experience clearly indicates that a national reform to the student assessment system has to be supported by a strong and stable political leadership. Focusing on one strategic reform (in the case of Russia, the Unified State Examination) can open the door to reforms in other areas. Setting regulations, creating institutions, building human capacity, and ensuring continuous financial support are key ingredients of success. However, the case study also shows that introducing reforms is not an



easy task. Changing traditional assessment practices means changing the teaching culture. This process affects different interest groups that benefit from the status quo. This is especially the case when reforms aim to eradicate corruption and different types of unprofessional practices. Strong political support and open public discussion are required to overcome barriers to change.

The case study is organized as follows. Following this introduction, the paper provides an overview of the enabling context for student assessment in Russia. It analyzes existing policies, institutions, human resources, and funding for assessment activities. The following section describes the different types of assessments that are conducted in Russia: international, national, and regional large-scale assessments; examinations; and classroom assessments. The fourth section reviews the driving factors that enabled (or not) reforms in the assessment system. The final section distills the main lessons learned in the course of the reforms.

## **Enabling Context**

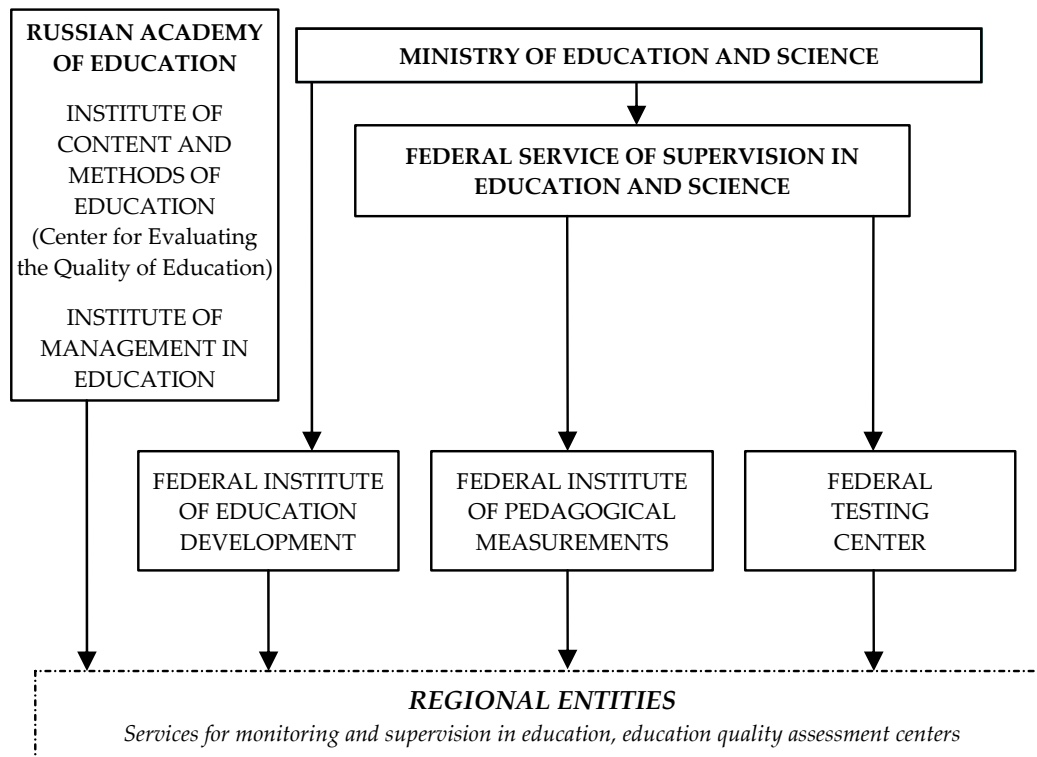
Russia has a well-established assessment system that provides regular information for different educational purposes. The system encompasses three main types of assessment: (1) primary and secondary school examinations for certification and, in the case of the latter, selection into tertiary education; (2) national and international large-scale assessments for monitoring educational quality; and (3) classroom assessments for the purposes of grading students and improving teaching and learning. Other uses of assessment information include raising awareness about the quality of education; informing decision making at the federal, regional, local, school, and classroom level; recording student performance; advising parents about student progress; informing teacher training programs; school certification; and teacher evaluation.

This comprehensive assessment system is supported by clear policies. The Law on Education (passed in 1992 and amended in 2010), together with regulations of the Ministry of Education and Science and the Federal Service on Education and Science Supervision, provide the legal basis for the national examination in grade 9 (Final State

Certification, or FSC-9) and the unified examination in grade 11 (Unified State Examination, or USE). The Law on Education makes individual schools responsible for classroom assessment activities. Federal regulations mandate Russia’s participation in international assessments.

The national assessment system operates through a complex network of federal and regional institutions. Each institution is in charge of a different aspect of the system, including policy decisions, design, implementation, and research. The institutions that form this network are shown in figure 1 and listed below:

**Figure 1. Institutions Involved in the Student Assessment System in Russia**



- The Ministry of Education and Science, which defines national priorities in education and is responsible for the development and implementation of national educational policy.
- The Federal Service of Supervision in Education and Science, which is responsible for the control and supervision of education laws that ensure young people’s access to quality education and

that accredits and licenses educational institutions. It is also responsible for organizing the implementation of the USE in the Russian Federation and providing the Final State Certification of basic and secondary school graduates.

- The Federal Institute of Education Development, which produces policy research, organizes expert meetings, and generates knowledge used for policy advice.
- The Federal Institute of Pedagogical Measurements, which develops standardized assessments, trains examination administrators, and analyzes examination results.
- The Federal Testing Center, which provides technology and information support in administering assessments, including data processing and maintenance of the examinations database. It is also responsible for producing and distributing examination materials to the regions.
- The Russian Academy of Education, which pursues research in assessment and evaluation through the activities of two institutes: The Institute of Content and Methods of Education and the Institute of Management in Education. The Center for Evaluating the Quality of Education, part of the Institute of Content and Methods of Education, is responsible for administering international assessments (e.g., TIMSS, PIRLS, PISA). It also provides methodological support to align examinations with national educational standards. The Institute of Management in Education uses assessment and evaluation results to manage the overall education system.

In addition to federal institutions, regional entities are in charge of core assessment activities. Every Russian region sets up a State Examination Board during the secondary school examination period. This board is responsible for preparing and implementing examinations and for approving the results. Regional centers provide technological support to administration sites. These regional centers are connected to primary information processing offices, which in turn are connected to the actual administration sites. The regional centers consolidate the information they receive from the latter. This information is then used to

produce a regional database, which is then submitted to the Federal Testing Center to produce the national examination database.

Russia has developed sufficient local capacity to implement all core assessment activities, including test development, assessment administration, data processing, and reporting. This capacity has been developing since 2000 at the federal, regional, and municipal levels. The main institutions that have been in charge of providing capacity building in this sphere are the Federal Institute of Pedagogical Measurements, the Federal Testing Center, and the Center for Evaluating the Quality of Education of the Russian Academy of Education.

Assessment activities in Russia are supported by permanent and qualified staff. For example, the Federal Institute of Pedagogical Measurements has 84 permanent staff members. In 2011, the team in charge of developing the USE examination consisted of 150 people, a group that included assessment specialists, test developers, test item-writers, statisticians, and other specialists. The committees for the 14 school subjects included 125 persons (70 percent of whom had a Ph.D.). Some 51 percent of this group were professors or assistant professors from different universities; 29 percent, subject matter specialists; and 20 percent, teachers.

Permanent funding supports the student assessment system in Russia. Examinations are funded from federal, regional, and municipal budgets and are free of charge to students. National and regional large-scale assessments are financed from national and regional budgets. Russia's participation in the international PIRLS, TIMSS, and PISA assessments is financed from the national budget under the Federal Targeted Program of Development of Education. According to this program for the period 2006–2010, the federal budget for developing a national system to evaluate the quality of education amounted to over 2 billion rubles (about US\$75 million).

## **Student Assessment Programs**

This section reviews Russia's participation in international large-scale assessments, the status of national and regional large-scale assessments

initiatives, reforms to existing examination programs, and efforts to revamp classroom assessment activities.

### **International Large-Scale Assessments**

Russia has been an avid participant in almost all large-scale international assessments. Since the 1990s, the country participated in the Progress in International Reading and Literacy Study (PIRLS) for grade 4 in 2001, 2006, and 2011; in the Trends in International Mathematics and Science Study (TIMSS) for grades 4 and 8 in 1995, 1999, 2003, 2007, and 2011, and for grade 11 in 2008; and in the Programme for International Student Assessment (PISA) for students aged 15 (regardless of grade) in 2000, 2003, 2006, 2009, and 2012. All of these assessments use nationally-representative samples of students to draw inferences at the national level.

Participation in different but complementary international large-scale assessments has provided a comprehensive picture of the education system in Russia. Together, these assessments provided a global perspective on student learning and educational quality in Russia for the first time. They also provided valuable information about the social and school contexts in which students learn. For instance, they gathered information about student beliefs and motivations; the socioeconomic background of their families; their school facilities and teachers; and school pedagogy. This information has been very relevant, considering that Russia does not have a permanent national large-scale assessment program (see next subsection).

Results from international large-scale assessments have raised concerns about educational quality in Russia. They show that although the performance of Russian students is relatively strong in the early grades, it worsens as students progress through school. Russian fourth-graders ranked at the top of the 40 countries participating in the PIRLS assessment in 2006 and 2011. The performance of Russian eighth-graders has also been consistently above the international average score on TIMSS. The performance of 15-year-old Russian students on PISA, however, has been consistently below that of their peers from developed countries. Educational tradition in Russia emphasizes the acquisition of conceptual knowledge and theories, not the application of that knowledge in order to solve real-life problems. This has translated into

students who know theory, but who are weak in solving practical problems.

Federal and regional institutions are involved in implementing international assessments. At the federal level, the administration of such assessments is the primary responsibility of the Center for Evaluating the Quality of Education of the Russian Academy of Education. The Russian Federal Ministry of Education and Science and the Federal Service of Supervision in Education and Science also support international assessments. At the regional level, these assessments are implemented with the support of ministries and/or departments of education, regional assessment and examination centers, teacher training institutes, and some other entities.

### **National and Regional Large-Scale Assessments**

As noted earlier, Russia does not have a permanent, national large-scale assessment program to monitor educational quality. Instead, several independent assessment projects have been implemented in the country, most importantly:

- National assessments that use the same sample as international assessments, but administer national tests (such as mathematics in grade 9 in 1995; mathematics in primary school in 1999; and mathematics and Russian in primary school in 2011).
- National assessments to monitor educational quality that were performed as an experiment for modernizing the structure and content of general education (2001–04).
- National assessments that monitored student achievement in primary school during the development of new educational standards (2007–13).

These assessments have provided valuable information on educational quality and student learning. Nevertheless, they were not institutionalized and ended up being discontinued. This was in part the consequence of the federal government not taking the lead in national assessments. Instead, leadership has usually come from nongovernmental research organizations, and from the Russian Academy of Education, which conducted a series of once-off assessment exercises.

Leadership for conducting large-scale assessments has also come from the regions. In fact, regional governments seem to be taking faster steps to build student assessment programs than the federal government. Several regions are gradually building their own programs; some already have permanent programs in place. In a decentralized federal structure, the information provided by regional assessments has been important for strengthening the management capacity of local government. This information is used for multiple purposes, including monitoring educational quality, informing management, and as an input into teacher evaluation and school accreditation.

The Russian regions have also implemented large-scale assessments for specific research projects. The “Study of Mathematics and Science Performance of School Students in the Republic of Tatarstan (1999–2000),” for example, aimed to know how educational quality in Tatarstan compared to that in Russia and other countries. The study answered this question by administering the TIMSS 1999 instruments to a representative sample of students in the region. A second assessment was run in Smolensk region in 2002, with the aim of evaluating the effectiveness of newly-introduced pedagogical resources in the schools.

However, at the federal level there is concern about the development of such regional programs. The main reasons for this concern are: (1) regional “home-made” assessments do not seem to meet international quality standards; (2) the local capacity to develop these assessments may be absent; and (3) these assessments are more costly than national assessments, i.e., there is no economy of scale. To reduce these risks, more leadership from the federal government seems to be required.

## **State Examinations**

### ***Unified State Examination (USE)***

Russia recently introduced a new examination program that overcomes many of the weaknesses of previous examinations. Before the reform, Russia did not have country-level standardized examinations to certify students at the end of secondary education and select them into tertiary education. Instead, it had secondary school examinations and university entrance examinations that were led by each institution separately. This process had several problems, including the central government’s lack of

control over what was measured, the requirement that students take several examinations in a short time span, and the incomparability of student results from different institutions. Moreover, in the case of university entrance examinations, there were serious equity problems, given that many students from rural areas could not afford to take the examinations, corrupt practices to gain access to universities were pervasive, and overall, the system was very inefficient.

In order to overcome these weaknesses, in 2001 the Ministry of Education and Sciences was given the mandate to develop a new, combined unified secondary school leaving and university entrance examination. This reform was at the core of efforts to modernize education in Russia and has become the foundation of a national student assessment system. The Unified State Examination (USE) was gradually introduced after extensive piloting. It was first piloted from 2001 through 2008 in Russian regions and universities that volunteered to implement it. Since 2009, it has been implemented nationwide in a centralized, standardized manner. In 2010, 878,000 students took over 3 million examinations in 83 subjects (each student has to take at least two examinations, since Russian language and mathematics are mandatory).

The USE is aligned with new federal education standards for secondary education. It uses a wider range of assessment formats, including oral questions, multiple-choice questions, short-answer questions (which require students to write a word combination or a number), and extended-response questions (which require students to write an essay or explain their reasoning or state a personal point of view).

A combination of automated and expert judgment is used to score student responses. Multiple-choice items and items requiring a short answer are automatically scored by computer. Extended-response items are scored by the State Examination Board of the Russian regions. Scorers are selected so as to avoid potential conflicts of interests (e.g., they should not come from the same school as the student or from the university to which the student is applying). Test scores are scaled and equated on a 0–100 scale. In order to obtain a secondary school certificate, students have to pass both the Russian language and mathematics examinations. Those students who want to continue to



tertiary education or a secondary professional school must also pass the USE in the disciplines required by each institution.

Beyond selection and certification, USE results are used for a variety of different purposes. Regional governments, regional assessment centers, teacher training institutions, and municipal centers use them to monitor school and teacher performance, as well as to inform teacher training programs. At the federal level, examination results are used (together with other performance indicators) to evaluate the effectiveness of regional education authorities and distribute performance-based monetary incentives to the regions. Unfortunately, incentive schemes do not take into consideration the social context in which a school operates. There have thus been cases of tough and unfair sanctions imposed on schools and municipalities based on USE results, usually those that serve the poorest students.

### *Final State Certification (FSC-9)*

The reform of examinations in Russia also encompassed the introduction of a new, central-level standardized examination at the end of basic education (grade 9): the Final State Certification (FSC-9). This examination will replace school- and regional-level examinations that are not standardized and suffer from problems of quality, fairness, and efficiency. The main purpose of the FSC-9 will be to certify that students have met the requirements of basic education, which themselves will be aligned with new federal learning standards. Examination results will be comparable across schools and regions, thus ensuring a more fair and valid certification process. Beyond certification, examination results are expected to be used to “track” students in secondary education (i.e., the vocational or academic track), as well as in school accreditation and teacher evaluation processes.

The FSC-9 will be managed at the federal level, using a similar approach to that used for the secondary school leaving examination, the USE. Students in grade 9 will be required to pass at least four examinations in order to obtain their certification. These include Russian language and mathematics, plus two examinations chosen by the student from among all subjects taught in grade 9. Completed examination papers will be scored by the Regional Examination Board. Examination results will be reported on a 20–45 scale, which will be converted to a 5-

score scale used by schools to report student grades. Examination results will be included in the student certification diploma.

The FSC-9 has been piloted in a growing number of Russian regions that have volunteered to conduct the tests. So far, students from 78 regions of the Russian Federation have taken part in the FSC-9. In 2008, the number of students that took the FSC-9 mathematics examination reached 530,000 (around 59 percent of grade 9 students); in 2009, it reached 820,000 (more than 75 percent of grade 9 students) and the total number of examinations administered reached 1.9 million. The new examinations are expected to be implemented nationwide by 2014.

### **Classroom Assessments**

Russian teachers use classroom assessments for different purposes. As shown by PIRLS 2006, teachers mainly use classroom assessments to assign student marks or grades, identify students in need of remedial instruction, and as a way of informing parents about student progress. To a lesser extent, teachers also use classroom assessments to adapt their instruction, group students for specific instruction, and provide data for national or local educational monitoring. However, teachers do not use classroom assessments to monitor students' non-academic skills, such as team work, creativity, and character development. Moreover, these assessments do not seem to be used by local or regional officers to monitor the quality of education.

Regarding official classroom assessment information, teachers enter students' grades in the class book, following a standard procedure. The class book is the official instrument for reporting grades at the regional level. Based on current marks recorded in the class book, teachers award students final grades for a certain period (i.e., a 3-, 6-, or 12-month period). Grades are recorded in student report cards, which are the main tool for informing parents of student achievement. School e-books are a new form of recording student performance. This mechanism has been introduced only in some schools, but the prospects are high for their universal implementation due to a gradual shift to electronic documents.

Teachers currently grade students using a scale system that has been highly criticized. The system was established by the Commissariat of Public Education in 1937, and applies to both formative and summative assessments. Scoring is based on a five-point scale that goes

from “1” (poor) to “5” (excellent). The grading system is justly criticized by teachers, students, and the pedagogical community for: (1) lack of transparency, together with the subjectivity of assessment criteria and procedures, (2) the fact that in practice, only positive marks are given, reducing the effective range of the scale to 3–5 points; and (3) the scant information provided by the grades.

To improve classroom assessment practices and overcome the above-mentioned weaknesses, the following efforts to reform classroom assessments have been made:

- *Aligning classroom assessments with the new federal learning standards.* These standards were developed by the Russian Academy of Education and were adopted by primary schools starting in 2011. The introduction of new learning standards is part of a push to assess a broader range of skills using a more applied, problem-solving approach.
- *Using explicit criteria to judge student learning.* Efforts have been made to use criterion-referenced assessments, where assessment criteria linked to learning standards are explicit, known, and agreed upon in advance by teachers and students. This change will ensure transparency, while avoiding judgments based on the relative standing of individual students among their classmates.
- *Employing self-assessment and peer assessment.* It is important for students to internalize assessment criteria and to monitor their own work. Self- and peer assessments are only possible when explicit and agreed-upon criteria are used to judge student work.
- *Ongoing assessment of student learning.* A major emphasis has been put on using assessments for formative purposes, i.e., for the purposes of informing teaching and learning.
- *Assessing personal skills.* To overcome the criticisms that traditional assessment practices were excessively focused on academic subjects, a stronger emphasis has been put on assessing students’ social skills, character, creativity, and team work, among other qualities.
- *Using a wider range of assessment instruments and formats.* This change was needed to measure a wider range of knowledge and

skills. Assessment instruments and formats may include written papers, oral questions, projects, practical work, observations, self-assessments, and peer assessments.

- *Using student portfolios.* A portfolio provides evidence of student performance because it represents a comprehensive sample of student work. Portfolios may include student biographical data; a profile of student achievement over time; information on extracurricular activities; certificates received at academic competitions; papers and creative work (e.g., poems, essays, drawings). Portfolios of secondary school students were introduced in dozens of Russian regions during an experiment conducted by the Russian Ministry of Education over the period 2002–05.
- *Introducing a non-grade system.* In 2003, the Ministry of Education and Science issued guidelines for schools to adopt a non-grade system for primary education.
- *Using a new scale to describe learning levels.* A ten-point scale describing five levels of knowledge acquisition, from 1 (*very poor*) to 10 (*excellent*), was introduced. Student knowledge acquisition ranges from recognition and memorization of factual information to the application of theoretical knowledge to practical, non-routine situations.
- *Communicating assessment information to students and parents.* Reforms in classroom assessment have emphasized the communication of assessment results to students and parents so that they “own” the assessments. This communication should go beyond grades to provide descriptions of what the students know and are able to do. The Ministry of Education and Science has proposed using student assessment sheets to reflect levels of key skill attainment in each subject area. The new sheets will allow for ongoing recording of student grades.

Reforms to classroom assessments in Russia still have a long way to go before being fully adopted by schools. Resistance from school teachers, school administrators, and the media has made it hard to introduce changes. New regulations favoring formative, non-grade assessments have been carried out in a small number of schools that

adopted a revised curriculum. Primary schools seem to make fair use of formative assessments, but these practices do not seem to transfer to secondary schools. Some schools have adopted reforms as part of specific projects to improve educational quality. Others have adopted them as part of the pedagogical approach used in their respective school networks. In schools that belong to the International Baccalaureate program, the use of criterion-referenced assessments and self-assessments seems to be common practice; however, these schools do not represent typical Russian schools.

## **Drivers for Change**

The entire national assessment system in Russia was reformed thanks to strong political support and leadership. The federal government supported the reforms in order to transition to a more decentralized, performance-based budgeting and management system. This required valid and trustworthy information with which to monitor educational quality at the national level. Continuing political support was also critical for the USE to gain official status as a state examination. The leadership of the Ministry of Education and Science was also critical to push for reforms in classroom assessment.

The introduction of new federal educational standards in 2010–11 pushed the development of the student assessment system forward, as examinations, national and regional assessments, as well as classroom assessments had to be aligned with the new standards.

Pervasive quality problems originally led to the reform of secondary school examinations. These included misalignment of previous examinations with new educational standards, the lack of comparability of student results, inefficiency, inequality, and corruption. The unequal representation of students from lower-income families in rural areas and small towns was another factor that drove the introduction of the USE. Before this reform, the higher education system in Moscow worked mostly to support local elites. For instance, only 25 percent of students in Moscow universities originated from other regions of the country. After the USE was implemented, high school graduates had the chance to take the examinations in their hometowns rather than in the cities where

universities are located. Hence, students from other Russian regions started to take the examinations and be proportionally represented at the universities. As reported by the Ministry of Education and Science, the number of university students originating from rural areas increased by 15–20 percent over the period 2001–08.

The USE therefore became a genuine mechanism of modernization in educational assessment in Russia. Secondary school graduates now have a real chance to choose a preferred higher education institution. Every student, regardless of a social background or place of residence, has the chance to send his/her USE certificate and scores to any university (or several simultaneously) and take part in an admissions competition. The leading universities from Moscow, St. Petersburg, and other major cities began to work for the benefit of the entire country.

Another main driver of reforms was the pervasive corruption surrounding university entrance examinations. Unfortunately, higher education officials did not trust the testing methods proposed by the USE. Instead, they wanted to assess the knowledge and capabilities of candidates in an in-person dialogue during the entrance exam. In reality, there were hidden motives for this preference. One was the wish to control additional cash flow, that is, the “entrance ticket” (fees paid to private tutors, or preparatory courses from a university to guarantee enrollment) and direct bribes paid to university staff. Another reason was the unwillingness of rectors to lose their “administrative currency,” that is, their ability to guarantee the enrollment of applicants from the “rector’s list” (generally children of highly placed, “useful” parents). The USE accordingly struck a blow at corruption in university entrance examinations.

The introduction of the USE required broad public discussion. This process was critical to attract additional support for the examinations and provide well-reasoned answers to opponents’ objections. The USE was discussed openly and systematically in printed and electronic media; on TV, the radio, and the Internet; and during various meetings, conferences, and forums. The discussion was not limited to the professional educational community; it included students, parents, employers, and members of different social groups.

However, the public debate did not result in a dramatic change of attitude towards the USE. Even now, public opinion does not

unanimously support its introduction. Many politicians and mass media outlets did their best to persuade the public that the USE was killing world-class Russian education, claiming that the examinations focused only on basic skills and failed to assess student creativity.

As the USE started to become a high-stakes assessment, several professional groups immediately started to oppose it openly. This occurred when USE results began to be used as an input into teacher evaluation and school certification processes.

Poor student assessment results were another driver that allowed for the reform of the assessment system. Findings from national and international assessments, and from regional examinations, all showed that Russian students had difficulties applying basic skills to practical, real-life situations. This despite the fact that they had the theoretical knowledge to do just that. Accordingly, there was a push to introduce real-life problems into teaching and classroom assessments. There also was greater emphasis given to the use of projects and observations as more authentic ways to evaluate student learning.

Russia's active participation in international assessments was also instrumental in strengthening its national assessment system. The involvement of staff from various organizations and almost all Russian regions helped build an assessment culture in the country. The use of international quality standards for student assessment "raised the bar" and put pressure on Russian educators to improve local assessment programs.

Participation in international large-scale assessments thus allowed Russia to build local capacity in assessment. International consortiums and the Russian Academy of Education provided training to the staff of federal agencies, national examinations offices, educational standards units, and regional coordinators from at least 45 regions. Professional development opportunities offer hands-on experience in conducting large-scale assessments, including designing instruments, item writing, data processing, and analysis. Participants also benefit from participating in international and local conferences and workshops.

In addition, participation in international assessments allowed for the transfer of up-to-date technology to national and regional assessment programs in Russia. Federal and regional assessments, as well as the FSC and the USE, benefited from changes in test design, sampling,

administration procedures, and scaling methods. All of these changes were possible thanks to the increased experience and skills of local staff.

Financial support from the federal government also enabled the regions to develop their own regional assessment programs. To date, more than 30 regions have benefited from this support.

Reforms in classroom assessments have been harder to introduce, although some changes are taking place. The lack of teachers trained in the use of new assessment strategies has made it difficult for reforms to be implemented. Pre- and in-service teacher training programs do not provide training in this topic. Traditional assessment practices used for bureaucratic and control purposes are deeply engrained in the culture of Russian teachers; changing that culture takes time, training, and resources. Nevertheless, the weaknesses of current classroom assessment practices have created room for change. For instance, the lack of comparability of student grades across schools has created support for the adoption of criterion-referenced assessments.

## Lessons Learned

Other countries can draw several lessons from the Russian experience of reforming its student assessment system, as discussed below.

*Sustained political support is needed.* Sustained political support is a key factor in reforming national student assessment systems. Countries need this support first, to make the political decision to launch an assessment reform, and then to move forward with reforms. Political support is especially important because reforms will always be subject to some criticism by stakeholders and because reforms must be institutionalized by translating changes into new rules, practices, and procedures.

*Introduce focused reforms.* While successful reforms usually require a systemic approach, countries need to identify a clear priority or core program that leads the reforms. Russia strategically focused its educational reform on the USE. This clear focus ensured both political support and resources for successful implementation.

*The federal government needs to play a leading role.* The federal government should play the key role in leading assessment reform in



large, federally-governed countries. This is important to ensure that both federal and regional assessment programs meet high-quality technical standards. In Russia, the federal government has not always taken this lead. As a consequence, the regions seem to be leading reforms to large-scale assessments for monitoring purposes and accountability. Although regional assessments have their merits, they also have several limitations, including lack of local capacity; the poor-quality, “home-made” nature of the assessments; and prohibitive costs.

*Implement reforms in a step-by-step manner.* The successful implementation of complex projects requires a long-term approach. Time is needed to introduce gradual adjustments, build local capacity, and gain support for reform. In Russia, the USE was piloted for nine years before its full implementation at the national level. Time was needed to adjust the structure and content of the examinations, test the new technology, and solve implementation problems. This phased implementation, based on pilot testing, was a key factor in its success.

*Build an assessment culture.* Countries aiming to reform their assessment systems need to build an assessment culture with stakeholders that are “assessment literate.” This means having a critical mass of teachers, educators, and political leaders that can understand assessment information and have positive attitudes towards it. Russia has taken important steps to build this culture. The reform of examinations favored the development of an assessment culture and provided impetus for training. Nevertheless, many teachers in Russia still require assessment training. Two critical pieces seem to be missing: (1) training in assessment topics in pre- and in- service teacher training institutions, and (2) the development of teaching standards that include student assessment.

*Be prepared to overcome resistance to change.* In contrast to a number of countries with well-established standardized testing systems (e.g., the United States, United Kingdom, and Australia), standardized tests were practically nonexistent in Russia prior to the introduction of the USE. For the most part, standardized measurements became a point of discussion after it was implemented. University professors held on to their beliefs that university acceptance is best granted based on a face-to-face interaction with a candidate, instead of an entrance examination. School teachers held on to the belief that they were best positioned to judge

student performance. Teachers continue to distrust new assessment practices and have difficulties interpreting and using objective assessment data. These beliefs and negative attitudes made it a challenge to introduce student assessment reforms. Changing teacher practices will require changing teaching culture *and* overcoming resistance to change.

*Gain public support for introducing reforms.* Countries aiming to introduce reforms to their national examinations should consider launching a big campaign to gain public support for the reforms. In Russia, a skillfully organized, broad, and open public discussion was critical for gaining public support for the USE. This campaign included the media (TV, newspapers) and open forums that featured examination specialists, teachers, and researchers. The horizontal interaction of ideas was a necessary component of the introduction of the USE.

*Ensure fair use of assessment information.* Student results are affected by a myriad of factors, some internal to the school (e.g., teachers), some external to the school (most notably, poverty). When student results are used to rank or compare schools, especially in the context of high-stakes accountability programs, it is extremely important to take into account poverty levels. Failure to do so is unfair to quality schools that serve the country's poorest students. In Russia, there has been concern about the unfair use of USE results to distribute school incentives without considering the poverty level in which schools and municipalities operate. One way to prevent this oversight is by publishing guidelines that signal appropriate and inappropriate ways of comparing schools and/or enforcing comparisons (and rewards) of schools that operate in similar socioeconomic contexts.

*Avoid corruption by establishing a system of public control.* Bribery, misbehavior, falsification, and impersonation are all widely reported forms of corruption in national examinations. Russia has taken big steps to prevent corruption by establishing an open and transparent system of public control at all stages of examinations, engaging independent observers during examinations, and discussing cases of corruption with reputable nongovernmental organizations, among other steps. Countries aiming to combat similar corruption can learn from these reforms to improve their assessment systems.

\* \* \*

In conclusion, Russia has taken important steps to develop a national student assessment system. Nevertheless, important steps remain ahead in order to consolidate these reforms. The authors of this case study hope that the lessons of Russia's experience will be useful to other countries in building their own national assessments systems.

## Bibliography

- Arora, A., P. Foy, M. O. Martin, and I. V. S. Mullis, eds. 2009. *TIMSS Advanced 2008 Technical Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Bolotov, V. 2012. "Quality Assessment in Education." *Journal of Management in Education* [Moscow]. Vol. 6 (Issue 2).
- Bolotov, V., G. Kovaleva, and I. Valdman. 2012. "Russian Quality Assessment System: What Have We Learned for 10 years?" Materials of the VIII International Conference, "Tendencies of Education Development: Problems of Management," February 18–19, 2011. Moscow.
- Bolotov, V., and I. Valdman. 2012. "Conditions of Effective Utilization of Assessment Results." *Journal of Pedagogy* 6.
- Clarke, M. 2012. "What Matters Most for Student Assessment Systems: A Framework Paper." SABER–Student Assessment Working Paper 1. World Bank, Washington, DC.
- Federal Commission on Evaluation of Experimental Results on Introduction of the Unified State Examination. 2009. "Report on Evaluation of Experimental Results on Introduction of the Unified State Examination." Federal Commission on Evaluation of Experimental Results on Introduction of the Unified State Examination, Moscow.
- Kovaleva, G., et al. 2012. *Main Results of the International Assessment PISA 2009; Analytic Report*. Moscow: MAX-PRESS, Moscow.
- Kovaleva, G., and K. Krasnianskaia. 2012. "The Russian Federation." In *TIMSS 2011 Encyclopedia: Education Policy and Curriculum in Mathematics and Science*, ed. I. V. S. Mullis, M. O. Martin, C. A. Minnich, G. M. Stanco, A. Arora, V. A. S. Centurino, and C. E. Castle. Vol. 2, L–Z and Benchmarking Participants. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College

- Minina, E. 2010. "The Unified National Test for Student Admission to Higher Education in Russia: A Pillar of Modernisation?" In *Politics, Modernisation, and Education Reform in Russia: From Past to Present*, ed. D. Johnson. Oxford Studies in Comparative Education. Oxford, UK: Oxford University.
- Mullis, I. V. S., M. O. Martin, P. Foy, and A. Arora. 2012. *TIMSS 2011: International Results in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., M. O. Martin, P. Foy, and K. T. Drucker. 2012. *PIRLS 2011: International Results in Reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., M. O. Martin, A. M. Kennedy, and P. Foy. 2007. *PIRLS 2006: International Report; IEA's Progress in International Reading and Literacy Study in Primary Schools in 40 Countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., M. O. Martin, C. A. Minnich, K. T. Drucker, and M. A. Ragan, eds. 2012. *PIRLS 2011 Encyclopedia: Education Policy and Curriculum in Reading*. Vols. 1 and 2. TIMSS & PIRLS International Study Center, Boston College.
- Russian Federation, Ministry of Education and Science. 2006. *Education in Russia—2006* (in Russian). Moscow: Ministry of Education and Science. <http://stat.edu.ru/Edu2006/contents.htm>.
- — —. 2010. *Federal Educational Standards for Primary Education* (in Russian). Moscow: Ministry of Education and Science. [http://xn--80abucjiibhv9a.xn--p1ai/документы/922/файл/748/ФГОС\\_НОО.pdf](http://xn--80abucjiibhv9a.xn--p1ai/документы/922/файл/748/ФГОС_НОО.pdf).
- — —. 2008. *Unified State Examination: Final State Certification of the Graduate of Grade 9* (in Russian). Moscow: Ministry of Education and Science. [http://ege.edu.ru/ru/main/legal-documents/index.php?id\\_4=17622](http://ege.edu.ru/ru/main/legal-documents/index.php?id_4=17622).
- UNDP (United Nations Development Programme). 2009. *Human Development Report 2009: Overcoming Barriers—Human Mobility and Development*. New York: UNDP.

### **Useful websites**

[www.centeroko.ru](http://www.centeroko.ru) – Center for Evaluating the Quality of Education (in Russian)

[www.edu.ru](http://www.edu.ru) – Federal State Educational Standards (in Russian)

[www.fipi.ru](http://www.fipi.ru) – Federal Institute of Pedagogical Measurements (in Russian)

[www.gks.ru](http://www.gks.ru) – Federal State Statistics Service of Russian Federation (in Russian)

[www.iea.nl](http://www.iea.nl) – International Association for the Evaluation of Educational Achievement

[www.obrnadzor.gov.ru](http://www.obrnadzor.gov.ru) – Federal Service of Supervision in Education and Science (in Russian)

[www.oecd.org](http://www.oecd.org) – Organisation for Economic Co-operation and Development

[www.rustest.ru](http://www.rustest.ru) – Federal Testing Center (in Russian)

[timss.bc.edu](http://timss.bc.edu) – TIMSS and PIRLS International Study Center

***List of papers published in the current Working Paper series***

1. Clarke, M. 2012. "What Matters Most for Student Assessment Systems: A Framework Paper." (*Available in Russian and Portuguese*).
2. Ramirez, M. 2012. "Developing the Enabling Context for Student Assessment in Chile." (*Available in Russian*).
3. Ramirez, M. 2012. "Disseminating and Using Student Assessment Information in Chile." (*Available in Russian*).
4. Liberman, J., and M. Clarke. 2012. "Review of World Bank Support for Student Assessment Activities in Client Countries." (*Available in Russian*).
5. Flockton, L. 2012. "The Development of the Student Assessment System in New Zealand." (*Available in Russian*).
6. Allen, R. 2012. "Developing the Enabling Context for School-Based Assessment in Queensland, Australia." (*Available in Russian*).
7. Castro, M. 2012. "Developing the Enabling Context for Student Assessment in Brazil." (*Available in Russian*).
8. Kanjee, A., and S. Acana. 2013. "Developing the Enabling Context for Student Assessment in Uganda." (*Available in Russian*).
9. Bolotov, V., G. Kovaleva, M. Pinskaya, and I. Valdman. 2013. "Developing the Enabling Context for Student Assessment in the Russian Federation."

***Forthcoming papers***

10. Tyumeneva, Y. 2013. "Disseminating and Using Student Assessment Information in Russia."

This case study shows how Russia—a large, federally-governed country—introduced important reforms to strengthen its student assessment system. These reforms enhanced the enabling context that supports student assessment by putting in place the right policy framework, creating a network of institutions that took responsibility for assessment activities, building local capacity, and securing stable funding. As a result, Russia’s assessment programs and activities improved. The most important reform was the introduction of a unified examination for secondary school certification and selection into tertiary education. This country-level standardized test replaced institutional-level examinations that suffered from quality and corruption problems. This case study shows that reforms are more likely to succeed when there is sustained political support, a clear focus or priority area for change, changes are introduced gradually, and reforms are openly discussed with stakeholders.

**Viktor Bolotov**, *Vice President, Russian Academy of Education*

**Galina Kovaleva**, *Head, Center for Evaluating the Quality of Education, Institute of Content and Methods of Education, Russian Academy of Education*

**Marina Pinskaya**, *Lead Research Associate, Institute for Educational Studies, National Research University Higher School of Economics*

**Igor Valdman**, *Director, Russian Training Center; Head, Laboratory for Educational Monitoring, Institute of Management in Education, Russian Academy of Education*