Science Education in Russian Schools as Assessed by TIMSS and PISA

A. Pentin, G. Kovaleva, E. Davydova, E. Smirnova

Received in January 2018

Alexander Pentin
Candidate of Sciences in Mathematics and Physics, Associate Professor, Head of the Center for Science Education, Institute for Strategy of Education Development, Russian Academy of Education (Federal State Budgetary Research Institution). Email: pentin@mail.ru

Galina Kovaleva
Candidate of Sciences in Pedagogy, Head of the Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education (Federal State Budgetary Research Institution). Email: galina_kovaleva_rao@mail.ru

Elena Davydova
Junior Researcher, Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education (Federal State Budgetary Research Institution). Email: centeroko@mail.ru

Elena Smirnova
Junior Researcher, Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education (Federal State Budgetary Research Institution). Email: centeroko@mail.ru

Address: 5/16 Makarenko St, 101000 Moscow, Russian Federation.

The article was prepared as part of the project "Modernization of the School Science Curriculum and the Methods of Teaching Natural Sciences in the Present-Day Information Contexts". Project No. 27.6122.2017/БЧ. Translated from Russian by I. Zhuchkova

Abstract. The article investigates the specific features of science education in Russian schools as they are manifested in international assessments: the decrease in science achievements between elementary and middle school and the startling difference between the TIMSS and PISA scores of eighth- and ninth-graders. Conclusions are drawn from analysis of data obtained from the international studies (test results, participant surveys, item analysis) as well as from the characteristics of science education guidelines contained in the federal education standards and curricula. Factors affecting science performance in elementary, middle, and high school are identified. For instance, the high TIMSS scores of fourth-graders are largely explained by active acquisition of scientific knowledge beyond the school walls at this age. The sharp difference between the TIMSS and PISA results of eighth- and ninth-graders has to do, on the one part, with the close correspondence between Russian science curricula and the TIMSS conception, and on the other part with the considerable disagreement between science education in Russia and the PISA conception, since the former is little oriented toward developing scientific literacy in students. The decrease in eleventh-graders' performance in TIMSS Advanced-2015 (advanced physics) as compared to the previous cycles may be due to, among other things, the increase in the percentage of items on atomic/nuclear physics, which turned out to be the cause of more difficulties for students. The TIMSS, PISA, and TIMSS Advanced results of 2015 indicate that science education in Russian schools is aimed more at acquiring and demonstrating knowledge rather than applying it or learning the scientific procedures and practices, i. e. evaluating and designing scientific enquiry, interpreting data and evidence scientifically. The novelty of this study is that it brings together international assessment data on the quality of science education in Russia at all levels of school education for the first time.

Keywords: science education, international comparative studies on educational quality, PISA, TIMSS, educational attainment, scientific literacy.

DOI: 10.17323/1814-9545-2018-1-79-109

The year 2015 witnessed a coincidence of regular cycles of as many as three international studies of secondary education quality: TIMSS, PISA, and TIMSS Advanced. All three are designed, among other things, to assess school student attainment in natural sciences across countries. The TIMSS, administered in four-year cycles, measures the quality of science education among students from the fourth and eighth grades. The PISA (conducted every three years) tests the so-called scientific literacy of 15-year-olds. The TIMSS Advanced evaluates every seven years the educational attainment of students in their final year of secondary school (eleventh grade in Russia) enrolled in special advanced physics programs or tracks. Russia participates in all three assessments, so the 2015 results can be regarded as a fairly informative dimension of horizontal data on school science education.

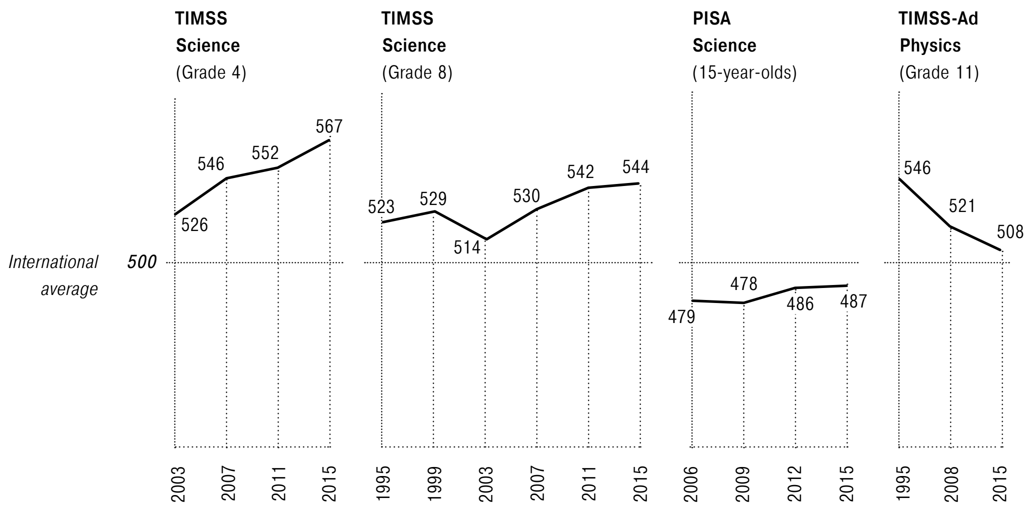
Figure 1 presents Russia's science education results in the cycles of all three studies, measured by the national sample average; the continuous line corresponds to the international average of 500 points.

The graphs present a rather controversial picture:

- Russia's results are higher than the international average in all the assessments except for PISA;
- Not only are the TIMSS results in the fourth grade high in absolute terms but they also reveal a positive trend;
- The 2015 results in the eighth grade show very little change as compared to the previous TIMSS cycle in 2011, i. e. there is no trends in this time segment;
- Fifteen-year-old students (mostly ninth-graders) performed lower than average in the PISA Science tests, and there is almost no difference between the results of 2012 and 2015. However, if we compare the 2006–2015 period, when science literacy was PISA's top priority, we will see some progress over the period (8 points);
- The eleventh grade demonstrates an essential downward trend in performance in advanced physics, yet Russia's 2015 score is still higher than the international average.

Obviously, these oppositely-directed trends need to be analyzed in depth and commented upon.

Fig. 1



The Phenomenon of Russian Elementary School

Table 1 shows the TIMSS-2015 results of fourth-graders in Science [Martin et al. 2016]. Only two countries, Singapore and South Korea, demonstratesd better performance, and Japanese students performed nearly the same as their Russian peers. The results in the other 43 countries are significantly lower than in Russia. Moreover, Russian elementary school graduates have shown continuous progress since 2003, when Russia participated in the TIMSS fourth grade for the first time (Fig. 2) [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016b]. Russian scores have grown by 41 points on the international scale over this period. As compared to the previous study cycle (2011), Russian students scored 15 points higher in Science in 2015, which is not typical of any country. Seventeen countries improved their performance between 2011 and 2015, while 16 remained where they were, and eight showed a decrease.

Meanwhile, the high results of Russian fourth-graders in science are even more surprising than, say, the unsatisfactory results of 15-year-olds in PISA science literacy. The truth is that the TIMSS Science Framework for the fourth grade [Martin et al. 2016], designed in collaboration with all the participating countries, goes far beyond the curriculum of the “The World Around Us” (TWAS) course delivered in Russian elementary schools[1].

[1] Elementary Education Curriculum Guidelines (2015). Approved by the Resolution of the Federal Academic Association for School Education of April 8,

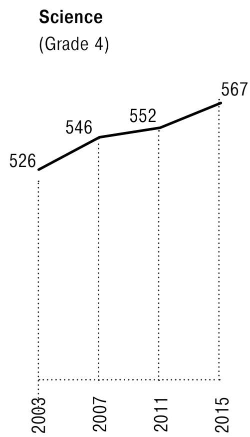
Table 1. **TIMSS-2015 Results in Fourth Grade Science**

Country	Average Score	Country	Average Score
1. Singapore	590 (3.7)	25. Australia	524 (2.9)
2. Republic of Korea	589 (2.0)	26. Slovak Republic	520 (2.6)
3. Japan	569 (1.8)	27. Northern Ireland	520 (2.2)
4. Russian Federation	567 (3.2)	28. Spain	518 (2.6)
5. Hong Kong (SAR)	557 (2.9)	29. Netherlands	517 (2.7)
6. Taiwan	555 (1.8)	30. Italy	516 (2.6)
7. Finland	554 (2.3)	31. Belgium (Flemish)	512 (2.3)
8. Kazakhstan	550 (4.4)	32. Portugal	508 (2.2)
9. Poland	547 (2.4)	33. New Zealand	506 (2.7)
10. United States	546 (2.2)	**TIMSS Scale Average**	**500**
11. Slovenia	543 (2.4)	34. France	487 (2.7)
12. Hungary	542 (3.3)	35. Turkey	483 (3.3)
13. Sweden	540 (3.6)	36. Cyprus	481 (2.6)
14. Norway	538 (2.6)	37. Chili	478 (2.7)
15. England	536 (2.4)	38. Bahrein	459 (2.6)
16. Bulgaria	536 (5.9)	39. Georgia	451 (3.7)
17. Czech Republic	534 (2.4)	40. UAE451	(2.8)
18. Croatia	533 (2.1)	41. Qatar	436 (4.1)
19. Ireland	529 (2.4)	42. Oman	431 (3.1)
20. Germany	528 (2.4)	43. Iran	421 (4.0)
21. Lithuania	528 (2.5)	44. Indonesia	397 (4.8)
22. Denmark	527 (2.1)	45. Saudi Arabia	390 (4.9)
23. Canada	525 (2.6)	46. Morocco	352 (4.7)
24. Serbia	525 (3.7)	47. Kuwait	337 (6.2)

According to the international analytical report [Ibid.], TIMSS-2015 tests could assess knowledge in all natural sciences taught in school: life science (biology), physical science (physics and chemistry), and Earth science (geography). These content domains were represented in the TIMSS-2015 tests for the fourth grade in the following ratio: life science-45 percent, physical science-35 percent, and Earth sci-

2015. Minutes No. 1/15. Register of School Education Curriculum Guidelines. http://mosmetod.ru/files/dokumenty/primernaja-osnovnaja-obrazovatelnaja-programma-osnovogo-obshchego-obrazovanija.pdf

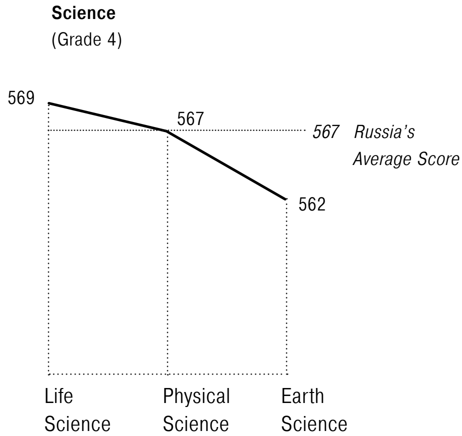
Fig. 2. **The Progress of Russian School Students in the TIMSS Fourth Grade Assessment over the Entire Period of Participation** (International Scale Average)



ence-20 percent. Russia's Elementary Education Curriculum Guidelines (2015) shows prevalence of the biology and geography components, often interrelated on the level of natural communities, which together account for about 95 percent of all the scientific content in TWAS. As little as 5–6 percent is left to physical sciences, which only give the most general idea of the diversity of matter and the properties of the three main states of matters. Meanwhile, the TIMSS-2015 tests asked fourth-graders to demonstrate the following knowledge and skills in physical science [Ibid.]:

- Compare and classify objects and materials by their physical properties (weight/mass, volume, state of matter, thermal or electrical conductivity, whether an object sinks or floats in water);
- Know the properties of metals (electrical and thermal conductivity) and relate them to metal applications;
- Give examples of mixtures and explain how they can be separated into ingredients using physical methods (screening, filtering, evaporation, or magnetism);
- Know ways of speeding up the rate of dissolving matters in a given amount of water (heating, stirring, increasing surface area) and compare concentrations of two solutions with differing amounts of solvent/solute;
- Recognize the observed transformations of substances that result in new substances with different properties (decay, burning, rusting, boiling);
- Associate familiar physical phenomena (shadows, reflection, rainbow) with the properties of light;

Fig. 3. **Russian Fourth-Graders' TIMSS-2015 Scores in Three Content Domains of Science** (International Scale Average)



- Know that vibrating objects can produce sound;
- Know that magnets have "north" and "south" poles, that like poles repel and opposite poles attract;
- Know that electrical energy in an electric circuit can be converted into other forms of energy, e. g. light and sound;
- Explain that a closed electric circuit is required for the operation of simple electrical systems, e. g. a pocket lantern;
- Know that forces (pushing or pulling action) can change the direction of motion and compare the effects of forces of different magnitudes when they are directed the same or oppositely.

This list of knowledge and skills specified in the TIMSS Physical Science Module but lacking in the Russian program is nowhere near complete. There is much more matching in biology and geography, but even here, for example, the extensive topic of "Life Cycles, Reproduction, and Heredity" is not represented in the Russian "The World Around Us" course. For the complete list of TIMSS topics lacking in the Russian science curriculum, see, for instance, [Demidova 2017].

The TIMSS-2015 results of Russian fourth-graders in three content domains of science are shown in Figure 3 [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016b].

Russia is among those countries with the lowest allocation of science education time in elementary school. Russian pupils receive slightly under 50 hours of science education per annum in the fourth grade, as compared to 96 hours in Singapore, 92 in South Korea, and 91 in Japan. Besides, there are countries that allocate significantly more time to science education for elementary pupils than Russia: 162 hours in Portugal, 161 in Chili, 135 in Qatar, 110 in Georgia [Martin et al. 2016]. However, they are not ranked among the top TIMSS performers (see Table 1).

To summarize the phenomenon, Russian pupils do not study much of what is assessed in TIMSS and what is part of the science curricula in most countries. In addition, elementary schools in Russia allocate much less time to science education than most countries. However, Russian fourth-graders are among the highest achieving students in the TIMSS ranking, which means they succeed in performing the tests of this international study.

How can this phenomenon be explained? The report [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016b] identifies the following factors that differentiate Russian fourth-graders from their foreign peers (data was obtained from surveys of students, their parents, teachers and administrators of the participating schools):

1) Russian fourth-graders (aged 10.8 on average) are more than six months older than their peers in the other participating countries (whose average age is 10.2), with the exception of two countries, Norway and Denmark. Empirical studies have proved that fourth-graders' scores in PIRLS (reading literacy) reveal an obvious correlation with pupil age even when the difference is as small as a few months [Van Damme et al. 2010];
2) The level of education of elementary school graduates' parents in Russia is one of the highest in the world;
3) Russian parents are ranked first in the amount of time devoted to preschool home education activities;
4) Ninety-three percent of Russian fourth-graders do their homework every day, which is only second to Kazakhstan and Japan (94 percent in both);
5) Russian parents render a great deal of support to their children in learning: 72 percent of parents make sure their children do their homework on a daily basis, which is also one of the highest rates across all the participating countries;
6) Russian elementary school teachers are among the oldest ones: 81 percent of teachers are aged above 40 (as compared to the international average of 58 percent), and only 5 percent are younger than 30. An average Russian teacher has 25 years of teaching experience, and 78 percent have been teaching for more than two decades, as compared to the international average of 17 years (15 years in the top-ranking countries). Elementary school teachers in Russia are 100 percent female, while the other countries have on average 18 percent of male teachers;
7) In all the countries, pupils who like studying natural sciences demonstrate on average better performance. However, the cor-

relation is not applicable to a number of countries including Russia: 58 percent of Russian fourth-graders who liked the TWAS course scored only 4 points higher in TIMSS than the eight percent of children who did not like the class. Moreover, the proportion of Russian pupils who like natural sciences has even dropped from 62 to 58 percent since 2011;

8) Eighty percent of pupils in Russia are fully engaged in science lessons (as compared to the international average of 69 percent) and 18 percent show weak engagement (the international average is 25 percent). No difference was observed in the results between the two groups.

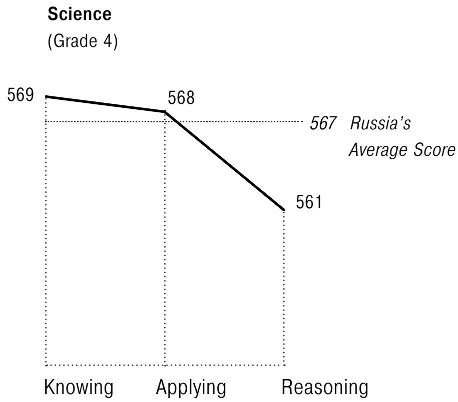
Naturally, these factors are not enough to explain the observed phenomenon of Russian fourth-graders' excellent performance in science. However, they reveal the grounds for a simple and rather obvious hypothesis: with regard to Russian fourth-graders, the TIMSS Science test measures the quality of extracurricular education rather than that of education in the classroom.

Apart from the lack of many TIMSS topics in TWAS and the small amount of time allocated to science education, the hypothesis is supported by the following:

- Russian fourth-graders are older than their peers in other countries (point 1), which means they have slightly more life and learning experiences;
- Parents devote more time to the education of their children (points 3 and 5) and have high levels of education themselves (point 2);
- There is almost no difference in the scores between children who like science lessons and those who do not (point 7) as well as between pupils who engage more and less in the learning process (point 8);
- Although the interest in science lessons has decreased since 2011, the TIMSS results have improved (see Fig. 2);
- It is hard to expect from Russian elementary school teachers, who are older than their foreign colleagues (point 6), that they will apply advanced science teaching methods to compensate for content deficiencies.

Of course, all of this raises a question, where do elementary pupils get their science knowledge from, if not from school lessons? And why could Russian pupils find themselves in a more advantageous position than students in most other countries? The answer to these questions implies in-depth analysis of the learning environment (in its broad sense) of this age group, which obviously includes family, the system of supplementary education (which is well-developed in Russia), and science fiction TV shows and cartoons. It could be that the inadequate science component of elementary school education has

Fig. 4. **Russian Fourth-Graders' Scores in TIMSS Tests Assessing Scientific Thinking in Various Cognitive Domains** (International Scale Average)

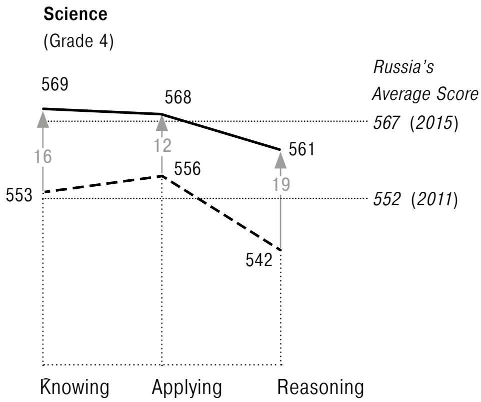


had a positive effect by making room for children's natural curiosity beyond the school walls.

The TIMSS also discriminates between tests assessing scientific thinking in different cognitive domains: reproducing factual knowledge ("knowing"), applying knowledge in real-life contexts ("applying"), and explaining phenomena or describing one's own observations ("reasoning") [Demidova 2017]. Russian fourth-graders tend to perform better in knowing and applying (Fig. 4) [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016b], while the reverse is true for the three countries ranking above Russia (see Table 1), where pupils are more successful in reasoning. Perhaps, this effect is also a consequence of the extracurricular nature of science education in Russia. Children absorb information and knowledge from various sources easily but are less likely to learn systematically how to explain, reason, or solve problems.

Figure 5 shows the contribution of different thinking processes to the overall improvement in TIMSS results between 2011 and 2015. Reasoning appears to have improved most of all (+19 points). Development of the cognitive skills that can be listed in the "reasoning" domain—make inferences, determine relationships between objects, analyze information, and create simple models—is defined by the new standard of elementary education (Federal State Education Standard of Elementary Education), introduced in 2011, as expected educational outcomes. Such meta-subject skills could manifest themselves in science performance. However, by promoting the development of a number of qualities required for scientific reasoning, the standard does not eliminate the deficiencies in the content of science education in Russian elementary schools.

Fig. 5. **The Contribution of Different Cognitive Skills to the Improvement in TIMSS Results Between 2011 and 2015**
(Points on a 1,000-Point Scale)



Analysis of Russian fourth-graders' TIMSS-2015 scores and the changes in them provides the grounds for identifying the reforms that science education in Russian elementary schools needs in order to conform to the present-day standards of teaching science to elementary pupils. Such reforms should not be limited to expanding curricula by including new content modules. Children of this age group happily obtain most of the knowledge lacking in the curriculum beyond the school walls. Consequently, changes should just as much, or even to a greater extent, affect science teaching methods, redesigning them to maintain and encourage pupils' curiosity and natural desire to explore nature.

Why PISA and TIMSS Results Differ That Much

Another curious feature of science education in Russian schools that reveals itself in the international assessments consists in the huge gap between results in PISA (15-year-old students) and TIMSS (eighth grade). Indeed, while Russia was ranked in thirties with a below-average score in PISA 2015 [OECD 2016; Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016a], Russian students ranked seventh in TIMSS-2015 with a score high above the average [Martin et al. 2016]. A similar pattern was observed in the previous cycles of both assessments. For analysis of the differences in Russian school students' results in mathematics and science between PISA-2003 and TIMSS-2003, see [Kovaleva 2006], and between PISA-2006 and TIMSS-2007 [Polivanova 2015].

To ensure a meaningful comparison, it makes sense to examine only the countries that participated in both studies in 2015. In Table 2,

Table 2. **Science Performance in Countries that Participated in PISA-2015 and TIMSS-2015 (Grade 8)**

#	Country	PISA Score	PISA RANK	TIMSS Score	TIMSS RANK	DIFF RANK
1	Canada	528	4	526	12	8
2	New Zealand	513	7	513	15	8
3	Australia	510	9	512	16	5
4	Norway	498	12	509	17	5
5	Chili	447	21	454	25	4
6	Italy	481	16	499	19	3
7	Hong Kong (SAR)	523	5	546	6	1
8	Malta	465	20	481	21	1
9	Singapore	556	1	597	1	0
10	Japan	538	2	571	2	0
11	Taiwan	532	3	569	3	0
12	UAE	437	22	477	22	0
13	Thailand	421	24	456	24	0
14	Georgia	411	26	443	26	0
15	Jordan	409	27	426	27	0
16	Lebanon	386	28	398	28	0
17	Sweden	493	14	522	13	–1
18	Israel	467	19	507	18	–1
19	Republic of Korea	516	6	556	4	–2
20	Great Britain	509	10	537 (England)	8	–2
21	Ireland	503	11	530	9	–2
22	Qatar	418	25	457	23	–2
23	Slovenia	513	8	551	5	–3
24	United States	496	13	530	10	–3
25	Turkey	425	23	493	20	–3
26	Lithuania	475	18	519	14	–4
27	Hungary	477	17	527	11	–6
28	Russian Federation	487	15	544	7	–8

these countries are listed using the rank-ordering method suggested in [Grønmo, Olsen 2006].

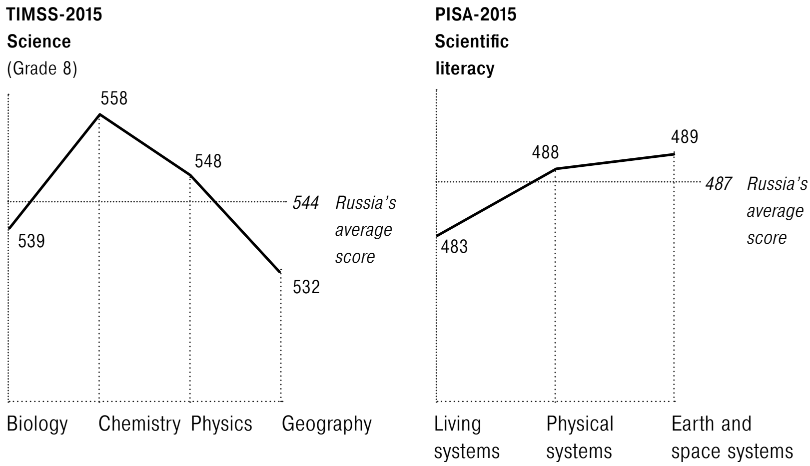
The rankings in Table 2 are determined by the relative position of these 28 countries in both studies. The countries are sorted by the decreasing difference between their ranks in PISA and TIMSS, positive numbers corresponding to higher rank in PISA.

Most countries appear to have close results in both assessments (the difference in rank being 2 at the most), the top-rankers in both PISA and TIMSS being Singapore, Japan, and Taiwan. Russia, Canada, and New Zealand show the greatest difference in ranks, but Russia scores better in TIMSS, in contrast to Canada and New Zealand.

What are the reasons for such a wide gap between Russian school students' science scores in TIMSS and PISA, and what inferences about science education in Russian schools can be made based on this difference? It is unlikely that the key lies in the age difference of 12–18 months or a sharp decline in quality of science education between the eighth and ninth grades, although these hypotheses need to be tested, too. A more plausible assumption is that TIMSS and PISA goals and, consequently, instruments are profoundly different and, crucially, that this difference has become critical in the polarization of advantages and disadvantages of Russian school science education.

The TIMSS and PISA approaches to student assessment and performance of different countries in these two assessments have been compared in a number of studies, e. g. [Olsen 2005; Grønmo, Olsen 2006; Hutchison, Schagen 2007; Tyumeneva, Valdman, Carnoy 2014; Polivanova 2015; Carnoy et al. 2016; Klieme 2016]. However, most of these, with the exception of [Olsen 2005], have been based on analyzing the mathematical competencies of students in both TIMSS and PISA. The aim of TIMSS is defined as to assess knowledge and skills acquired as a result of learning mathematics and science [Martin et al. 2016]. Otherwise speaking, TIMSS demonstrates how students assimilate mathematic and scientific material from the perspective of some international standards agreed among the participating countries. The aim of PISA with regard to science is defined in a different way: to study and measure scientific literacy of students. PISA documents (e. g. [OECD 2016; Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016a]) define scientific literacy by the three competencies to: explain phenomena scientifically; evaluate and design scientific enquiry; and interpret data and evidence scientifically. Of no less importance—this is also part of the definition—is that these competencies are needed for students to be informed critical consumers of scientific knowledge and engage in critical discussion about issues that involve science and technology [OECD 2016]. Thus, PISA is designed to assess how successfully scientific competencies are applied to real-life problems and contexts—relevant but extending beyond the school curriculum.

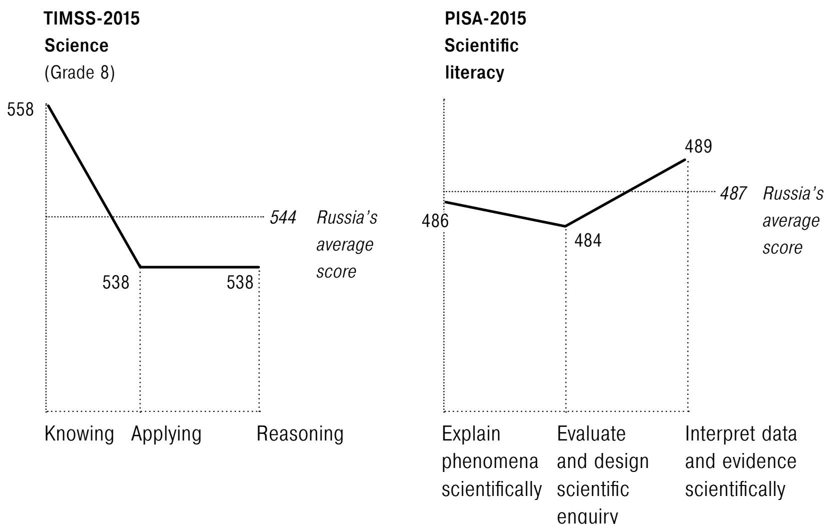
Fig. 6. **Russian Pupils' Results in Subjects and Content Domains of TIMSS-2015 and PISA-2015** (International Scale Average)



The differences in the aims of TIMSS and PISA determine the differences between their tests (instruments). In terms of content, TIMSS tests for the eighth grade are classified distinctly by the subject assessed, while PISA tests fall into three scientific fields—"living systems", "physical systems", and "earth and space systems"—that do not fit precisely into the school subjects. Yet, the fit is rather close between "life science" (TIMSS) and "living systems" (PISA), "physical science" (TIMSS) and "physical systems" (PISA), "earth science" (TIMSS) and "earth and space systems" (PISA). Russian pupils' results in subjects and content domains of both assessments are shown in Figure 6 [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016c; Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016a].

Perhaps, the only thing that these results have in common is lower performance in biology (life science / living systems) as compared to the average scores of Russian pupils in both TIMSS and PISA. Meanwhile, performance in this domain contributes a lot to the overall test score, accounting for the best part of tests in both studies: 36 percent of the TIMSS assessment focus on life science, and 40 percent of PISA is about living systems. Today, out of the three scientific subjects (physics, chemistry, biology) only biology is studied throughout the whole period of middle school in Russia (5th-9th grades) and it receives the greatest number of teaching hours during this period, as compared to physics and chemistry. However, the effects observed are rather negative.

Fig. 7. **Russian Pupils' Science Scores in TIMSS and PISA, broken down by Cognitive Domains and Competencies**
(International Scale Average)



It is much harder to compare TIMSS and PISA by the types of competencies that they assess, as their tests are classified on essentially different bases. The TIMSS embraces the cognitive domains of "knowing" (reproducing factual knowledge and applying it in standard learning contexts), "applying" (applying knowledge in more complex contexts), and "reasoning" (explaining phenomena or describing observations and experience). The PISA, in its turn, is designed to assess three competencies: "explain phenomena scientifically", "evaluate and design scientific enquiry", and "interpret data and evidence scientifically". Figure 7 presents Russian students' results in both studies, broken down by cognitive domains and competencies [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016c; Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016a].

TIMSS tests were analyzed in the "coordinates" of PISA-assessed competencies in order to establish the correspondence between the two studies. It was possible to conventionally classify about 45 percent of TIMSS tests under the "explain phenomena scientifically" competency (as compared to 48 percent in PISA), about 5 percent of TIMSS tests under the "evaluate and design scientific enquiry" competency (21 percent in PISA), and about 9 percent of TIMSS tests under the "interpret data and evidence scientifically" competency (31 percent in PISA). The remaining roughly 40 percent of TIMSS tests do not fit into

Table 3. **The Main Differences in Science Questions Between TIMSS and PISA**

	TIMSS	PISA
1	Normally offers a standard, formalized context	Normally offers a novel, unfamiliar real-life context
2	The amount of information that needs to be processed to answer the question is small, fitting easily into the familiar, "What is known?"	The amount of information that needs to be processed to answer the question is considerable; this information rather resembles a scientific or science fiction text than a typical school problem setting
3	Items are isolated and include one or two questions	Items are grouped into topical modules, most of which include from 3 to 5 tasks interconnected by some kind of a plot
4	Few items (5 percent) assess how students evaluate and design scientific enquiry	One fifth (21 percent) of the items assess how pupils evaluate and design scientific enquiry
5	Very few items imply analysis of data from graphs or tables; the graphs look perfect instead of being drawn using test points	A number of items involve the handling of real-life scientific data presented as graphs, tables, or diagrams
6	Very few items touch upon environmental issues	Over one third of all items touch upon environmental issues one way or another

the PISA competency-based classification at all. Most of this 40 percent is represented by the so-called reproductive tasks, when a student only needs to reproduce the relevant piece of scientific content to answer a question, e.g. assign an animal with described characteristics to one of four taxonomical groups offered or say which type of energy is possessed by a compressed spring by choosing the correct answer from the four options. For a detailed analysis of TIMSS tests, see [Kamzeeva 2017].

The inverse procedure, i.e. analysis of PISA tests in the "coordinates" of TIMSS-assessed cognitive domains, shows that nearly all PISA tests can be classified under "reasoning" (although the boundary between "reasoning" and "applying" often looks rather conventional), which has been revealed by researchers comparing the TIMSS and PISA approaches to mathematical literacy assessment [Grønmo, Olsen 2006].

Therefore, the two studies overlap in the cognitive domains and relevant competencies only partially, by roughly 60 percent if judged from the PISA's perspective. However, it is probably not this factor that determines the main difference between TIMSS and PISA test questions. Item comparison between both assessments allows for identifying a number of additional differentiating parameters (Table 3).

It would appear reasonable that all these differences together are responsible for the gap in Russian pupils' results between TIMSS and PISA. To put it another way, it is the specified characteristics of PISA items that explain difficulties experienced by Russian school students. One may suggest a number of reasons behind such difficulties, like the insufficient level of teacher knowledge. However, Martin Carnoy and his colleagues [Carnoy et al. 2016] have proved that the level of teacher knowledge has an insignificant effect even on the TIMSS scores of Russian pupils in mathematics, although the content of TIMSS tests is quite in line with the Russian standards here. As for PISA, the authors believe that a number of competencies (understand and interpret texts, model real-life contexts mathematically) required to solve PISA items in mathematics have never been taught in Russian schools. It means that PISA results cannot be regarded as an indicator of Russian teachers' knowledge, as developing those competencies has never actually been part of their education or experience.

Probably, neither is teacher knowledge a decisive factor in pupils' TIMSS and PISA scientific achievements. In recent years, a series of studies have sought ways of increasing teacher influence on the academic performance of students, in particular they have discussed possible domains of teacher training to improve the development of scientific literacy among students [Pentin 2012]. A number of factors potentially affecting PISA performance have been identified, teaching methods being one of the most important of these [Kovaleva, Loginova 2017]. Student-centered teaching practices can be more useful for PISA results than the teacher-centered approach. However, Russian schools have been using passive learning methods much more extensively so far. In addition, Russian science education curricula, textbooks, and assessment instruments have the characteristics typical of TIMSS items (left column of Table 3), attending little to the tasks and material presenting methods applied in PISA (right column of Table 3).

The peculiar features of PISA items are not important so much per se; rather, they act as concrete indicators of science education orientation toward the development of scientific literacy. At the same time, the perceptions of the aims and expected outcomes of TIMSS and PISA do not have to be opposed to each other. The performance of such countries as Singapore, Japan, and Taiwan demonstrates the possibility of being equally successful in both (conventionally) applied and pure mathematics, assessed by PISA and TIMSS, respectively. Having analyzed the test results of pupils from Singapore, Japan, and Taiwan—outstanding in both assessments—Liv S. Grønmo and Rolf V. Olsen come to a conclusion that "mathematics in school in the East Asian countries to a great extent focus on pure mathematics in all topics, while at the same time they also give some attention to the full cycle of applied mathematics" [Grønmo, Olsen 2006]. And further: "Our analysis and comparisons between TIMSS and PISA support that in order to do well in daily life mathematics, students need a basis of knowledge and skills in pure mathematics. <…> This indicates that it is important in school curriculum that mathematical literacy is not seen as an alternative to pure mathematics. A reasonably high level of competence in pure mathematics seems to be necessary for any type of applied mathematics. On the other hand, if too little attention is given to the full cycle of applied mathematics, it is unlikely that students will develop the type of competence we may call mathematics literacy." The inferences that Grønmo and Olsen made about mathematics seem to be entirely applicable to natural sciences.

The root of problems that Russian pupils tend to have with PISA items is apparently that students (as well as teachers) deal little with even raising the question of applying scientific knowledge and skills to solve real-life problems. Meanwhile, the very use of real-life contexts determines nearly all the peculiarities of the PISA tasks presented in Table 3, because:

- A real-life context needs a detailed description, which entails the need to understand and process all the related information;
- The problem contained in the context normally falls into a sequence of consecutive problems or sub-problems, which is captured in the way PISA items are grouped into topical modules;
- A real-life problem often needs a researcher's approach, analysis and interpretation of data available;
- Real-life contexts that make sense for every member of society, including students, often have to do with environmental issues and human health (the "living systems" content domain).

Russia's performance in the other two PISA categories, mathematical and reading literacy, has been improving very rapidly. In maintaining high positions in TIMSS, i.e. "pure" mathematics (ranked 6th), Russian school students have shown considerable progress in applied mathematics, or mathematical literacy, over the latest two PISA cycles (between 2009 and 2015). This achievement, just like the improved results in reading literacy, is sometimes attributed to the newly integrated education standard. Indeed, the education quality requirements stipulated in the Federal State Education Standard of Middle School Education often echo the competencies assessed in PISA. However, the same educational outcomes are expected in natural sciences, yet there has been no progress in PISA-assessed scientific literacy (see Fig. 1). Possibly, the problem here lies in the disintegration among teachers of different scientific disciplines and their understandings of the present-day objectives of science education in school.

By now, we have identified a number of factors describing science education in Russian schools and potentially affecting how pupils score in TIMSS and PISA. Further research must revolve around finding correlations between these factors and TIMSS and PISA results.

Table 4. **Results in TIMSS Advanced (Advanced Physics)**

Country	Average Score	Coverage Index	Years of Formal Schooling	Average Age
1 Slovenia	531 (2,5) ↑	7,6 percent	13	18,8
2 Russian Federation	**508 (7,1)**	**4,9 percent**	**11**	**17,7**
3 Norway	507 (4,6) =	6,5 percent	13	18,8
TIMSS Scale Average	**500**			
4 Portugal	467 (4,6) ↓	5,1 percent	12	18,0
5 Sweden	455 (5,9) ↓	14,3 percent	12	18,8
6 United States	437 (9,7) ↓	4,8 percent	12	18,1
7 Lebanon	410 (4,5) ↓	3,9 percent	12	17,8
8 Italy	374 (6,9) ↓	18,2 percent	13	18,9
9 France	373 (4,0) ↓	21,5 percent	12	18,0

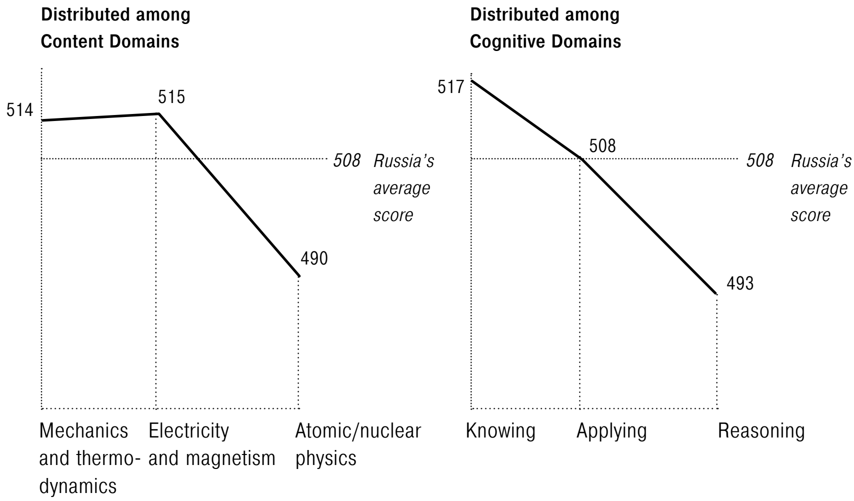
↑ Country average statistically significantly higher than Russia's average.
= No statistically significant differences between country average and Russia's average.
↓ Country average statistically significantly lower than Russia's average.
( ) Standard error of measurement

What Do Russian High School Students' Scores in Advanced Physics Indicate?

TIMSS Advanced is designed to assess final year high school students' performance in advanced physics. Therefore, results in this test can be treated as a small though important fragment in the big picture of science education in school shaped by international studies. Only nine countries participated in TIMSS Advanced in 2015, and Russia scored second highest (Table 4) [Mullis et al. 2016].

The sample of Russian students consisted of eleventh-graders (final year students) who had at least four lessons of physics per week. The key test results contained in the report [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016d] are largely in line with expectations and thus come as no surprise, except for the steady decrease in Russia's average score over the last three assessment cycles: in 1995, 2008, and 2015. However, a considerable decrease has been observed in all the participating countries except the U.S. and Slovenia. The reasons for the negative trends are yet to be established, but for Russia specifically, one of the factors could be the increased coverage index, which indicates the percentage of young people enrolled in advanced physics programs or tracks in the country's population of the given age group. The coverage index was 1.9 percent in 1995, 2.6 percent in 2008 [Mullis et al. 2009], and 4.9 percent in 2015, i. e. nearly twice as high as in 2008. Of course, the hypothesis about the cov-

Fig. 8. **Results of Russian Eleventh-Graders' Enrolled in Advanced Physics Programs and Tracks, broken down by Content and Cognitive Domains** (International Scale Average)



erage index affecting average scores needs to be tested, but the increase in the number of students enrolled in advanced physics tracks probably indicates an increase in the number of advanced physics classes and teachers within that period. It is not inconceivable that "newly-arrived" teachers of advanced physics did not have enough knowledge and experience at the beginning, which could result in lower levels of student attainment. Anyway, analysis of the reasons for the negative trend in physics in the three cycles of TIMSS Advanced would be one-sided without considering the change in the percentage of high school students enrolled in advanced physics programs.

Results in advanced physics reveal some general peculiarities and deficiencies of science education in Russian schools. In particular, they can be seen from the data on TIMSS Advanced performance in different content and cognitive domains assessed (Fig. 8) [Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016d].

In 2015, Russian high school students scored lower in atomic/nuclear physics than in other content domains. One of the possible explanations is that the relevant domain (usually referred to as "quantum physics") in the Russian advanced physics curriculum is currently the least covered. A very small proportion of USE items focus on quantum physics (Table 5). Conversely, TIMSS Advanced is focusing more and more on this domain at the expense of "mechanics". This balance of the TIMSS Advanced body of tests captures the important trend in physics and other branches of scientific education, which

Table 5. **The Distribution of Items among Physics Content Domains in TIMSS Advanced 2008 and 2015 and in USE-2017**

	TIMSS Adv 2008	TIMSS Adv 2015	USE-2017
Mechanics	29%	40%	30–35%
Molecular physics and thermodynamics	24%		23–25%
Electrical dynamics	29%	25%	30–35%
Atomic/nuclear physics	19%	35%	13–16%

consists in increasing the percentage of the latest scientific knowledge in the content of school classes, especially advanced programs. Indeed, the modern-day avenues of fundamental research (e. g. particle physics or quantum gravity) and advances in technology innovation (e. g. quantum computers, nanotechnology, nuclear and thermonuclear fusion physics) require competence in quantum physics in the first place, which cannot but be considered in the curriculum of advanced school programs.

Curiously, it was in atomic/nuclear physics that Russian eleventh-graders scored better in TIMSS Advanced in 2008 [Mullis et al. 2009]. Perhaps, the focus on the domain decreased between 2008 and 2015 due to adopting a more pragmatic orientation toward the USE requirements and structure. Simple estimates show that if the results of Russian students had been as high in "atomic/nuclear physics" as they were in "mechanics and thermodynamics" and "electricity and magnetism" in 2015 (they used to be even higher in 2008), the average score would have risen to 516–517 and would not have differed significantly from that of 2008.

Russian eleventh-graders perform much better in reproducing knowledge than in tests which assess applying and, even more so, reasoning (see Fig. 8). The distribution of TIMSS Advanced 2015 items among these cognitive domains looks as follows: 30 percent in "knowing", 40 percent in "applying", and 30 percent in "reasoning". However, this distribution appears to be questionable at times. It is quite understandable when constructed-response items asking students to explain, substantiate, or prove something using their own words are classified under the "reasoning" domain. In a number of cases, however, the same domain is assigned in TIMSS Advanced to multiple-choice or short-response items as well. Of course, some kind of internal reasoning always precedes any decision or inference, but classifying such items under the cognitive domain of "reasoning" makes the very boundaries of TIMSS Advanced thinking processes too ambiguous. Nonetheless, the results on advanced physics confirm that science education in Russian schools is oriented toward teaching students to reproduce their knowledge much more than to apply this knowledge or interpret evidence scientifically.

There is hardly any single factor on which to blame the decrease in Russia's average score on advanced physics over the three cycles of TIMSS Advanced (1995, 2008, and 2015). A comparison of TIMSS Advanced data between 2008 and 2015 does not confirm, for instance, the opinion that students have become less interested in physics and physics-related professions. In fact, the opposite is true: the percentage of students selecting engineering and computer sciences as their future majors among Russian eleventh-graders enrolled in advanced physics tracks increased between 2008 and 2015, contrary to those who chose to major in finance and business [Mullis et al. 2009; Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Education 2016d]. Probably, the negative trend in TIMSS Advanced performance has to do with other factors, such as the increased number of students enrolled in advanced physics programs or the decreased focus on the specific content domains which are given ever more attention in the TIMSS Advanced assessment.

Conclusion

A review of Russian school students' science performance in the international assessments of educational quality in 2015—TIMSS for the fourth grade, TIMSS for the eighth grade, TIMSS Advanced for the final grade, and PISA—allows for drawing some conclusions.

Analysis of not only the results but also the frameworks of these studies, developed through a process of collaboration among the participating countries, reveals that the content and methods of science education in Russia are neither unique nor unorthodox, being basically consistent with the global trends.

At the same time, the Russian approach to teaching natural sciences is rather centered around reproducing knowledge than applying it or learning to design scientific enquiry and interpret evidence scientifically. This peculiarity (which has probably developed historically) manifests itself as early on as in elementary school (see Fig. 4), becomes more conspicuous in the eighth grade (see Fig. 7), and persists even in students who choose science (physics) as their major (see Fig. 8). While the TIMSS results only reveal some minor falling behind in "reasoning" and "applying" as compared to "knowing", the overall level of science education being pretty high, in PISA this "peculiarity" becomes a sore point, which shows itself in the lack of scientific literacy among 15-year-old students, who find it the most difficult to understand the practices and procedures associated with scientific enquiry (see Fig. 7).

A separate issue in science education is how biology, or the content domain of "living systems" in PISA terms, is taught in Russian schools. According to the TIMSS data, Russia's average score in biology is somewhat better than in other content domains in elementary school (see Fig. 3), which should come as no surprise since virtually no material on other scientific subjects is contained in the TWAS curriculum. However, pupils begin to fall behind in biology as compared to other scientific subjects by middle school (see Fig. 6). The PISA "living systems" items present even more difficulty to Russian pupils as most of them fall under the competency "evaluate and design scientific enquiry", which has been the least developed in Russian school students (see Fig. 7). The main difficulty with learning biology probably stems from its content in the Russian curriculum, which mostly consist of descriptions overloaded with information that is hard to make sense of. Meanwhile, students do not develop any good idea of the methods of scientific enquiry that are needed to acquire knowledge in biology.

The results of TIMSS Advanced (advanced physics) indicate that it is not only high school physics but other sciences as well that need their curricula to be modernized to focus more on the recent scientific findings, methods of scientific enquiry, and new technology. Only then will school education satisfy society's growing demand for engineering and research professionals to ensure the development of the innovation economy.

International assessments also serve as indirect evidence of the effectiveness of making science education in school continuous. Most developed countries in the world make it imperative that science be part of the curriculum from the first year of elementary school through to the final year of high school, whether as an integrated course or as a set of systematic disciplines. Throughout the whole period of schooling, science education should always include modules of physical science, life science, and Earth science. In this regard, Russia is definitely at a disadvantage. Under the FSES of Middle School Education, there is no integrated course of science in the fifth and sixth grades, only biology and geography. Meanwhile, it is at the age of 10–12 (which corresponds to the fifth and sixth grades in Russia) that children become highly inquisitive, eager to explore nature, try out and even devise curious experiments. Normally, it is mostly at this age that students develop actively their first research skills and learn the fundamentals of scientific literacy and a scientific worldview, all of this being done within the framework of the integrated course "Science", an analogue of which used to exist in the Russian (Soviet) curriculum for many years. One-lesson-per-week courses in biology and geography cannot solve this problem completely, as they do not allow for setting up a sufficient number of short laboratory investigations demonstrating the specific aspects of the scientific method of enquiry. Besides, the two- or three-year gap in full-fledged science education (Russian students only start learning physics in the seventh grade, and chemistry in the eighth grade) results in a lot of students losing their interest in natural sciences and even forgetting the basic scientific knowledge and skills obtained in elementary school as part of the TWAS course.

One of the main conclusions that can be drawn from the results of the international assessments is that it makes no sense to oppose the rich traditions of Russian education and the modern trends in foreign education, nor does it make sense to oppose fundamental and applied knowledge. This is evidenced by the example of some East Asian countries that demonstrate outstanding science achievements in both domains. Applied learning means that fundamental (purely theoretical) knowledge is applied to solve real-life (practical) problems. To make science education in Russia more "applied", it is necessary to ensure a higher quality and a wider range of problems that are presented to students. It does not require attracting enormous resources but naturally implies enhancing the teaching methods associated with the diverse forms of working with these new types of problems.

The novelty of these findings is that they bring together international assessment data on the quality of science education in Russia at all levels of school education for the first time, making it possible to see the general regularities and problems of science education in school and to identify some of the factors affecting the performance of Russia in TIMSS, PISA, and TIMSS Advanced.

However, the article does not provide any final conclusions about Russian school students' science performance in the international assessments; rather, it suggests hypotheses that can be tested in subsequent studies.

References

Carnoy M., Khavenson T., Loyalka P., Schmidt W. H., Zakharov A. (2016) Revisiting the Relationship between International Assessment Outcomes and Educational Production: Evidence from a Longitudinal PISA-TIMSS Sample. *American Educational Research Journal*, vol. 53, no 4, pp. 1054–1085.

Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Sciences (2016a) *Osnovnye rezultaty mezhdunarodnogo issledovaniya PISA-2015* [Key Findings from PISA 2015]. Available at: www.centeroko.ru (accessed 10 January 2018).

Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Sciences (2016b) *Rezultaty mezhdunarodnogo issledovaniya TIMSS-2015 4 klass* [TIMSS2015 International Results—Fourth Grade]. Available at: www.centeroko.ru (accessed 10 January 2018).

Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Sciences (2016c) *Rezultaty mezhdunarodnogo issledovaniya TIMSS-2015 8 klass* [TIMSS-2015 International Results—Eight Grade]. Available at: www.centeroko.ru (accessed 10 January 2018).

Center for Education Quality Assessment, Institute for Strategy of Education Development, Russian Academy of Sciences (2016d) *Rezultaty mezhdunarodnogo issledovaniya TIMSS Advanced 2015, 11 klass* [TIMSS Advanced 2015– Eleventh Grade]. Available at: www.centeroko.ru (accessed 10 January 2018).

Demidova M. (2017) Osnovnye resultaty estestvenno-nauchnoy chasti issledovaniya TIMSS dlya nachalnoy shkoly [The Main Results In Natural Science Part In TIMSS For Primary Schools]. *Pedagogicheskie izmereniya*, no 2, pp.30–34.

Gronmo L. S., Olsen R. V. (2006) TIMSS *versus* PISA*: The Case of Pure and Applied Mathematics*. Paper presented at the 2nd IEA International Research Conference, Washington, DC, 8–11 November, 2006.

Hutchison D., Schagen I. (2007) Comparisons between PISA and TIMSS: Are We the Man with Two Watches? *Lessons Learned: What International Assessments Tell Us about Math Achievement* (ed. T. Loveless), Washington, DC: Brookings, pp. 227–261.

Kamzeeva E. (2017) Osobennosti vypolneniya rossiyskimi vosmiklassnikami sadaniy po estestvoznaniyu mezhdunarodnogo issledovaniya TIMSS [Specifics Of Natural Science Items For Russian 8th Graders In TIMSS]. *Pedagogicheskie izmereniya*, no 2, pp. 56–62.

Klieme E. (2016) *TIMSS2015 and* PISA *2015. How Are They Related on the Country Level?* Deutsches Institut für Internationale Pädagogische Forschung Working Paper published online, 12 December, 2016.

Kovaleva G. (ed.) (2006) *Kachestvo obshchego obrazovaniya v rossiyskoy shkole: po rezultatam mezhdunarodnykh issledovaniy* [The Quality of Education in Regular Russian Schools: International Findings], Moscow: Logos.

Kovaleva G., Loginova O. (2017) Uspeshnaya shkola i effektivnaya sistema obrazovaniya: kakie factory pomogayut priblizitsya k idealu? (Po dannym issledovaniya PISA-2015) [SuccessFul School And Effective System Of Education: Which Factors Help To Approach The Ideal? (Based On PISA-2015)]. *Pedagogicheskie izmereniya*, no 2, pp 69–80.

Martin M. O., Mullis I. V.S., Foy P., Hooper M. (2016) TIMSS2015 International Results in Science. *Boston College,* TIMSS *&* PIRLS *International Study Center website*. Available at: http://timssandpirls.bc.edu/timss2015/international-results/ (accessed 10 January 2018).

Mullis I. V.S., Martin M. O., Foy P., Hooper M. (2016) TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics. *Boston College,* TIMSS *&* PIRLS *International Study Center website*. Available at: http://timssandpirls.bc.edu/timss2015/international-results/advanced/ (accessed 10 January 2018).

Mullis I. V.S., Martin M. O., Robitaille D. F., Foy P. (2009) *TIMSS Advanced 2008 International Report: Findings from IEA's Study of Achievement in Advanced Mathematics and Physics in the Final Year of Secondary School*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

OECD (2016) *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematics and Financial Literacy*. Paris: OECD.

Olsen R. V. (2005). *Achievement Tests from an Item Perspective. An Exploration of Single Item Data from the PISA and TIMSS Studies, and How Such Data Can Inform Us about Students' Knowledge and Thinking in Science* (PhD Thesis). Oslo: University of Oslo.

Pentin A. (2012) Ot zadachi formirovaniya estestvennonauchnoy gramotnosti k neobkhodimym kompetentsiyam uchiteley estestvennonauchnykh distsiplin [From the Need to Develop Scientific Literacy to Required Competencies of Science Teachers]. *Nepreryvnoe pedagogicheskoe obrazovanie.ru*, no 1, p. 158.

Polivanova K. (2015) Obrazovatelnye rezultaty osnovnoy shkoly v kontekste mezhdunarodnykh issledovaniy [Educational Outcomes of Middle School in the Context of International Research]. *Psychological Science and Education*, vol. 20, no 4, pp.19–30. doi: 10.17759/pse.2015200402

Tyumeneva Y., Valdman A., Carnoy M. (2014) Chto dayut predmetnye znaniya dlya umeniya primenyat ikh v novom kontekste. Pervye rezultaty sravnitelnogo analiza TIMSS-2011 i PISA-2012 [What Does Subject Knowledge Give for Its Applying in New Context. The First Results from Studies TIMSS-2011 and PISA-2012]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 1, pp. 8–24.

Van Damme J., Liu H., Vanhee L., Pustjens H. (2010) Longitudinal Studies at the Country Level as a New Approach of Educational Effectiveness: Explaining Change in Reading Achievement (PIRLS) by Change in Age, SES and Class Size. *Effective Education*, vol. 2, no 1, pp. 53–84.