IRSTI 04.51.53 https://doi.org/10.26577/JPsS202594306

**Zh. Aubakirova[1*], A. Amrin[1], S. Kaztayeva[1], K. Tenisheva[2]**

[1]Corporate Fund "El Umiti Fund", Astana, Kazakhstan
[2]European University, St. Petersburg, Russia
[*]e-mail: zhaniya.abkr@gmail.com

# EDUCATIONAL INEQUALITY AND RURAL STUDENTS' OPPORTUNITIES FOR HIGH ACHIEVEMENT

This article focuses on the factors influencing the participation of rural schools in Kazakhstan in the national "Myn Bala" Olympiad, as well as the individual and institutional characteristics that enable students to achieve high results. The key question is: which individual and structural factors are associated with the success of rural school students in the Olympiad? The study uses open data on school infrastructure, participants, and Olympiad results.

The research applies statistical modeling methods (regression analysis) with elements of spatial analysis to identify the mechanisms of inequality reproduction. The results show that, on average across the country, students living in district centers score higher than their peers from rural areas. Student participation in the Olympiad and their educational achievements are geographically uneven, while school infrastructure (libraries, computers, and internet speed) is not significantly associated with student academic performance. A gender gap is also observed: girls perform better in tasks related to native and English languages, whereas boys score higher in science and mathematics tests. Individual student characteristics (gender, language, place of residence) have a smaller impact on general ability tests compared to subject-specific tasks (mathematics, science, etc.).

This study contributes to the ongoing academic and practical discourse on educational inequality and the specifics of learning in rural schools. The practical significance of the research lies in developing effective strategies to support schools outside urban areas and providing arguments for shaping policies aimed at improving the quality of school education in the Republic of Kazakhstan.

**Keywords:** social inequality, educational inequality, rural schools, giftedness, Kazakhstan.

Ж.К. Аубакирова[1*], А.К. Амрин[1], С.С. Казтаева[1], К.А. Тенишева[2]

[1]«Ел үміті» қоры» корпоративтік қоры, Астана, Қазақстан
[2]Еуропа университеті, Санкт-Петербург, Ресей
[*]e-mail: zhaniya.abkr@gmail.com

## Білім беру теңсіздігі және ауылдық оқушылардың жоғары жетістіктерге жету мүмкіндіктері

Мақала Қазақстандағы ауылдық мектептердің «Мың бала» ұлттық олимпиадасына қатысуына әсер ететін факторларға, сондай-ақ оқушылардың жоғары нәтижелер көрсетуіне мүмкіндік беретін жеке және институционалдық сипаттамаларға арналған. Негізгі сұрақ – ауылдық мектеп оқушысының Олимпиададағы табыстылығымен байланысты жеке және құрылымдық факторлар қандай. Зерттеуде мектептердің инфрақұрылымы, қатысушылары және Олимпиада нәтижелері туралы ашық деректер қолданылған.

Мақалада теңсіздіктің қайта өндірілетін тетіктерін анықтау үшін статистикалық модельдеу (регрессиялық талдау) және кеңістіктік талдау элементтері пайдаланылды. Нәтижелер көрсеткендей, ел бойынша аудан орталықтарында тұратын оқушылар ауылдық жерлердегі құрдастарына қарағанда орта есеппен жоғары балл жинайды. Оқушылардың Олимпиадаға қатысу деңгейі мен олардың білім жетістіктері географиялық тұрғыдан біркелкі емес, мектеп инфрақұрылымы (кітапханалар, компьютерлер және интернет жылдамдығы) оқушылардың білім жетістіктерімен айтарлықтай байланысты емес. Сондай-ақ гендерлік алшақтық бар: қыздар ана тілі және ағылшын тілі тапсырмаларын жақсы орындайды, ал ұлдар жаратылыстану және математика тесттерінде жоғары нәтиже көрсетеді. Оқушылардың жекелеген сипаттамалары (жынысы, тілі, тұрғылықты жері) балалардың жалпы қабілет тесттерін орындауына пәндік тапсырмаларға (математика, жаратылыстану және т.б.) қарағанда аз әсер етеді.

Бұл зерттеу ауылдық мектептердегі білім беру теңсіздігі мен оқыту ерекшеліктері жөніндегі академиялық және практикалық пікірталасқа үлес қосады. Зерттеудің практикалық маңызы

 

қаладан тыс мектептерді қолдауға арналған тиімді стратегияларды әзірлеу және Қазақстан Республикасында мектептегі білім сапасын арттыру саясатын қалыптастыру үшін дәлелдер ұсыну.

**Түйін сөздер:** әлеуметтік теңсіздік, білім беру теңсіздігі, ауылдық мектептер, дарындылық, Қазақстан.

Ж.К. Аубакирова[1*], А.К. Амрин[1], С.С. Казтаева[1], К.А. Тенишева[2]

[1]Корпоративный фонд «Фонд "El Umiti"», Астана, Казахстан
[2]Европейский университет, Санкт-Петербург, Россия
[*]e-mail: zhaniya.abkr@gmail.com

**Образовательное неравенство и шансы сельских учеников на высокие достижения**

Статья посвящена факторам участия сельских школ Казахстана в национальной олимпиаде «Мың бала», а также индивидуальным и институциональным характеристикам, позволяющим учащимся показывать высокие результаты. Ключевой вопрос – какие индивидуальные и структурные факторы связаны с успешностью ученика в сельской школе на Олимпиаде. В исследовании используются открытые сведения об инфраструктуре школ, участниках и результатах Олимпиады.

В работе применены методы статистического моделирования (регрессионный анализ) с элементами пространственного анализа для поиска механизмов воспроизводства неравенства. Результаты показывают, что в среднем по стране школьники, проживающие в районных центрах, набирают более высокие баллы, чем их сверстники из сельской местности. Вовлечённость школьников в Олимпиаду и их образовательные успехи географически неоднородны, школьная инфраструктура (библиотеки, компьютеры и скорость интернета) не связаны с образовательными успехами школьников. Также обнаружен гендерный разрыв: девочки лучше справляются с заданиями по родному и английскому языку, мальчики – с тестами по естествознанию и математике. Индивидуальные характеристики учеников (пол, язык, место проживания) оказывают меньший эффект на решения тестов на общие способности детей, в отличие от задач по конкретным предметам (математика, естествознание и прочее). Исследование делает вклад в уже существующую академическую и практическую дискуссию об образовательном неравенстве и специфике обучения в сельских школах. Практическая значимость исследования заключается в выработке эффективных стратегий по поддержке школ за пределами городов и даст аргументы для выработки политики повышения качества школьного образования в Республике Казахстан.

**Ключевые слова:** социальное неравенство, образовательное неравенство, сельские школы, одарённость, Казахстан.

## Introduction

In modern social sciences, researchers identify many different factors that influence a child's future success. These can be individual abilities, family, environment, as well as school characteristics: infrastructure, quality of teaching staff, internal organizational climate, etc. In this context, inequality in the quality of school education can affect not only the career prospects of a particular child, but also entail global negative consequences in the context of the entire country. Lack of opportunities for quality education reduces the chances of higher-paying and stable work, which leads to economic difficulties and an increase in poverty. These factors can increase the uneven development of regions, as well as an increase in crime, undermining social stability and order.

To understand the structural features of inequality in different countries, large international studies are currently being conducted, for example, PISA, PIRLS, TIMMS. These projects are aimed at assessing the educational achievements of schoolchildren and provide a comparative analysis of education systems in different countries (PISA, 2018), (PIRLS, 2021), (TIMMS, 2023). Kazakhstan's disappointing position in these rankings has increased the attention of the state and the public to the problem of inequality. Although the topic of educational inequality in Kazakhstan, in particular the situation of rural schools, is actively discussed in society, these discussions are often based on expert opinions, but not on empirical research that takes into account the social, economic and cultural characteristics of the region.

*The object of the study* is 6th grade students of rural schools in Kazakhstan. The work uses the results of the National Olympiad "Myn Bala", which helps to identify talented schoolchildren in rural areas and provides the winners with the opportunity to study in the best schools for gifted children. The

Olympiad is aimed at reducing the gap between rural and urban schoolchildren (National Olympiad "Myn Bala", 2023).

*The subject of the study* is the individual characteristics of students and the structural characteristics of schools associated with the achievements of schoolchildren at the Olympiad. The work uses data on the characteristics of rural schools from the information system "National Educational Database" of the Republic of Kazakhstan (National Educational Database of the Republic of Kazakhstan, 2023).

*The aim of the study* is to determine the individual and structural factors associated with the success of a student in a rural school at the Olympiad. The answer to it will allow the development of effective strategies to support rural schools and justify measures for developing a policy to improve the quality of school education in the Republic of Kazakhstan.

The data used allow us to draw conclusions that are relevant specifically for Kazakhstan, which is important due to the lack of similar works on local material, recognized by the expert community.

**Literature review**

Inequality in education is a pressing global issue, common to both developed and developing countries. According to research by Porta et al., countries with high levels of inequality in education show low levels of innovation, low levels of production efficiency, and a tendency to transmit poverty from generation to generation (Porta, 2011). International researchers identify various factors that influence the academic performance of schoolchildren. Among them are individual student characteristics, the socioeconomic status of the family, school resources, and the qualifications and experience of teachers.

*Individual characteristics of the student and the socio-economic status of the family*

A number of studies in the literature have examined the influence of socio-demographic characteristics such as gender, location, and family socio-economic status on academic performance. Hyde et al. conducted a study analyzing psychological reports and found that there were both gender similarities and differences in student performance. For example, girls outperform boys in calculus in primary and secondary schools, while only a small proportion of boys in high school excel in problem solving. The differences in girls' and boys' performance between countries are greater than within countries (Hyde, 2007: 599).

A study by Duckworth et al. noted that psychological characteristics, namely self-discipline, were found to be a stronger predictor of academic performance than IQ. Students with high levels of self-control received higher grades, had better test scores, and were more likely to achieve long-term academic goals (Duckworth, 2005: 940).

Hanushek et al., analyzing the scores of international school tests, come to the conclusion that the place of residence of the student and his family – the education of parents, home resources (especially the number of books), immigrant status, the language spoken at home are strong predictors of academic performance. (Hanushek, 2011: 117). At the same time, the presence of books at home is one of the most stable predictors of academic performance in different countries. This is confirmed by other studies. Thus, Budiongan et al. among the factors contributing to the improvement of academic results are the presence of a large number of books at home, high professional status and income of parents, participation of children in extracurricular activities, as well as higher education of parents. These factors not only directly affect the opportunities of children in the educational environment, but also create a favorable cultural and social climate in the family (Budiongan, 2024: 395). Schütz et al. note that the presence of books at home has a particularly significant impact in countries where schools are divided depending on the academic abilities of students. (Schütz, 2008: 283).

In the context of Kazakhstan, the place of residence, in particular the unattractive economic situation of the village, the difficult basic living conditions of families in the village are directly related to access to educational opportunities for children, educational resources, quality infrastructure and qualified teaching staff.

*School Resources and Influence*

In his work, Fini examines educational inequality using three types of educational institutions (vocational school, technical school, academic high school) as an example and concludes that the difference in the achievements of students in these institutions could hypothetically be neutralized if two conditions were met. First, if it were possible to achieve equality of opportunity for children with different family environments, i.e. with different family sizes, levels of education of parents and family members, professional status, and behavioral orientations toward children. Second, the same equality is necessary in the school environment: the social composition of students, and the attitude of

teachers toward students. He concludes that "differentiation of school types makes a decisive contribution to maintaining a high level of social stratification and a low level of intergenerational mobility" (Fini, 2007: 504). McPherson suggests that primary examination and test results may be an inaccurate indicator of school performance unless they are adjusted for differences in the composition of the student body, such as previous achievement and socioeconomic status. As a solution, the author applies a 'value added' model, which compares student performance across stages, using baseline assessments to estimate expected progress and then measuring the difference between expected and actual results, thereby isolating the contributions of schools and teachers (McPherson, 1993: 45).

A study by Somers et al. in Sweden found that independent schools, which are publicly funded but privately run, perform better academically than municipal schools. This is explained by their greater operational autonomy and competitive allocation of resources. For example, students in independent schools scored higher on standardized tests than their peers in municipal schools (Somers, 2001: 69). These findings are supported by the results of Wikstrom, who analyzed the Swedish SweSAT test of academic ability. The study found a statistically significant increase in academic performance in regions with a higher proportion of independent schools, where innovations are more quickly introduced into the educational process (Wikstrom, 2005: 34).

*Qualifications and experience of teachers*

Sociologist Coleman explains differences in school performance not only by the social background of students. He came to the conclusion that improving the quality of teaching, creating a healthy social climate in the educational institution and the practical orientation of school education can help children from poor families improve their academic performance. The author acknowledges that family background plays a decisive role, but the school and qualified teachers can soften this influence, thereby ensuring the student's chances for success (Coleman, 1968: 12).

According to Fuller et al., school factors have a significant impact on the effectiveness of learning: infrastructure, class size, experience and qualifications of teachers, availability of teaching materials (Fuller, 1994: 120). According to the study by Rivkin et al., an increase in the indicator "quality of teacher preparation" by one point is equivalent to a reduction in the average class by 10 students. It is important to remember that this effect is also expected in the opposite direction. Even the best teacher will not be able to convey the material fully to a class that is too large, but at the same time, a poorly prepared teacher will not be able to do the same for a small number of children (Rivkin, 2005: 424).

A review of the existing literature highlights the many factors that influence the academic performance of schoolchildren. These factors can be of an individual nature, as well as a wider range of reasons related to the quality of school education and regional inequality. Although there are studies devoted to educational inequality, this work focuses on rural youth, who are significantly inferior to their urban peers in a number of indicators. The study included the entire territory of the country, while most previous studies were limited to analysis within a single school or district. However, comprehensive studies in Kazakhstan that consider this issue on a large scale, taking into account socio-economic, infrastructural and cultural factors are still lacking.

## Materials and methods

The main research question was: "In which schools are students more likely to demonstrate high academic achievement?".

In addition, 4 sub-questions were put forward:

1. What are the structural characteristics of the schools that took part in the Olympiad?
2. Is there a relationship between the individual characteristics of the student and his final scores in spatial and logical thinking in the first stage of the Olympiad?
3. Is there a relationship between the individual characteristics of the student and his final score in English, native language, mathematics and natural science in the second stage of the Olympiad?
4. Is there a relationship between the structural characteristics of the school and the average scores that its students scored in the second stage of the Olympiad?

We put forward the following hypotheses:

1. The higher the level of infrastructure provision of a school, the higher the probability of the school's participation in the Olympiad;
2. The student's success is associated with his individual factors, such as language, gender and place of residence;
3. The higher the school's infrastructure provision, the higher the students' results in the Olympiad;

4. The smaller the number of students per teacher in a school, the higher the average scores of students in the Olympiad.

The study used data from the results of the National Intellectual Olympiad for Rural Schools "Myn Bala" in 2023, including both stages of the Olympiad, and data on schools from the "National Educational Database" of the Republic of Kazakhstan (NEDB). The Olympiad is held annually in the format of online testing and includes two stages. The first stage includes tests to determine the level of cognitive abilities. After passing to the second stage, the academic knowledge of students is assessed. Here, students are tested in mathematical literacy, languages (Kazakh / Russian, English) and natural science. All students of grades 6 of rural schools across the country are eligible to participate in the Olympiad. At the first stage, 5,000 finalists are selected. At the second stage, 1,000 winners are selected, who receive the opportunity to study free of charge in the best specialized schools in the country.

In addition to the scores received by the participants, the Olympiad data also contained information on the language of the test and the language of instruction of the school from which the student submitted the application. The Olympiad data were divided according to the number of stages.

The school data were obtained from the source – NEDB – and contain a large number of various school characteristics, including the size of the school classes, their infrastructure, accessibility, etc. After pre-processing, the following characteristics were included in the analysis: the total number of students in the school, the total number of teachers in the school, the distribution of students by gender, the number of books in the school, the availability of computers in schools, the number of computers with Internet access, Internet speed, the number of infrastructure facilities at the school and the years of their construction, the distance of the school to the district center.

The Olympiad data were linked to school data from the NEBD. A total of 4,007 schools across the country participated in the 2023 Olympiad. Based on the results of the first and second stages, 3,965 and 1,235 schools were linked to the NEBD database, respectively. In total, the study used data from 51,955 Olympiad participants.

The study used the regression analysis method. The dependent variable for testing the first hypothesis is the fact of school participation ("1" – participated, "0" – did not participate). To test the second hypothesis, the dependent variables were the number of points scored by the student (in tens), and for the third and fourth, the average number of points scored by the school's students (in tens). Accordingly, for the first question, logistic regression was used, which estimates the chance of an event occurring (in this case, the school's participation in the Olympiad), while in the other cases, linear regression was used to linearly estimate the relationships between one numerical variable and many other variables.

Independent variables were of two types. For questions 2 and 3, individual characteristics of the child were used:

- Language of writing of the Olympiad;
- Gender of the child;
- Distance category from the district center (up to or more than 100 km.);
- Control over the region of location.

For questions 1 and 4, the following school indicators were used:

- Number of students per teacher (in tens);
- Distance category from the district center (up to or more than 100 km.);
- Number of books in the school library (in thousands);
- Number of computers in the school (in tens);
- Number of shifts in the school;
- Internet speed in the school (in tens of kb);
- First language of instruction in the school;
- Controls on the region of location and the number of children in the school who wrote the Olympiad.

For each model, robust errors were presented, clustered by region of location of students or schools.

**Results and discussion**

The results of the descriptive analysis show the expected trend that the average score increases with the time spent writing the Olympiad (Figure 1). The red line is shown for those students who wrote the Olympiad in Kazakh, the blue line is for those who wrote it in Russian. This may indicate that students who spend more time thinking about questions and tasks demonstrate better results. However, when looking at the regional variation, it is noticeable that in some regions (North Kazakhstan Region and East Kazakhstan Region) there is a drop in scores at certain time intervals when taking the test in Russian, which may indicate factors such as student fatigue or difficulty of tasks at the end of the test. In both cases, this trend can be seen after 60 minutes of the

test. However, the supposed reason may also be the speed of the Internet, because when looking at the data, you can notice a small difference of a couple of points in such an indicator as Internet "lag" and at the same time a large load. In addition, it is worth noting that it was in these regions that a greater number of schoolchildren took the test in Russian, that is, their number and simultaneous delivery of the students could interfere with each other and alternately slow down the Internet speed, increasing the load.

Also, in the Zhambyl and Karaganda regions, students who took the test in Kazakh, on average, show higher results compared to those who took the test in Russian. However, this trend is noticeable only when a certain threshold of points is reached, after which the differences in results between the languages of the test become less pronounced. We believe that the key reason for this trend is the higher quality of education in the Kazakh language in local rural schools. However, we cannot say this with complete certainty due to the fact that this change does not go through the entire graph, and, apparently, this explanation can only work for a limited number of regions and schools. Another possible reason for the differences in students' basic knowledge that have been discussed is that the students who took the test in 2023 were studying remotely at the time of their primary school education due to the outbreak of the coronavirus (COVID-19) pandemic. This suggests that schools had different levels of digitalization at the time of the outbreak of the pandemic. Of the regions shown in the graph, East Kazakhstan Region appears to be the one that stands out from the general trend. Its graph shows an unusual pattern, where the average scores for students who took the test in Kazakh initially increase over time, but then decline before rising again.

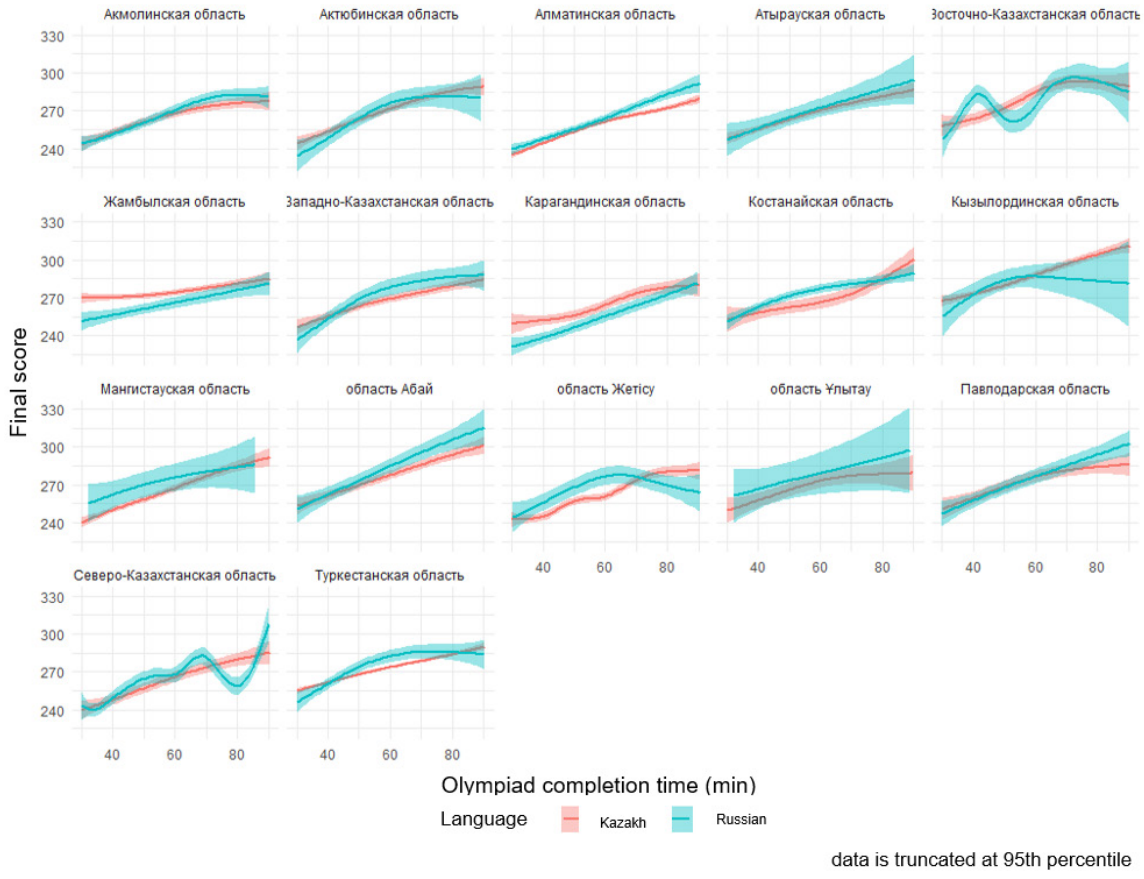


**Figure 1 –** The relationship between the time of passing the 1st stage of the Olympiad "Myn Bala" and the final score by region

Let's move on to the difference in scores between boys and girls (Figure 2). It shows the relationship between the students' final scores for spatial and logical tasks. According to the graph, both groups show a positive correlation between the scores for logical and spatial tasks – that is, students who did well in logical tasks, as a rule, also showed good results in spatial tasks. The graph shows that the line representing female results (red) in the middle of the graph is slightly higher than the blue line representing male results. This indicates that schoolgirls who scored from 110 to 150 points for "logical thinking" were on average slightly higher than schoolboys who, in turn, wrote slightly better in "spatial thinking" in the same range.

However, as the total score for logical tasks increases, this difference decreases and the lines converge. These are interesting results, since research shows that at an early age, girls can show higher abilities in spatial and verbal tasks. For example, Halpern's work shows that girls are, on average, superior on many memory tasks, including object and location memory, episodic memory, literacy, oral language, and writing (Halpern, 2006). As scores on the reasoning tasks increase, as the level of difficulty of the problems increases, the gender differences disappear, which may indicate that the existing difference in the middle of the graph is not the result of stable cognitive differences, but rather reflects initial testing conditions or teaching methods.

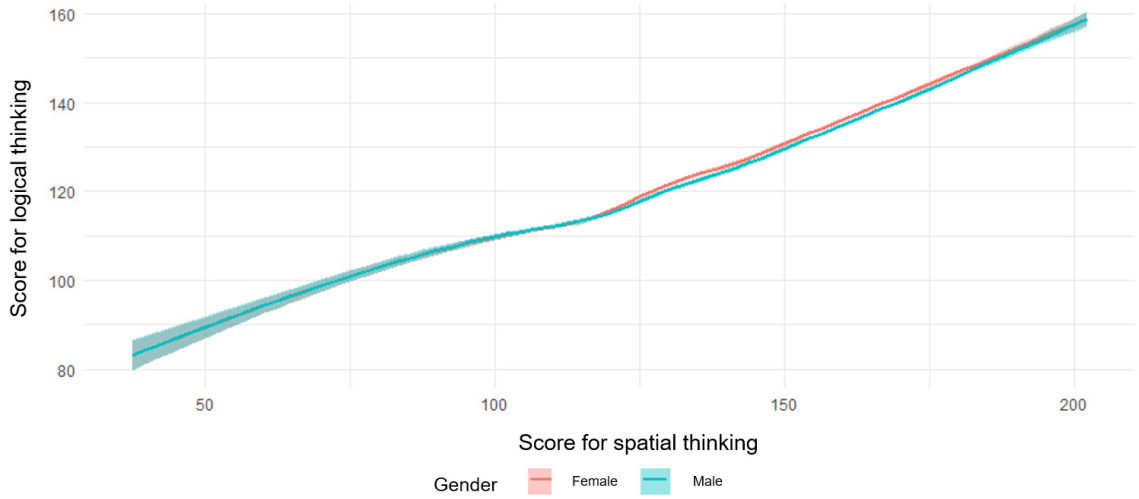


**Figure 2** – The relationship between the scores for the spatial and logical thinking blocks of the 1st stage of the "Myn Bala" Olympiad for boys and girls

In addition to language and gender differences, geographic variation is also expected (Figure 3). Three types of geographic location of students are distinguished: in the district center (red), within 100 km from the district center (green), and those who wrote the Olympiad more than 100 km from the district center (blue). Results are given for each of the regions. Students from district centers in many regions tend to score higher than students from more remote areas. This may be due to better access to educational resources and higher quality education in central areas. Students living close to district centers (up to 100 km from the district center) also show results comparable to those of students from district centers, especially at longer writing times. This may reflect the spread of educational opportunities beyond the center, but with some dependence on proximity to the center. While students living further from district centers (more than 100 km from the district center) show lower results on average, which may indicate the importance of the role of distance and the degree of remoteness from the district center in obtaining education.

Some regions, such as East Kazakhstan Region, show an unusual trend where students living far from the district center show average scores comparable to or even higher than students from nearby and central regions. This may indicate the presence of effective educational programs in remote areas or special motivation of students in these regions.

North Kazakhstan Region also shows an unusual trend where students living far from the dis-

trict center sometimes show results better than their peers from the district center or living closer to it. In Turkestan Region, there is a significant spread of results among students living more than 100 km from the district center, especially at the upper end of the time range. These deviations from the general trend may be due to the fact that the region has an individual feature of the educational system, innovative technologies are being introduced in remote schools.

Following the outlined sub-questions, the properties of the schools that decided to participate in the Olympiad were identified. Let us turn to Table 1.

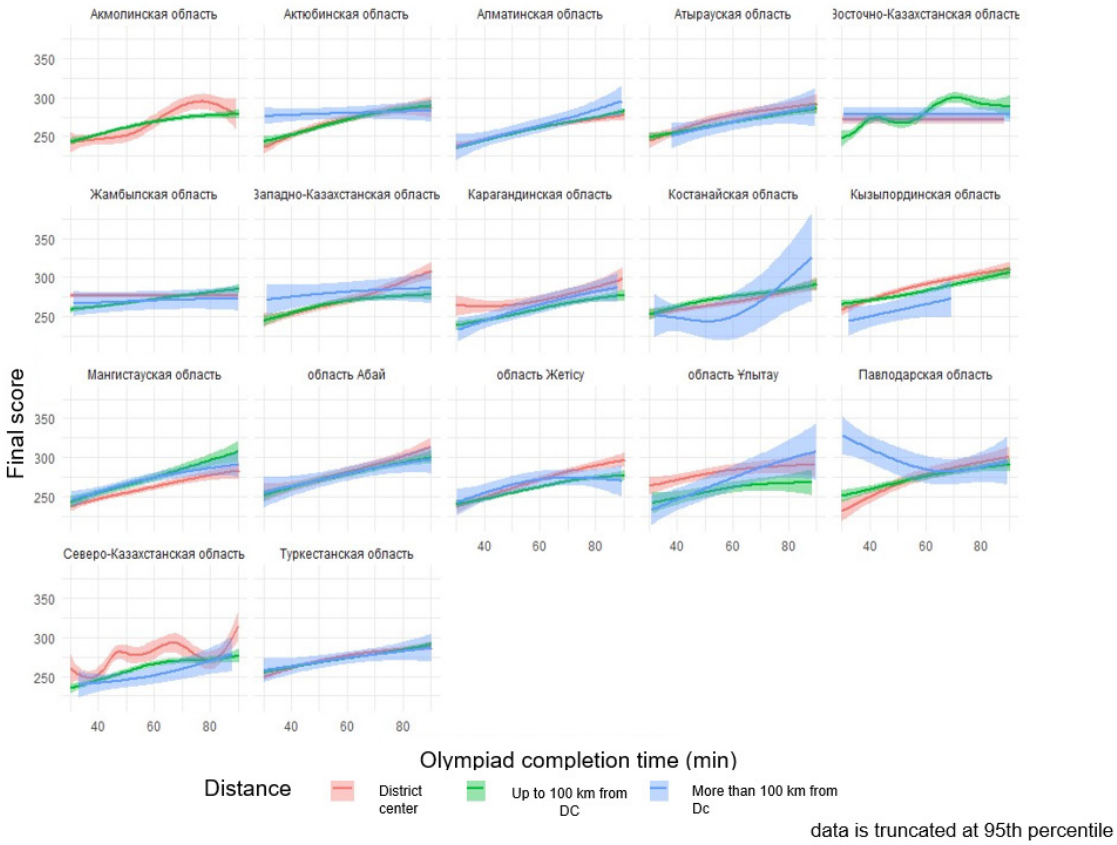


**Figure 3** – Relationship between the time of completing the 1st stage of the Olympiad "Myn Bala" and the final score for schoolchildren at different distances from the district center

**Table 1 –** Logistic regression of the chance of a school participating in the Olympiad

| School characteristics | School participation in the Olympiad (odds ratio) |
|---|---|
| Number of students per teacher (in tens) | 1.182*** (0.006) |
| Distance of school from district center (5-100 km) | 1.319*** (0.033) |
| Distance of school from district center (more than 100 km) | 0.722*** (0.099) |
| Number of books in school (in thousands) | 1.000 (0.001) |
| Number of computers in school (in tens) | 1.018*** (0.005) |

*Continuation of the table*

| School characteristics | School participation in the Olympiad (odds ratio) |
|---|---|
| 2 school shifts in school | 1.335*** (0.020) |
| 3 school shifts in school | 1.053 (0.078) |
| Internet speed in school (in tens) | 0.878*** (0.010) |
| First language in school (Kazakh) | 1.245*** (0.032) |
| Constant | 0.000 (0.000) |
| Region | Yes |
| Observations | 4597 |
| *Note:* *p<0.1; **p<0.05; ***p<0.01<br>Robust standard errors clustered by location region are given in parentheses. | |

An increase in the number of students per teacher in a school is associated with an increase in the chance of school participation. It can be assumed that large schools demonstrate greater activity, while small schools operate in conditions of deprivation and limited resources. With an increase in the number of students, teachers and parents of schoolchildren have more incentives for their child to participate in the "competition", since they will not be able to receive a quality education with the same heavy workload on teachers in the future.

In general, distance from the district center is one of the most important predictors in this model. Compared to district centers, the chance that schools located no more than 100 km away from them will participate in the Olympiad increases by 32%. At the same time, the chance that schools further than this radius will participate, on the contrary, is 28% lower. This may be due to the fact that teachers in such schools are less aware of the Olympiad and the organizers need to pay more attention to popularizing the event among such educational institutions.

A similar pattern to distance is observed for the number of shifts. If 2 shifts at school are one of the most important predictors of participation, then three shifts are not a significant category. We tend to interpret this in the same logic – with an average quality of education, teachers and parents may have incentives for a child to win such an important competition as "Myn bala". However, with a large workload on teachers, we expect less involvement of school administration and teachers in additional initiatives.

Infrastructure predictors did not show a statistically significant effect on participation. Thus, the size of the school library turned out to be insignificant, and the addition of 10 computers in the school determines about a 2% increase in the chances of participation.

It is worth noting that schools with Kazakh as the first language of instruction are also more likely to participate in the Olympiad, which may also determine the success of children from Kazakh-language schools (see below).

If school representatives decide to participate in "Myn Bala", schoolchildren who have signed up for the Olympiad first go through the first online stage – they write tests on logical and spatial thinking. Table 2 shows the relationship between the results of these tests and the individual characteristics of the child.

We have to admit that the model, according to R-square, doesn't provide enough information to figure out how a student's individual characteristics relate to their final score. Thus, the strongest predictor is student's gender, where boys are statistically significantly better at spatial thinking problems, but worse at logic. However, in both cases the difference is only one point (out of 200).

The fact of taking the Olympiad in the Kazakh language, which is associated with a 2-point increase in logical thinking tasks, also has an insignificant effect. The distance from the regional center up to 100 km has the same degree of significance, which, all other things being equal, reduces the final score by 1.

**Table 2** – Linear regression of students' success in the 1st stage of the Olympiad "Myn bala"

| | **The student's final score for:** | |
|---|---|---|
| | Logical thinking (in tens) | Spatial thinking (in tens) |
| Distance of school from district center (5-100 km) | -0.077 | -0.131* |
| | (-0.077) | (-0.131) |
| Distance of school from district center (more than 100 km) | 0.146 | 0.156 |
| | (0.146) | (0.156) |
| Participant gender (male) | -0.058*** | 0.072*** |
| | (-0.058) | (0.072) |
| Language in which participant took the test (Kazakh) | 0.116* | -0.075 |
| | (0.116) | (-0.075) |
| Constant | 12.521*** | 13.863*** |
| | (12.521) | (13.863) |
| Region | Yes | Yes |
| Observations | 47,262 | 47,262 |
| $R^2$ | 0.043 | 0.024 |
| Adjusted $R^2$ | 0.043 | 0.024 |
| Residual Std. Error (df = 47241) | 2.054 | 2.509 |
| F Statistic (df = 20; 47241) | 106.5*** | 59.13*** |
| *Note:* *p<0.1; **p<0.05; ***p<0.01<br>Robust standard errors clustered by location region are given in parentheses. | | |

Perhaps we will find a more pronounced effect of individual characteristics among the best schoolchildren of the first stage who got to the final and wrote tests in their native (Kazakh or Russian) and English, as well as mathematical literacy and natural science. The answer to this is given in Table 3.

**Table 3 –** Linear regression of students' success in the 2nd stage of the Olympiad "Myn bala"

| | **The student's final score (in tens) for:** | | | |
|---|---|---|---|---|
| | Natural science | English language | Mathematical literacy | Native language |
| Distance of school from district center (5-100 km) | -0.029 | -0.360*** | -0.466*** | -0.206*** |
| | (-0.029) | (-0.360) | (-0.466) | (-0.206) |
| Distance of school from district center (more than 100 km) | 0.124 | -0.132 | -0.135 | -0.105 |
| | (0.124) | (-0.132) | (-0.135) | (-0.105) |
| Participant gender (male) | 0.259*** | -0.256*** | 0.672*** | -0.481*** |
| | (0.259) | (-0.256) | (0.672) | (-0.481) |
| Language in which participant took the test (Kazakh) | -0.629*** | -0.265*** | -0.634*** | 0.816*** |
| | (-0.629) | (-0.265) | (-0.634) | (0.816) |
| Constant | 10.223*** | 9.487*** | 17.798*** | 10.245*** |
| | (10.223) | (9.487) | (17.798) | (10.245) |
| Region | Yes | Yes | Yes | Yes |
| Observations | 4,022 | 4,022 | 4,022 | 4,022 |

*Continuation of the table*

| | **The student's final score (in tens) for:** | | | |
|---|---|---|---|---|
| | Natural science | English language | Mathematical literacy | Native language |
| $R^2$ | 0.074 | 0.053 | 0.038 | 0.057 |
| Adjusted $R^2$ | 0.070 | 0.048 | 0.033 | 0.052 |
| Residual Std. Error (df = 4001) | 1.330 | 1.464 | 3.216 | 1.871 |
| F Statistic (df = 20; 4001) | 16.047*** | 11.097*** | 7.926*** | 12.072*** |
| *Note:* *p<0.1; **p<0.05; ***p<0.01<br>Robust standard errors clustered by location region are given in parentheses. | | | | |

Thus, compared to girls, boys are statistically significantly worse at writing English and their native language (by 3 and 5 points, respectively), but they do better at answering tests in natural science (by 3 points) and mathematics (by 7 points). This distance is all the more interesting because we are analyzing schoolchildren who have already demonstrated outstanding abilities. Among other things, this raises questions about the need to pay more attention to the gender aspect in the quality of education for students in primary and secondary schools. We observe the same curious dynamics for the language of the Olympiad. Thus, Kazakh-speaking students write tests in mathematics, natural science, and English worse than Russian-speaking students (by 6, 6, and 3 points, respectively), but they write the section on knowledge of their native language better by 8 (!) points.

In addition, it is worth noting that schoolchildren living in the district center write all the Olympiad blocks except for natural science better (by 2-5 points) than students at a distance of up to 100 km from the district center.

We still have to answer the last sub-question about the relationship between the structural characteristics of schools and the average scores that their students get in the second stage of the Olympiad. Table 4 shows the average scores of students by school and the indicators of the educational institutions themselves.

An important finding is that the predictors of school infrastructure are consistently insignificant (with the exception of a borderline positive relationship between Internet speed and the score in natural science). However, we can interpret this in the context of the fact that infrastructure is less important for quality education than the motivation and training of teachers.

In this sense, it is curious that the number of students per teacher, as in the first model, has a positive effect. In this case, we are also inclined to argue in favor of the non-linearity of the effect, but, in addition, a possible explanation is the non-obvious quality of the data in the National Educational Database, since, as far as we know, the database is filled with information from representatives of the schools themselves and, therefore, may contain distortions. Perhaps, with alternative and independent estimates, we would have obtained a different effect in this case.

We also find an effect of language, but not at the level of individual choice of students, but as the first language of instruction at school. If this is Kazakh, then students at the school, all other things being equal, will score lower than in Russian-language schools in all subjects except their native language.

It is also worth noting that the distance from the district center up to 100 km. is statistically associated only with average scores in English.

**Table 4 –** Linear regression of school success in the 2nd stage of the Olympiad "Myn bala"

| | **Average score (in tens) for:** | | | |
|---|---|---|---|---|
| | English language | Native language | Mathematical literacy | Natural science |
| | (1) | (2) | (3) | (4) |
| Number of students per teacher (in tens) | 0.788*** | 0.716*** | 1.382*** | 0.351** |
| | (0.788) | (0.716) | (1.382) | (0.351) |

*Continuation of the table*

| | Average score (in tens) for: | | | |
|---|---|---|---|---|
| | English language | Native language | Mathematical literacy | Natural science |
| Distance of school from district center (5-100 km) | -2.042** | -0.796 | -2.058 | -0.749 |
| | (-2.042) | (-0.796) | (-2.058) | (-0.749) |
| Distance of school from district center (more than 100 km) | -1.426 | 1.593 | 0.290 | 1.749 |
| | (-1.426) | (1.593) | (0.290) | (1.749) |
| Number of books in school (in thousands) | 0.028 | -0.001 | 0.096 | 0.017 |
| | (0.028) | (-0.001) | (0.096) | (0.017) |
| Number of computers in school | 0.154 | 0.212 | 0.408 | 0.019 |
| | (0.154) | (0.212) | (0.408) | (0.019) |
| (in tens) | 1.778* | 0.065 | -1.101 | -1.154 |
| | (1.778) | (0.065) | (-1.101) | (-1.154) |
| 2 school shifts in school | 5.713** | 3.699 | 4.774 | -0.700 |
| | (5.713) | (3.699) | (4.774) | (-0.700) |
| 3 school shifts in school | 0.050 | 0.205 | 0.306 | 0.251* |
| | (0.050) | (0.205) | (0.306) | (0.251) |
| Internet speed in school (in tens) | -1.095 | 4.996*** | -5.157*** | -3.218*** |
| | (-1.095) | (4.996) | (-5.157) | (-3.218) |
| First language in school (Kazakh) | 87.048*** | 101.464*** | 178.903*** | 104.355*** |
| | (87.048) | (101.464) | (178.903) | (104.355) |
| Constant | Yes | Yes | Yes | Yes |
| Number of participants | 1,221 | 1,221 | 1,221 | 1,221 |
| $R^2$ | 0.128 | 0.094 | 0.101 | 0.106 |
| Adjusted $R^2$ | 0.109 | 0.075 | 0.082 | 0.087 |
| Residual Std. Error (df = 1194) | 12.123 | 15.203 | 26.814 | 10.704 |
| F Statistic (df = 26; 1194) | 6.722*** | 4.792*** | 5.167*** | 5.451*** |

*Note:* *p<0.1; **p<0.05; ***p<0.01
Robust standard errors clustered by location region are given in parentheses.

Thus, we can talk about the presence of significant effects, which, however, are not clearly manifested. At the individual level, we observe a difference only for outstanding schoolchildren who were able to pass to the second stage of the Olympiad (it is worth highlighting the effects of gender and language). The absence of differences in the first stage of the Olympiad, during which cognitive abilities were assessed, is logical, since this stage tested knowledge that does not depend on the learning conditions. At the same time, the characteristics of schools also show ambiguous results. With the exception of the advantage of regional centers, the effects found are either inconsistent or counterintuitive. This forces us to appeal to the need for a more sophisticated analysis in the future, or to clarify the data used for the analysis.

## Conclusion

The aim of this study was to identify individual and structural factors that explain differences in students' performance in rural schools.

On average across the country, students living in district centers score higher than their peers in rural areas. This effect is particularly noticeable when compared to students who live at a relatively short distance from the district center (between 5 and 100 km).

The involvement of students in the Olympiad and their educational achievements are geographically heterogeneous: in some regions, students from remote areas show the same of higher results compared to students from district centers.

School infrastructure (libraries, computers and internet speed) is not significantly related to stu-

dents' educational achievements. It can be assumed that the level of teachers' qualifications and their motivation are more important than the quality of infrastructure.

Children from medium-sized schools (two shifts) statistically perform better in academic competitions. At the same time, excessive workload (three shifts) is associated with lower scores. The number of students per teacher is also a significant factor. We tend to interpret this in terms of the incentives available to teachers and parents. Perhaps children, who study in single-shift schools, receive a relatively good education and are more likely to have better career prospects in the future. With average quality education, teachers and parents may have incentives for their children to win such an important competition as "Myn Bala". However, with a heavy workload on teachers (three-shift schools), we expect less involvement of school administrators and teachers in additional initiatives.

The results of the Olympiad show a gender gap, primarily among students who advanced to the final stage and took subject tests, i.e., those who showed the most outstanding results. Girls perform better on tasks in their native language and English, while boys perform better on tests in science and mathematics. In this regard, we believe that more attention should be paid to the gender aspect in provision of education for primary and secondary school students.

Kazakh-speaking students participate in Olympiad more often and perform better on native language tests than Russian-speaking students, but the latter have higher results in mathematics, English and natural sciences.

Individual characteristics of students (gender, language, place of residence) have less effect on test results for general abilities than on tasks in specific subjects.

The hypotheses put forward regarding the relationship between a student's success and individual factors such as language, gender and place of residence, as well as structural features of the schools, were partially confirmed. At the same time, the degree of their expression varied depending on the characteristics under consideration and regional features. It's important to note that this work is positioned as the first attempt to analyze the problem using local data. The results emphasize the need for further study of individual and structural factors, taking into account the specifics of the educational environment and social context.

*The article was prepared with grant funding from the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (BR28713139 «Computational Social Sciences: New Horizons for Methodology and Evidence-Based Policy»).*

**Reference**

Budiongan K., Reyes G., Liporada K., Galo M., Bangeles D. (2024) Socioeconomic Status, Parental Expectations, and Student Motivation as Predictors of Academic Performance of Grade 12 Senior High School Students. *European Journal of Theoretical and Applied Sciences*, vol.2(4), pp.385-396. http://dx.doi.org/10.59324/ejtas.2024.2(4).32

Coleman J. (1968) The concept of equality of educational opportunity. *Harvard educational review*, vol.38(1), pp.7-22. https://doi.org/10.17763/haer.38.1.m3770776577415m2

Duckworth A.L., Seligman M.E. (2005) Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological science*, vol.16(12), pp.939-944. http://dx.doi.org/10.1111/j.1467-9280.2005.01641.x

Fini R. (2007) School achievement in Italy: An empirical view. *In International studies in educational inequality, theory and policy.* Dordrecht: Springer Netherlands, pp. 490-508.

Fuller B., Clarke P. (1994) Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. *Review of educational research*, vol.64(1), pp.119-157. https://doi.org/10.3102/00346543064001119

Halpern D., Institute of medicine; national academy of engineering; national academy of sciences; committee on science, engineering, and public policy; committee on maximizing the potential of women in academic science and engineering. (2006, November). Biopsychosocial contributions to cognitive performance. *In Biological, Social, and Organizational Components of Success for Women in Academic Science and Engineering: Report of a Workshop*. Washington: National Academies Press, pp. 20-20.

Hanushek E.A., Woessmann L. (2011) The economics of international differences in educational achievement. *Handbook of the Economics of Education*, vol.3, pp.89-200. https://doi.org/10.1016/B978-0-444-53429-3.00002-8

Hyde J.S., Linn M. C. (2006) Gender similarities in mathematics and science. Science, vol.314(5799), pp.599-600. https://doi.org/10.1126/science.1132154

IEA TIMSS & PIRLS. URL: https://timssandpirls.bc.edu

McPherson A. (1993) Measuring added value in schools. *Education Economics*, vol.1(1), pp. 43-51. https://doi.org/10.1080/09645299300000006

National Educational Database of the Republic of Kazakhstan (2023). URL: https://seda.iac.kz/map/

National Olympiad «Myń Bala» (2023). URL: https://1000bala.elumiti.kz/

PIRLS 2021 International Results in Reading. (2021). URL: https://pirls2021.org/results/

PISA 2018 results. (2018). URL: https://www.oecd.org/en/about/programmes/pisa/pisa-publications.html

Porta E., Arcia G., Macdonald K., Radyakin S., Lokshin M. (2011) Analyzing education inequality with ADePT Edu (Report). World Bank. https://openknowledge.worldbank.org/handle/10986/2346

Rivkin S.G., Hanushek E.A., Kain J.F. (2005) Teachers, schools, and academic achievement. *Econometrica,* vol.73(2), pp.417-458. http://dx.doi.org/10.1111/j.1468-0262.2005.00584.x

Schütz G., Ursprung H.W., Wößmann L. (2008) Education policy and equality of opportunity. *Kyklos*, vol.61(2), pp.279-308. https://doi.org/10.1111/j.1467-6435.2008.00402.x

Somers M.A., McEwan P. J., Willms J.D. (2004) How effective are private schools in Latin America? *Comparative education review*, vol.48(1), pp.48-69. http://dx.doi.org/10.1086/379841

Wikström C. (2005) Criterion-referenced measurement for educational evaluation and selection: doctoral dissertation, Umeå University, Department of Educational Measurement.

**Литература**

Budiongan K., Reyes G., Liporada K., Galo M., Bangeles D. Socioeconomic Status, Parental Expectations, and Student Motivation as Predictors of Academic Performance of Grade 12 Senior High School Students // *European Journal of Theoretical and Applied Sciences*. – 2024. – Vol.2. – № 4. – P. 385-396. http://dx.doi.org/10.59324/ejtas.2024.2(4).32

Coleman J. The concept of equality of educational opportunity // *Harvard educational review*. – 1968. – Vol. 38. – № 1. – P. 7-22. https://doi.org/10.17763/haer.38.1.m3770776577415m2

Duckworth A.L., Seligman M.E. Self-discipline outdoes IQ in predicting academic performance of adolescents // *Psychological science*. – 2005. – Vol. 16. – № 12. – P. 939-944.doi: http://dx.doi.org/10.1111/j.1467-9280.2005.01641.x

Fini R. School achievement in Italy: An empirical view //*International studies in educational inequality, theory and policy*. – Dordrecht : Springer Netherlands, 2007. – P. 490-508.

Fuller B., Clarke P. Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy //*Review of educational research*. – 1994. – Vol. 64. – № 1. – P. 119-157. https://doi.org/10.3102/00346543064001119

Halpern D., Institute of Medicine; National academy of engineering; National academy of sciences; Committee on science, engineering, and public policy; Committee on maximizing the potential of women in academic science and engineering. Biopsychosocial contributions to cognitive performance // *Biological, Social, and Organizational Components of Success for Women in Academic Science and Engineering: Report of a Workshop*. – Washington: National Academies Press, 2006. – P. 20-20.

Hanushek E.A., Woessmann L. The economics of international differences in educational achievement // *Handbook of the Economics of Education*. – 2011. – Vol. 3. – P. 89-200. https://doi.org/10.1016/B978-0-444-53429-3.00002-8

Hyde J. S., Linn M. C. Gender similarities in mathematics and science // *Science*. – 2006. – Vol. 314. – № 5799. – P. 599-600. https://doi.org/10.1126/science.1132154

IEA. TIMSS & PIRLS [Electronic resource]. Available at: https://timssandpirls.bc.edu (accessed 10.07.2025).

McPherson A. Measuring added value in schools // *Education Economics*. – 1993. – Vol. 1. – № 1. – P. 43-51. https://doi.org/10.1080/09645299300000006

National Educational Database of the Republic of Kazakhstan [Electronic resource]. Available at: https://seda.iac.kz/map/ (accessed: 10.07.2025).

National Olympiad «Myń Bala» [Electronic resource]. Available at: https://1000bala.elumiti.kz/ (accessed: 10.07.2025).

PIRLS 2021 International Results in Reading [Electronic resource]. Available at: https://pirls2021.org/results/ (accessed: 12.07.2025).

PISA 2018 Results [Electronic resource]. Available at: https://www.oecd.org/en/about/programmes/pisa/pisa-publications.html (accessed: 10.07.2025).

Porta E., Arcia G., Macdonald K., Radyakin S., Lokshin M. Analyzing education inequality with ADePT Edu. Report. – Washington, DC: World Bank, 2011. – 118 p.

Rivkin S.G., Hanushek E.A., Kain J. F. Teachers, schools, and academic achievement //*Econometrica.* – 2005. – Vol. 73. – №2. – P. 417-458. http://dx.doi.org/10.1111/j.1468-0262.2005.00584.x

Schütz G., Ursprung H. W., Wößmann L. Education policy and equality of opportunity //*Kyklos.* – 2008. – Vol. 61. – № 2. – P. 279-308. doi: https://doi.org/10.1111/j.1467-6435.2008.00402.x

Somers M.A., McEwan P.J., Willms J.D. How Effective Are Private Schools in Latin America? Occasional Paper. – 2001. http://dx.doi.org/10.1086/379841

Wikström C. Criterion-referenced measurement for educational evaluation and selection: Doctoral dissertation. – Department of Educational Measurement Umeå universitet, 2005.

***Information about authors:***

*Aubakirova Zhaniya (corresponding author) – Kazakhstan Sociology Lab researcher, Corporate Fund "El Umiti Fund" (Astana, Kazakhstan, e-mail: zhaniya.abkr@gmail.com*)

*Amrin Aslanbek – Director of the Corporate Fund "El Umiti Fund", Astana, Kazakhstan, e-mail: aslanbekamrin@gmail.com;*

*Kaztayeva Sulushash – PhD, Kazakhstan Sociology Lab researcher, Corporate Fund "El Umiti Fund" (Kazakhstan, Astana, e-mail: ss.kaztayeva@gmail.com*)

*Tenisheva Ksenia – Associate Professor, Director of Sociology Programs, School of Computational Social Sciences, Candidate of Sociological Sciences, European University (St. Petersburg, Russia, e-mail: ktenisheva@eu.spb.ru)*

***Авторлар туралы мәлімет:***

*Аубакирова Жания Кайыржановна (корреспондент-автор) – Kazakhstan Sociology Lab зерттеуші, «El Umiti» корпоративтік қоры (Астана, Қазақстан, e-mail: zhaniya.abkr@gmail.com)*

*Амрин Асланбек Кеменгерович – «El Umiti» корпоративтік қорының директоры (Қазақстан, Астана, e-mail: aslanbekamrin@gmail.com*)

*Каztаева Сулушаш Сеитовна – PhD, Kazakhstan Sociology Lab зерттеуші, «El Umiti» корпоративтік қоры (Астана, Қазақстан, e-mail: ss.kaztayeva@gmail.com*)

*Тенишева Ксения Алексеевна – әлеуметтану ғылымдарының кандидаты, есептік әлеуметтік ғылымдар мектебінің әлеуметтану бағдарламалары бөлімінің доценті, директоры, Еуропа университеті (Санкт-Петербург, Ресей, e-mail: ktenisheva@eu.spb.ru)*

***Сведения об авторах:***

*Аубакирова Жания Кайыржановна (автор-корреспондент) – научный сотрудник Kazakhstan Sociology Lab Корпоративного фонда «Фонд "El Umiti"» (Астана, Казахстан, e-mail: zhaniya.abkr@gmail.com)*

*Амрин Асланбек Кеменгерович – директор Корпоративного фонда «Фонд "El Umiti"» (Астана, Казахстан, e-mail: aslanbekamrin@gmail.com*)

*Каztаева Сулушаш Сеитовна – PhD, научный сотрудник Kazakhstan Sociology Lab Корпоративного фонда «Фонд "El Umiti"» (Астана, Казахстан, e-mail: ss.kaztayeva@gmail.com)*

*Тенишева Ксения Алексеевна – кандидат социологических наук, доцент, директор программ по направлению социология, Школа вычислительных социальных наук, Европейский университет (Санкт-Петербург, Россия, e-mail: ktenisheva@eu.spb.ru)*

*Келіп түсті: 14 наурыз 2025 жыл*
*Қабылданды: 21 тамыз 2025 жыл*