

APPLICATION OF INFORMATION TECHNOLOGIES AND METHODS FOR PROCESSING BIG DATA TO THE MANAGEMENT OF THE EDUCATIONAL PROCESS DURING THE PANDEMIC

MUKHIYADIN A.¹, MAKHAZHANOVA U.¹, SERIKBAYEVA S.¹, KASSEKEYEVA A.¹,
MURATOVA G.², KARAUYL BAYEV S.², MURATKHAN R.³, KENZHEBAY A.¹

¹L. N. Gumilyov Eurasian National University, Kazakhstan, Astana

²Taraz Regional University named after M.Kh.Dulaty, Kazakhstan, Taraz

³Buketov Karagandy University, Kazakhstan, Karaganda

E-mail: amukhiyadin@gmail.com¹, umt87@mail.ru¹, inf_8585@mail.ru¹, aibike_7474@mail.ru¹,
gauchar70@mail.ru², saparkalymbet@mail.ru², raikhan.muratkhani@mail.ksu.kz³, salik2000@icloud.com¹

ABSTRACT

The spread of coronavirus infection has changed all spheres of human life, forcing a new look at traditional forms of activity. Education as the most important social sphere has taken the blow of the pandemic transformation, resulting in a significant spread of distance forms of interaction between the teacher and the learner, involving the reliance on digital learning tools. Due to the need for participants in the educational process in these conditions to operate with large amounts of data and process information electronically, the use of the potential of artificial intelligence in education. The article deals with new opportunities and problems of integration of education and artificial intelligence under conditions of Covid-19 pandemic. The object of the research is education in the epoch of Covid19. The article describes fundamental changes, which took place in the educational system during the pandemic period, outlines promising directions of artificial intelligence usage in the modern educational process and corresponding problems. The necessity of further research of possibilities of integration of education and big data under conditions of acceleration of processes of digitalization of society is marked.

COVID-19 epidemic caused a great number of human casualties and destruction of economic, social, public and health systems all over the world. Combating such an epidemic requires an understanding of its characteristics and behavior, which can be determined by collecting and analyzing relevant big data. Big data analysis tools play a vital role in generating the knowledge necessary for decision-making and precautionary measures. However, because of the vast amount of COVID-19 data available from various sources, there is a need to review the role of big data analysis in controlling the spread of COVID-19 by presenting the main challenges and directions of COVID-19 data. analysis, as well as providing a framework for relevant existing applications and research to facilitate future research on COVID-19 analysis. Therefore, in this paper, we conduct a literature review to highlight the contributions of several studies in the field of COVID-19-based big data analysis.

Keywords: *COVID-19, Big Data, Relationship Between Traits, Social Data, Regression Analysis, Multinomial Logistic Regression, Secondary School Students, Educational Activity*

1. INTRODUCTION

The entire world is currently engaged in an unprecedented fight against the global pandemic of the COVID-19 coronavirus. The world's accumulated human capital, which is widely acknowledged to be one of the major sources of wealth of nations, is under threat. The consequences

for the wider economy will be felt for many years to come. During this period, it is a well-designed education policy that must be the cornerstone of the response to the pandemic crisis. Education empowers governments, individuals and their families to multiply human capital, protect the poor and vulnerable, and drive economic recovery, while promoting sustainable economic growth and poverty reduction. The COVID-19 pandemic primarily

affects the health sector, but spillover effects can already be seen in the education systems of many countries, largely due to school closures. It is easier to catch up if short periods of time are spent out of school, while longer absences will have a negative impact on learning.

Today, in the context of a pandemic, the issue of the quality of education received by students and schoolchildren in a distance format, and, in particular, the control of students' knowledge, is becoming particularly relevant. It is obvious that the forms of control that were used in full-time training must inevitably be adapted or even changed taking into account the current learning conditions. In this regard, in this article the author attempts to consider two alternative control tools that can help a modern teacher to check the degree of assimilation of the studied material by students. It is also necessary to pay attention to the fact that these control tools should not only perform their direct function, but also increase the interest and motivation of students to the subject being studied, which is also important at a time when students are forced to study from home.

School closures due to the pandemic lasted (and continue) for more than two months in most countries worldwide. Globally, more than 70% of students were affected. Half of them had no access to a home computer and 43% had no home Internet connection. The focus in the initial period of school closures was on remote learning, including measures to bridge the digital divide.

The time has now come to reflect on effective learning in the context of pandemic closures. Specific studies on this are not yet available, but most experts acknowledge that learning has been less effective during this period than it has been during regular school closures. Children "incurred losses", significant both in the short term (passing the UNT, even if they are postponed/cancelled in the current year) and in the long term (educational trajectories after school, employment). Children from families with low socio-economic status (SES) are particularly likely to "lose out", even in the traditional model of education they perform lower than children with high SES [OECD, 2018], and in this situation they are in comparatively worse conditions for distance learning. Thus, the focus of the publication is to assess and predict the possible effects of the pandemic on children's education and countries' human capital. Therefore, special studies are needed - and in many countries they are already underway - to plan measures to offset losses and minimize risks.

The challenges of our time require a constant reassessment of the methods, forms and means of education. The advent of the digital age and the emerging information space have highlighted the need to systematize existing approaches, including in education. The constant increase in the quantity of data required the development of skills for effective processing and extraction of useful information from the boundless and haphazard field of information. According to the common understanding, the category "big data" is used to refer to structured and unstructured data of huge volumes and significant diversity, efficiently handled by horizontally scalable software tools that emerged in the late 2000s and are alternative to traditional database management systems and Business Intelligence solutions. A similar interpretation of big data, but formulated more succinctly, can be found in the Concept of Digital Analytical Platform for Statistical Data. Broadly speaking, big data is referred to as a socio-economic phenomenon associated with the emergence of technological capabilities to analyze vast amounts of data. It should be noted that big data technology is actively used in the legal environment as well, demonstrating the possibilities of use in different directions. When applied to education, big data is a technology. It can be used to do at least three important things: create methodologies adapted to a large number of students; personalize content; and tailor learning modes.

The digitalization of education was made possible by the fact that several general-purpose technologies, such as mobile devices, the Internet, biometric technologies, cloud computing, big data analytics platforms, and artificial intelligence, had been developed and implemented simultaneously before. The pandemic period, however, was marked by a real breakthrough: the trend for universities to make management decisions in the educational sphere based on big data has clearly emerged.

2. OVERVIEW OF THE WORK

The article [1] examines the main forms of dysfunctionality of education, expressed in such types of unavailability as technological, infrastructural, personnel, skill, mental, their characteristics and main forms of manifestation are presented. It is concluded that, on the one hand, the coronavirus crisis was a powerful factor in the destabilization of all economic and social systems and caused their dysfunctionality. On the other hand, in relation to the Institute of education, this manifested itself most vividly: accelerated technological modernization and the rapid transition to innovative forms and technologies of education,

followed by an increase in technological literacy of teachers and students, revision of methodological and didactic materials, a change in the paradigm of educational process management directly in educational institutions and in general national education systems.

The article [2] discusses the global expansion of high-quality online learning resources and networks, offering real opportunities for university students to study in the open and distance education sector. However, such informal training, declaring the possibility of obtaining knowledge to anyone at any time in any place, with all its attractiveness, is associated with fundamental drawbacks: a large dropout of students (up to 95%), the lack of training regulations and, as a consequence, the problems of acquiring the necessary set of competencies by students for a specialist.

[3] Changes in the educational process in the context of a pandemic are considered. Many educational institutions began to introduce elements of distance learning into their educational process, to create special courses for distance learning, and then for full-time. The teaching staff had to master new e-learning technologies for themselves. Students also had to face new forms of acquiring knowledge when switching to distance learning

The article considers [4] the problem of a number of challenges faced by the Russian education system in connection with the transition to mass distance learning in the first half of 2020. The article raises two key issues. The first is the problem of ensuring compliance with the technical requirements necessary for the successful functioning of the programs through which distance learning is implemented with the technical capabilities of computers and other devices used by students to access distance education. The second issue is the need to exclude the possibility of students cheating during the intermediate certification in a remote format and misleading the teacher with references to technical problems.

The article [5] deals with the problem of forced transition to distance learning in the context of the 2020 pandemic declared by WHO. Based on the experience of practical classes in the senior classes of the Moscow Economic Lyceum and practical and lecture courses of Grodno universities, preliminary results of training in this format are analyzed.

The article [6] discusses specific problems, risks and challenges associated with the digital transformation of higher education: digital inequality, closely related to digital literacy, the lack

of a unified method of communication between students and teachers, the lack of generally accepted platform solutions for conducting classes online, the actual increase in the burden on the teacher in the conditions of digitalization of education, which while the norm is not taken into account in the wage rationing system, professional burnout in a sufficiently large number of teachers, lack of motivated and self-discipline of students, increased bureaucratization of the higher education system.

3. THE POTENTIAL OF USING BIG DATA IN EDUCATION

To design effective strategies for educating students during and after the pandemic, teachers and policy-makers must be grounded in the specific context of the local context. They should be guided by a context-specific understanding of how the pandemic is affecting students, their families, teachers, school staff, the communities in which they live, and the ability of the education system to perform its core functions. So, the first step is to develop a holistic understanding of these changes. Since the pandemic has affected different groups of students and individual schools in different ways, identifying specific needs should be based on a localized impact assessment of children and families. Creating a locally relevant picture of these educational needs does not mean that local authorities should be left alone to find and implement solutions. On the contrary, central and regional governments have a responsibility to provide differentiated support to regions and schools to ensure equality of educational outcomes. However, they should do so in support of those needs and strategies that have been identified locally. Failure to take account of local specificities would result in an inadequate response to the real situation, while the lack of compensatory roles for central and regional governments would exacerbate the unequal impact of the pandemic on education [7].

Indeed, in view of the spread of the new coronavirus (COVID-19), the traditional view of face-to-face educational processes was rapidly becoming irrelevant. As a result of universities moving to distance learning, the ways in which students acquire general and vocational competences have been rethought.

In our opinion, all the stages of transition from "analogue" to digital level of comprehension of educational reality under the COVID-19 pandemic, which Russian universities overcame, were carried out under the influence of big data on the construction of educational process of the university.

Let us emphasize that the main value of big data lies in the result of information analysis, not in the information itself in electronic form. The sources of big data are any information of analytical interest in terms of processing and studying the information obtained by the relevant services (e.g. Google Analytics or Yandex.Metrika). Well-trained artificial intelligence with the help of big data can replace traditional face-to-face forms of education involving a human teacher in terms of analytical data processing and the formation of the final decision. It should be acknowledged that this work can be done exclusively by artificial intelligence, not displacing the cognitive system of the human teacher, but relieving the latter of unnecessary work - processing of unstructured information accumulated by mankind. It is important for the educational process at universities to accumulate and present information from the content of big data. In this matter, the use of artificial intelligence is inevitable, the work of which is set up by university teaching staff based on the needs of the time and the profession, as well as the traditions of scientific schools and universities. It should be noted that the use of big data potential to improve the efficiency of educational activities is also a subject of discussion by foreign researchers. For example, Ray Saptarshi ⁸ stressed the possibility of identifying the interest of an educational course and the visualisation benefits of interactive educational processes as a result of big data analysis. Christos Vaitsis, Vasilis Hervatis and Nabil Zary⁹ explore the possibility of manipulating data with the appropriate use of visual big data analytics. Vidal Alonso Secades and Olga Arranz¹⁰ focus on the potential of big data as a source of information in distance learning for students and the automation of the educational process. Giti Javidi, Lila Rajabion and Ehsan Sheybani¹¹ reveal the benefits of big data analysis for making good pedagogical decisions, identifying learner behavior, and applying effective learning models [8].

4. METHODOLOGY AND METHODS

4.1 Educational activity – as social behavior

Educational activity is defined as the result of the integration of various types and manifestations of activity in the field of educational activity (intellectual, cognitive, volitional, etc.). In the most general form, educational activity is considered as a qualitative and quantitative measure of the interaction of the subject of learning with the learning environment. [9]

Most often, educational activity is understood as a certain interaction of various manifestations of cognitive, volitional, intellectual, emotional activity that characterize the degree of involvement of a student in the learning process. At the same time, educational activity is often identified either with cognitive activity, with an attitude to educational activity, or with volitional activity. Educational activity is also usually understood as a qualitative and quantitative measure of a child's involvement in the learning process, manifested, on the one hand, in the features of educational motivation, and, on the other, in the features of the implementation and regulation of educational activities.

In modern research, the emergency transition to online education and distance learning technologies is usually presented from a critical point of view, a point of view, as scientists seek to find out exactly how educational activity has suffered for distance learners.

A Danish study shows that at the first stage of quarantine, the reading activity of children of parents without higher education decreased compared to what it was before quarantine, and the transition to distance learning in primary schools in Malaysia has become a problem for both teachers and students and their parents.

To study the level of educational activity of school students during the pandemic, the factors influencing it were selected. Empirical methods of observation, survey of respondents (schoolchildren, parents, teachers) and analysis of the results were applied. The survey was conducted in the form of an online questionnaire, where 2 questions were asked. The first concerned the degree of their activity during distance learning. The second concerned the reasons that influenced the degree of educational activity during the study period. The respondents' attitude to the problem under study was assessed using an ordinal rating scale, otherwise called the Likert scale and a nominal scale.

As mentioned above, this technique involves the use of a scale that allows to identify the attitude of respondents to the problem under study, according to which they assess the level of achievements in the educational activities of students. There are various modifications of measuring scales, which include from two to seven points of estimation [10].

The use of Google Forms is possible for various purposes, for example, as a task for preparing for work, self-checking, as well as for conducting surveys. Results Section III. The modern contexts of digital transformation and improvement of the

educational process of the survey formed the basis for a practical study of the student's scientific project on a lexical topic, which made it possible to use fine-tuning statistics, as well as to reach a large number of respondents. To conduct a survey, the teacher will need to send invitation links to students and choose a method for processing materials. Remote survey allows you not to limit participants in time and place. The teacher is given the opportunity to issue a survey or assignment using the tools of replacing the background, inserting illustrations, explanations of correct and incorrect answers. In this case, even performing the task remotely remains personally oriented for the student. The teacher has at his disposal all the tools for automatic processing of the received data – from diagrams to comparative tables. The complexity of organizing distance learning is also related to the level of self-discipline of students. Researchers observe a significant drop in motivation and an increase in the influence of distractions among students, the same process occurs in schoolchildren [11]. The possibility of conducting an unconventional, or rather, a playful form of the lesson increases students' motivation and desire to actively participate in the educational process. Therefore, it is worth mentioning another tool that has become indispensable when conducting final lessons in Russian literature. Creating a quiz using a website myquiz.ru using various options for setting up and graphical shell tasks (start time, response time, question type, use of video, audio and illustrations, etc.) is intended to test students' knowledge in the classroom or as homework. After registering on the website, the teacher creates a quiz using all the proposed tools of precise settings and visual design, sets the time of the demonstration of the question and answer, selects colors and illustrations, attaches a video. After creating a quiz, the teacher shares the link or the quiz number in any way, for example, in the Google feed of the Class. Students register automatically using any account, or manually by specifying their first and last name. Participants are given time to connect. During the quiz, the correct answer and the statistics of the participants' responses are demonstrated. A leaderboard is displayed on which all participants see their position in the response rating. Additional points are given for the speed of responses. The quiz administrator can manage the process: pause, start the game again, etc. At the end of the survey, a rating table will be presented with all participants indicating points and correct answers. A report is generated in the virtual teacher's office, which can be downloaded in Excel spreadsheet format to carefully study the result of each participant [12].

The target audience of the study was represented by a total of 193543 respondents, namely 85946 students, 42903 parents and 64694 teachers from 17 regions.

In terms of territorial affiliation, the number of rural respondents was 100,551, urban – 92,991 people. The contingent of respondents with Kazakh language of instruction was 123 401, with Russian – 34 148 people.

4.2 Results

Big data processing in education has certain goals and is carried out to improve the quality of teaching of a particular student; to optimize management in the formation of classes with the closest possible abilities and style of activity; in order to focus on certain trends in changing knowledge, skills and abilities of students in the learning process; individualization of the learning trajectory in accordance with personal characteristics and needs; optimization of the management of the educational team in order to reduce educational risks for students. In the case of improving the quality of learning, the following can be analyzed: response time, student's reactions when responding, grades given, the style of solving the task, selected educational resources, methods for solving a group of tasks, and much more. When analyzing in real time, the data obtained allows the teacher to focus the student's attention on the optimal learning trajectory for him, and the student to see his typical ineffective actions and ways to optimize them. In the case of management optimization in the formation of classes with the closest possible abilities and style of activity based on an individual learning trajectory, big data allows us to assess the student's abilities not only for individual, but also for group activities [13]. Based on the preferences of each student, it is possible to identify and allocate a zone of cross-interests in order to optimally use it in teaching. As a feedback, the analysis system allows the teacher to recommend a certain group of professions to the student. Among the most well-known tools of this kind are Socrative, Nearpod and Classroom Monitor, which allow the teacher to monitor in real time.

Conduct a survey of key personnel every time there is a malfunction in order to assess the impact of the pandemic on the performance of key functions at the school and system level, such as the implementation of the curriculum, assessments, teacher training, appointment and career development of teachers, the provision of school meals and other services for students, etc. Governing structures in the field of education are often unable

to implement policy measures quickly due to excessive complexity, insufficient coordination between administrative levels, as well as due to norms and rules that slow down implementation in practice - in short, due to the lack of a conducting chain. The elimination of these limitations and the provision of a working conductive chain is extremely important at a time that requires immediate action [14].

Big data plays an important role in the development of the modern information society, representing both structured and unstructured arrays of various information. Structured data, according to J. Kelleher, are characterized by the possibility of

reflection and storage in the form of a table in which each object has an identical structure (set of attributes), for example, demographic data of the population, where each row in the table describes one person and consists of identical attributes (name, age, date of birth, etc.). Information is considered unstructured when each object in the dataset can have its own structure, and it will not necessarily be the same for all objects, for example, natural text arrays (emails and short text messages, etc.), various collections of audio, graphic and video files. Structured data is more suitable for analytical work than unstructured data, and is usually their derivative.

Table 1. One-Dimensional Statistics

One-dimensional statistics						
Questions	N	Average	Srednekv. deviation	Missed		Missed
				Quantity	Percentages	
Do you think you (your child, pupil) were able to be active in distance learning?	217114	2,00	,757	1639	0,7%	2,00
If you (your child, pupil) are active in distance learning what does it matter?	194274	2,33	1,105	24479	11,2%	2,33
a. Little's MCAR criterion (Little's Missing Completely at Random - missing completely arbitrarily): Chi-squared = 37781.215, DF = 2, Significance = ,000						

Summary report on observations						
	Observations					
	Valid		Missed		Total	
	N	Percent	N	Percent	N	Percent
Degree of activity * Reason for activity	193523	88,5%	25230	11,5%	218753	100,0%

Do you think you (your child, pupil) were able to be active in distance learning? * If you are active in distance learning, what does it matter? * respondents						
		If you are active in distance learning, what does it matter?				Total
		Personal control of the teacher	Strengthening the control of adults	It is easier for the student to answer personally	It is more interesting to study	
Do you think you (your child, pupil) were able to be active in distance learning?	High	11,5%	6,0%	7,2%	7,2%	32,0%
	Average	12,4%	14,4%	10,4%	9,8%	47,0%
	Without changes	5,8%	7,2%	4,8%	3,2%	21,0%
Total		29,7%	27,6%	22,4%	20,3%	100,0%

Having checked the data for the normality of the distribution according to the Kolmogorov-Smirnov criterion (n>50) (table 2), it can be seen that the

asymptotic significance is less than 5%, i.e. 0.05. This means that all the signs differ from the normal distribution.

Table 2. Kolmogorov-Smirnov Criterion

Kolmogorov-Smirnov single-selection criterion				
		respondents	Type of school	Language
N		218753	214616	215459
Parameters of the normal distribution, a,b	Average	2,09	1,50	1,43
	Srednekv.deviation	,720	,500	,552
The greatest extreme discrepancies	Absolute	,241	,343	,382
	Positive	,241	,339	,382
	Negative	-,231	-,343	-,249
Criterion statistics		,241	,343	,382
Asymptotic significance (2-sided)		,000 ^c	,000 ^c	,000 ^c
a. The distribution being tested is normal.				
b. Calculated from the data.				
c. Correction of the significance of Lilyfors.				

In our current work, we consider three possible variants of the learning curve, which can significantly affect the levels of education in countries where schools are closed (Fig.1). Each variant is due to a different mechanism that currently affects students. The first option is the simplest transformation, which is caused by a decrease in average learning levels throughout the distribution (blue curve). This is a very likely scenario, despite all the efforts made by school systems to provide distance learning. The change in school time is associated with learning losses. There is also evidence that natural shocks, such as floods, significantly affect learning outcomes in various classes. Children who do not attend school study less, despite the best intentions of distance education and those who study at home [15].

Consider the second option. The learning curve may flatten (or tilt) due to the extremely uneven consequences of the crisis. This is an option in which the children at the top will move forward, while the students at the bottom will lag behind more and more. Even if the virus doesn't care if you are rich or poor, the rich have much more opportunities to mitigate its effects. Wealthier families live in comfortable homes, have high-speed Internet, can hire a private tutor; in addition, well-educated parents can help a child with home schooling. Poor families, especially the extremely poor, often live in slums where there may not even be a radio, let alone the Internet or digital gadgets; there is also no need to talk about resources for hiring a tutor here, and parents are unlikely to cope with their children's homework. At the bottom of the income distribution, there may also be a sharp increase in poverty due to lack of work opportunities or unemployment. In this case, the rich will get ahead, and the poor will fall behind even more [16].

Consider the third option. The curve may change due to dropout. From earlier crises, such as the financial crisis in Asia in 1997– 1998 and the polio pandemic in 1916, we learned that school attendance can plummet due to both demand side factors and supply side factors. As for demand, a sharp drop in income forces families to ask their children to go to work, and they never go back to school. We are particularly concerned about the situation of girls, as they are usually the first to stop attending school. We may face an increase in the number of schools being closed. Governments will run out of cash as the global economic system is struggling. This may lead to the fact that the ministries of Education will be forced to dismiss teachers and close or merge schools. In addition, many countries have expanded the school enrollment of children through private schools with low tuition fees. These schools tend to operate on tiny margins, and we don't know if they will survive this crisis. It will take time to understand how significant the consequences of the COVID-19 crisis will be. But we can't wait that long, we need to start acting now, so we are modeling the impact of this crisis on the learning process. Based on the available data on the impact of crises, as well as on our databases, such as the "Agreed Learning Outcomes" and the "Data Set on Learning in Poverty", we will model the process of curve development in case appropriate actions are not taken. We will look at different scenarios, such as those presented above, and how different strategies can help in reducing the impact of the pandemic on education [17].

For signs that differ from normal, we output frequency statistics (Table-2) with a median. And build a graph of the normal distribution for the Types of respondents, by language and school status.

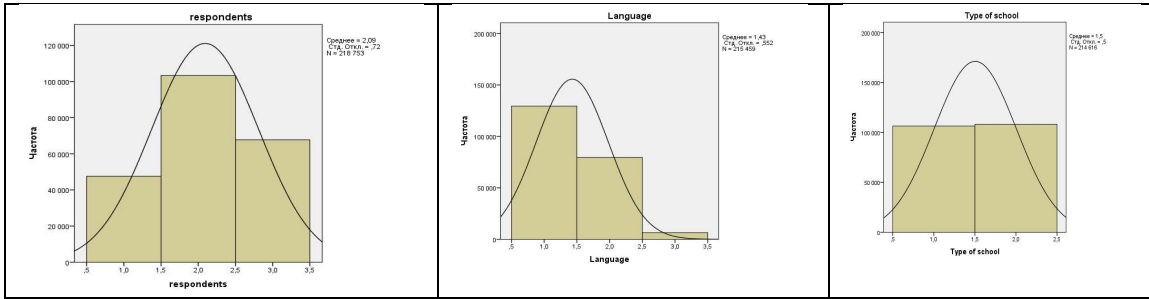


Figure 1. Curves Of Normal Distributions

Chi-square test and standardized residue analysis

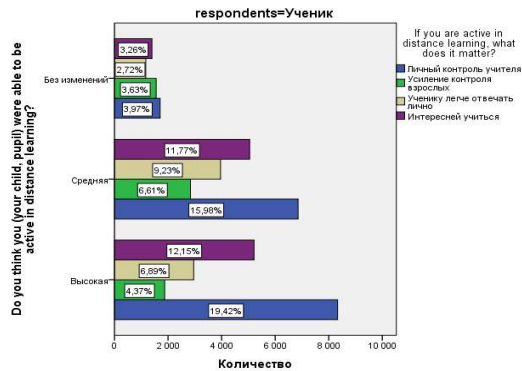
This test was conducted between the question of the degree of change in student activity and the reasons for this activity, as well as between these two

questions and the category of respondents (students, parents and teachers - Table-3). All tests showed a statistically significant result, which indicates that there is a relationship between the question of the benefits of distance learning.

Table-3. Pearson Chi-Square Test Results

Chi-square criteria					
		Value of the article	ст.св.	The asymptotic significance (2-sided)	
Respondent type.					
	Students	Pearson's Chi-square	5178,229 ^c	6	,000
		Likelihood relations	5350,339	6	,000
		Linear-linear connection	1712,674	1	,000
Number of allowed observations		85946			
Parents	Pearson's Chi-square	1350,035 ^b	6	,000	
	Likelihood relations	1273,429	6	,000	
	Linear-linear connection	66,534	1	,000	
	Number of allowed observations	42903			
Teachers	Pearson's Chi-square	897,423 ^d	6	,000	
	Likelihood relations	900,175	6	,000	
	Linear-linear connection	224,753	1	,000	
	Number of allowed observations	64694			

To see which of the response categories contribute to the relationship between variables, it is necessary to calculate standardized residuals. These residuals are an indicator of how much the observed and expected frequencies deviate from each other. Fields with a higher standardized remainder make a more significant contribution to the numerical value of the chi-square criterion and, consequently, to a significant result (Fig.2). The sign that has a standardized remainder is important: plus, means an overabundance, and minus a deficit in response frequency



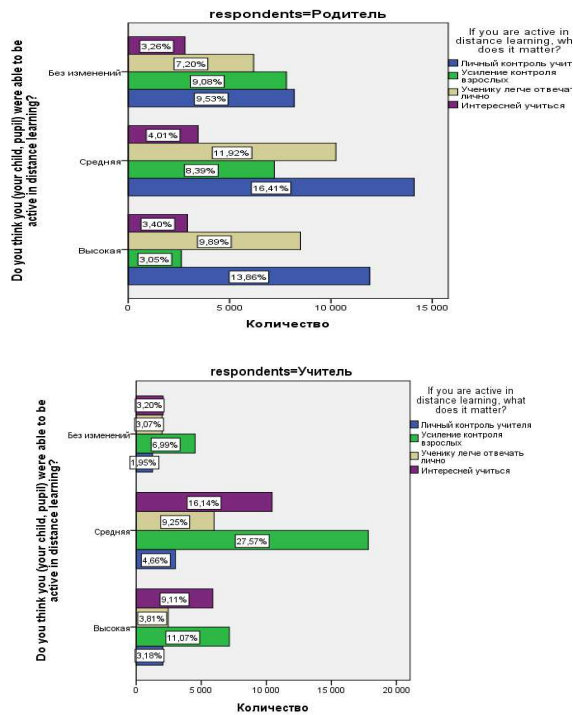


Figure 2. Standardized Residues Of Cross-Tabulation Of The Degree Of Activity And Its Causes

From the graph of standardized balances, it can be concluded that for the students who showed the most activity, the main reason is the factor of the teacher checking each task (blue column on the diagram). For this category of students, less important reasons are increased control by adults and the need to speak publicly in front of the class. Another important factor was that parents noticed an increase in children's interest in learning, due to the fact that they can study not only from textbooks, but also use numerous other resources [18].

For students who showed an average and low level of activity during distance learning, the main factor was increased control by adults, the opportunity not to speak publicly in front of the class. Of little importance for these two categories of students is the verification of each assignment by the teacher. And such a factor as an increase in interest in studying, due to the fact that it is possible to study not only according to textbooks, in a group of students who showed the weakest activity during distance learning – has a weak value for them.

Now let's move on to analyzing the relationship between the target questions and the type of respondents.

The following Figure 3 shows the standardized residuals between the degree of activity of students during distance learning and the type of respondents.

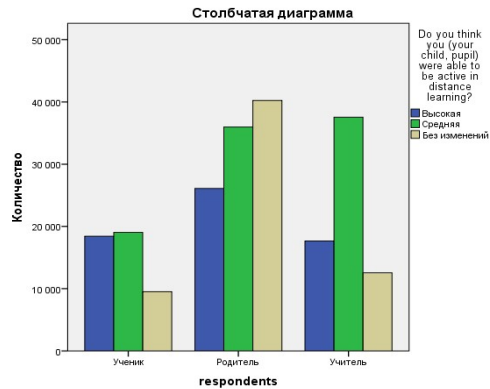


Figure 3. Standardized Remnants Of High-Activity Cross-Tabulation And Its Causes By Type Of Respondents

From this graph, it can be concluded that teachers consider the main reasons for the change in activity to be increased control by adults and an increase in interest in learning, due to the work of students not only with textbooks. To the least extent, teachers consider the main reason that each task of students is checked by the teacher. The opposite opinion is held by parents, who believe that the main reason for the change in the activity of their children is the implementation of the verification of each assignment of students by the teacher. The students themselves are not inclined to believe that the most important reason for the change in their activity is the increased influence from adults.

Conjugacy measures

The existence of a relationship between questions about the degree of change in the activity of students in distance learning and the causes of this activity, as well as between these questions and the categories of respondents, can be assessed by calculating the Kramer conjugacy measure V, which is an analogue of the correlation coefficient developed for nominal variables and has a numerical value [19].

The purpose of (table-4) Kramer's tightness indicator V is to obtain a visual interpretation of the relationship between variables in the form of a coefficient ranging from 0 to 1 and taking a zero value for uncorrelated variables and a single value for strictly related variables.

Table 4. Kramer's Conjugacy Measures V

		Activity	Reasons for activity	Type of respondents
1	Activity	1		
	Significance			
2	Reasons for activity	,127	1	
	Significance	,000***		
3	Type of respondents	,155	,307	1
	Significance	,000***	,000***	

***. Chi-squared statistics are significant at the level ,001.

All values of the conjugacy measures showed a statistically significant result with a possible error level of the first kind equal to 0.001. At the same time, the greatest correlation is observed between the causes of changes in student activity and the type of respondents, which indicates that different types of respondents evaluate the causes of changes in student activity during distance learning differently.

Multinomial logistic regression

Multinomial logistic regression is a general case of a logistic regression model in which a dependent variable has more than two categories. The dependent variable in the multinomial logistic regression model can be measured on an ordinal or nominal scale. The use of logistic regression allows us to try to determine the presence of a causal effect

of the causes of changes in student activity and the type of respondents on the assessment of the degree of change in student activity [20]. If the regression coefficients show a statistically significant result, accordingly, it can be concluded that they really affect the value that the dependent variable can take.

The following table shows the calculated regression coefficients and the statistical significance of the Wald criterion. If the criterion is statistically significant, then its significance will be less than 0.05. In fact, the variables showed an even more significant result than the selected significance level of 0.05 (table-5)

Table 5. Coefficients Of Multinomial Logistic Regression

		Parameter estimates							
Degree of activity ^a		b	Standard error	Wald	St. St.	Significance	Exp (B)	95% confidence interval for Exp(B)	
								Lower bound	Upper bound
High	Free member	1,010	,018	3027,196	1	,000			
	[Reason=1]	,424	,020	447,654	1	,000	1,528	1,469	1,589
	[Reason=2]	-,807	,020	1620,664	1	,000	,446	,429	,464
	[Reason=3]	-,613	,022	755,571	1	,000	,542	,519	,566
	[Reason=4]	0 ^b	.	.	0
	[Type=1]	-,918	,017	2930,057	1	,000	,399	,386	,413
	[Type=2]	,234	,020	130,598	1	,000	1,263	1,214	1,315
	[Type=3]	0 ^b	.	.	0
Average	Free member	1,528	,017	7843,814	1	,000			
	[Reason=1]	,221	,019	130,375	1	,000	1,247	1,201	1,296
	[Reason=2]	-,325	,018	314,308	1	,000	,722	,697	,749
	[Reason=3]	-,276	,021	179,650	1	,000	,759	,729	,790
	[Reason=4]	0 ^b	.	.	0
	[Type=1]	-1,155	,015	5799,516	1	,000	,315	,306	,324
	[Type=2]	-,307	,019	250,629	1	,000	,735	,708	,764
	[Type=3]	0 ^b	.	.	0

a. Reference category: Unchanged.

b. This parameter is set to zero because it is redundant.

The dependent variable is an assessment of the degree of change in the activity of students during distance learning. This is the ordinal type of variables. This variable has 3 categories that characterize the degree of change in activity (1 – high, 2 – medium, 3 – low). To predict the first two categories of the dependent variable, separate models are calculated, with their own regression coefficients. The third category of the dependent variable is equated to zero and is considered the reference category with which the first two categories are compared. Among the independent or influencing variables, the latter category is also selected as a reference and is equated to zero [21].

All categories of responses of independent variables showed statistically significant results. By the sign that has a regression coefficient for each category of responses, you can determine the nature of the relationship.

The first category of response, which characterizes the greatest degree of change in activity, is positively influenced by type 1 (student). This indicates that it is the students who highly appreciate the degree of change in their activity, compared with teachers, who are the reference category. A negative sign with a regression coefficient, such as with type 2 (parents), indicates that they are not inclined to believe that children have become very active during distance learning. This conclusion can be supported by a graph of standardized cross-tabulation balances of the degree of activity and the category of respondents, which shows that it is among parents that the most common opinion is that changes in activity have been achieved only slightly [22].

Thus, using 3 methods of statistical analysis, it was possible to prove the working hypotheses put forward in this paper.

The pandemic caused by COVID-19 has created many problems, one of which was the massive transition to distance learning. Training based on the use of modern information and telecommunication technologies is by no means an innovation. In addition, the development of distance education was recognized as one of the directions of UNESCO's main educational programs. Despite this, the higher school faced the need to promptly solve a number of tasks and difficulties, not only of an organizational and technical nature, but also related to the communicative interaction of participants in the educational process of the university. Such interaction involves two main communication

options: text correspondence and participation in webinars, online lectures.

The analysis of the features of the teacher's work in the distance format showed that, as with traditional, full-time training, teaching activities should be based on knowledge of their subject, possession of various educational methods and learning strategies. But at the same time, the online format requires the teacher to have a specific ability to own a virtual audience and confident skills of inter-active communication with students [23].

5. DISCUSSION

The transition of educational organizations to emergency distance learning in connection with the spread of coronavirus infection has become a stress test especially for schools. In this regard, the need to study the perception of the educational process by students in the conditions of distance learning has become more urgent.

As the results of our research have shown, the situation of an emergency transition to distance learning has significantly changed the usual forms of activities carried out by students during their academic activities [24].

The literature on school closures also provides some insight into the potential consequences of school closures due to COVID-19, especially given that such closures occur unexpectedly and disrupt class schedules. Although school closures due to inclement weather or natural disasters occur for a shorter time, this is analogous to school closures due to COVID-19. If it were not for weather events or a natural disaster, there would be classes in schools, and most students would be taught as usual. Hansen (2011) found that every day of school cancellations due to snow in Colorado reduced math performance in 8th grade by an amount from 0.013 to 0.039 CO, and the impact of snow days in Maryland ranged from 0.013 to 0.016 CO. [25]

Participants were asked to express the degree of agreement with the statements, the level of educational activity that was achieved in remote learning during the pandemic. It turned out that the reason "Increased control by adults" was chosen by a quarter of teachers (27.5%) and only a third of parents (32%).

According to a survey of parents of schoolchildren in Kazakhstan, the distance learning time of children has become more active due to increased personal control of teachers (16%), and there is a discrepancy

in the opinions of teachers and parents on the reasons for school classes halved during the school closure period:

Many authors [14-15] agreed that most of the students, teachers and other participants in the educational process did not have time to acquire significant experience for an emergency transition to a distance format.

In this study, the factors of the influence of distance learning on the educational activity of Kazakhstani schoolchildren were studied. Adult supervision, individual work with the teacher and aspects that could affect the learning process were considered as reasons. Compared with the opinions of other participants in education, students believe that they showed the greatest activity in their studies during distance learning.

Thus, the emergency transition to distance learning as a tool for the combination of material presentation and interaction between student and teacher during the quarantine regime had a positive impact on Kazakhstani schoolchildren.

With regard to the implications for research, it is recommended that qualitative research be conducted among students and their stakeholders (e.g. parents, teachers, and policy makers) to examine how the factors identified in this study have affected students' online learning. Further research is also needed to explore other potential factors, such as teaching design and family issues, as well as follow-up of students to understand the associations suggested in this study.

6. CONCLUSIONS

Currently, the full-fledged application of big data technology in the field of education is difficult due to a number of circumstances.

Firstly, a special electronic environment that generates big data is in its infancy in many universities. In this regard, big data of educational organizations does not yet represent "huge amounts of information of considerable diversity" in the sense of classical big data. Essentially, we are dealing with big data in miniature. Increasing the pace of formation and improvement of the electronic information and educational environment will accelerate the full use of big data.

Secondly, teachers as developers of electronic educational resources are often aware of the lack of technical knowledge and skills necessary for effective online interaction with students. Ways to

increase the level of digital literacy are diverse: from self-education to various advanced training courses.

Thirdly, universities and schools often do not have horizontally scalable software tools that allow them to efficiently process huge amounts of information. The technical means and software of an educational organization are capable of processing big data in a test mode rather than on an ongoing basis. As a result, only individual management decisions will be made using this advanced technology. The university's access to the big data analytics platform is seen as possible due to the conclusion of a partnership agreement with a company of the appropriate profile. It is obvious that big data with consumer value is important for an educational organization — the ability to bring informative benefits, including the final socio-economic effect, as well as having an applied nature.

ACKNOWLEDGMENTS

Doctoral student Mukhiyadin Ainur expresses gratitude to Associate Professor Mukasheva Manargul for her scientific guidance and assistance in her research work.

REFERENCES

- [1] Ananchenkova, P. I. Dysfunctionality of education in the conditions of the COVID-19 pandemic / P. I. Ananchenkova, N. M. Novikova // "Alma mater" (Bulletin of the Higher School). - 2020. - No. 11. - pp. 7-11.
- [2] Batura, M. P. Introduction of network learning in the Republic of Belarus / M. P. Batura, B. V. Nikulshin // Higher School. - 2017. — No. 1. — pp. 7-11
- [3] Golyakova I. V. Distance learning in the conditions of COVID - restrictions through the eyes of teachers and students of universities / I. V. Golyakova // Higher school. - 2021.—No. 5. - pp. 3-7.
- [4] Ivanov D. N. Distance education: challenges of mass transition to online learning / D. N. Ivanov // Alma mater (Bulletin of the Higher School). - 2021. - No. 7. - pp. 55-58.
- [5] Leskevich S. G. Distance learning: from the experience of forced use during the coronavirus pandemic / S. G. Leskevich, K. L. Leskevich, O. G. Secheiko // Higher School. - 2020.- No. 6. - pp. 3-5.
- [6] Lukashchuk V. I. Problems of digitalization of higher education in Russia / V. I. Lukashchuk // Alma mater (Bulletin of Higher School). - 2021.—No. 7. - pp. 24-28.

- [7] Gvozdenko Yu. V., Ishchenko A. A., Pilipenko A.V. Big data in the education system // International Student Scientific Bulletin: electronic scientific journal. — 2019. — Issue 5. — Part 1. — URL: <http://eduherald.ru/ru/article/view?id=19731>
- [8] Zenin S. S., Kuteynikov D. L., Yapyrntsev I. M. Big data in the legislative process // Actual problems of Russian law. — 2020. — Vol. 15. — No. 9. — pp. 19-31. — DOI: 10.17803/1994-1471.2020.118.9.019-031 .
- [9] Mayer-Schoenberger V., Kukier K. Big data. The revolution that will change the way we live, work and think / translated from English by I. Gaidyuk. — M. : Mann, Ivanov and Ferber, 2014. — 240 p.
- [10] Alonso V., Arranz O. G. Big Data & eLearning: A binomial to the future of the knowledge society // International Journal of Interactive Multimedia and Artificial Intelligence (IJIMAI). — 2016. — Iss. 3 (6). — P. 29–33. — DOI: 10.9781/ijimai.2016.364.
- [11] Javidi G., Rajabion L., Sheybani E. Educational Data Mining and Learning Analytics: Overview of Benefits and Challenges // International Conference on Computational Science and Computational Intelligence (CSCI). — 2017. — P. 1102–1107. — DOI: 10.1109/CSCI.2017.360.
- [12] Saptarshi R. Big Data in Education // Gravity, the Great Lakes Magazine. — 2016. — Iss. 20. — P. 8–10. — URL: https://www.researchgate.net/publication/259443597_BIG_DATA_IN_EDUCATION.
- [13] Vaitsis C., Hervatis V., Zary N. Introduction to big data in education and its contribution to the quality improvement processes // Big Data on Real-World Applications / S. V. Soto, J. M. Luna & A. Cano (Ed.). — London : In Tech, 2016. — P. 41–63. — DOI: 10.5772/63896
- [14] Ministry of Education and Science of the Republic of Kazakhstan National Academy of Education named after I. Altynsarin, Results of the monitoring of distance learning in secondary education, part 1 students, Nur-Sultan 2020
- [15] Ministry of Education and Science of the Republic of Kazakhstan National Academy of Education named after I. Altynsarin, Results of the monitoring of distance learning in secondary education, part 1 pedagogical, Nur-Sultan 2020
- [16] Mukasheva, M., Omirzakova, A. Computational thinking assessment at primary school in the context of learning programming (2021) World Journal on Educational Technology: Current Issues, 13 (3), pp. 336-353
- [17] Mukasheva, M., Chorosova, O., Zhilbayev, Z., Payevskaya, Y. Integrated approach to the development and implementation of distance courses for school computer science teachers (2020) 14th IEEE International Conference on Application of Information and Communication Technologies, AICT 2020 - Proceedings, статья № 9368817
- [18] Zhilbayev, Z.O., Syrymbetova, L.S., Mukasheva, M.Y., Zhetpisbayeva, B.A., Smagulova, G.T. Promotion of trilingual education in kazakhstan schools: Online monitoring results (2019) Journal of Siberian Federal University - Humanities and Social Sciences, 12 (2), pp. 285-30
- [19] Golenko, Y., Ismailova, A., Shaushenova, A., Mutalova, Z., Dossalyanov, D., Ainagulova, A., Naizagarayeva, A. Implementation of machine learning models to determine the appropriate model for protein function prediction (2022) Eastern-European Journal of Enterprise Technologies, 5 (4-119), pp. 42-49.
- [20] Bayegizova, A., Murzabekova, G., Ismailova, A., Aitimova, U., Mukhanova, A., Beldeubayeva, Z., Ainagulova, A., Naizagarayeva, A. Effectiveness of the use of algorithms and methods of artificial technologies for sign language recognition for people with disabilities (2022) Eastern-European Journal of Enterprise Technologies, 4 (2-118), pp. 25-31.
- [21] Sandugash, S., Jamalbek, T., Madina, S., Akbota, Y., Ainur, A. Building a Standard Model of an Information System for Working with Documents on Scientific and Educational Activities (2021) International Journal of Advanced Computer Science and Applications, 12 (9), pp. 445-455.
- [22] Weindorf-Sysoeva M.E. Methods of distance learning: a textbook for universities / M.E. Weindorf-Sysoeva, T.S. Gryaznova, V.A. Shitova. — M.: Yurayt, 2018. — 194 p.
- [23] Gamanovich A.I. Lumbosacral pain syndrome, comparison of views on the problem / A.I. Gamanovich, B.V. Drivotinov // Problems of health and ecology. — 2017. — № 2 (52). — Pp. 43-48.
- [24] Schenkova I.P. Problems of distance learning in the discipline "physical culture" // International Journal of Humanities and Natural Sciences. — 2020. — Vol. 5-3 (44). — pp. 190-193.
- [25] Tailakov U.N. Unified Information and educational school space as a factor of improving the quality of educational processes // Young scientist. — 2013. — № 5.